

19725735

Reserva

# Análisis Sintáctico Eficiente de Gramáticas de Adjunción e Inserción de Árboles

Departamento de Lenguajes y Sistemas Informáticos



UNIVERSIDAD  
de SEVILLA

Tesis  

---

38

Memoria de la tesis doctoral dirigida por Dr. Víctor Jesús Díaz Madrigal  
y desarrollada por D. Vicente Carrillo Montero para optar al grado de  
Doctor en Informática por la Universidad de Sevilla

Sevilla, Mayo de 2003



Yo, Vicente Carrillo Montero, con Documento Nacional de Identidad 34.034.248-K,

**DECLARO BAJO JURAMENTO**

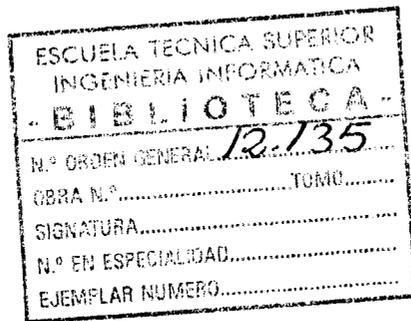
ser el autor del trabajo incluido en la memoria de esta tesis doctoral titulada

Análisis Sintáctico Eficiente de  
Gramáticas de Adjuncción e Inserción  
de Árboles

Tesis  

---

38



*Vicente Carrillo Montero*

Fdo. Vicente Carrillo Montero  
Dpto. de Lenguajes y Sistemas Informáticos  
Universidad de Sevilla  
Área de Lenguajes y Sistemas Informáticos

Sevilla, Mayo de 2003



Dr. D. Víctor Jesús Díaz Madrigal  
Profesor Titular de Escuela Universitaria  
Departamento de Lenguajes y Sistemas Informáticos  
Universidad de Sevilla

CERTIFICA:

Que la memoria titulada *Análisis sintáctico eficiente de gramáticas de adjunción e inserción de árboles* ha sido realizada por D. Vicente Carrillo Montero bajo mi dirección en el Departamento de Lenguajes y Sistemas Informáticos de la Universidad de Sevilla y concluye la Tesis que presenta para optar al grado de Doctor en Informática.



Sevilla, 20 de mayo de 2003

A handwritten signature in black ink, consisting of several loops and a long horizontal stroke.

Dr. D. Víctor Jesús Díaz Madrigal  
Director de la Tesis Doctoral



*A mis padres y Ángela*



# Agradecimientos

A Víctor Díaz, porque juntos iniciamos el camino de la investigación en los formalismos gramaticales basados en reescritura de árboles. Este trabajo es el fruto de esta andadura.

A Miguel Toro por sus consejos y su continuo apoyo.

A Miguel Ángel Alonso por compartir conmigo sus extensos conocimientos sobre gramáticas de adjunción de árboles, sin los cuales hubiera sido imposible el desarrollo de partes fundamentales de esta memoria.

A mis compañeros del grupo de investigación ITALICA del Departamento de Lenguajes y Sistemas Informáticos, especialmente a Francisco José Galán y José Antonio Troyano, por estar siempre que los he necesitado.

A Ángela por su cariño, paciencia y confianza. A mis padres y mi hermana por creer en mí y apoyarme en todo momento. Y sobre todo, a mi pequeña Paula por existir.

El trabajo descrito en esta tesis ha sido parcialmente financiado por el Departamento de Lenguajes y Sistemas Informáticos de la Universidad de Sevilla, la Xunta de Galicia (proyecto PGI1DT99XI10502B) y el Ministerio de Ciencia y Tecnología (proyectos FIT-150500-2002-416 y TIC 2000-2052-E).



# Resumen

Las gramáticas de adjunción de árboles (TAG, *Tree Adjoining Grammars*) es un formalismo gramatical que presenta características muy adecuadas para la descripción de lenguajes naturales. La mayoría de los analizadores tabulares definidos para ellas son adaptaciones de analizadores tabulares conocidos para gramáticas independientes del contexto (CFG, *Context Free Grammars*), sin embargo, el coste computacional teórico en el caso peor de estos algoritmos es bastante más elevado.

Las gramáticas de inserción de árboles (TIG, *Tree Insertion Grammars*) es un compromiso entre CFG y TAG. Las TIG se pueden analizar con el mismo coste que las CFG, manteniendo las propiedades lingüísticas fundamentales de las TAG. Como contrapartida, su potencia expresiva queda reducida a lenguajes independientes del contexto. La importancia de este formalismo radica en que la mayor parte de las estructuras que definen los lenguajes naturales se adaptan a él. Al igual que en las TAG, los analizadores tabulares definidos para TIG están basados en los de las CFG, aunque el número de ellos es significativamente menor que el de los analizadores para TAG, debido fundamentalmente a la corta historia con que cuentan las TIG.

Los esquemas de análisis sintáctico establecen un método general para la descripción de algoritmos de análisis sintáctico y constituyen la formalización de estudios previos sobre analizadores deductivos. Permiten definir, analizar y relacionar algoritmos de análisis, y se han empleado satisfactoriamente con un gran número de analizadores para CFG. Recientemente se han desarrollado trabajos que usan los esquemas de análisis para describir y relacionar analizadores tabulares para TAG, concretamente, se ha definido un conjunto de esquemas de análisis cuyo algoritmo central está basado en el algoritmo de Earley para CFG y se han mostrado las relaciones formales existentes entre ellos.

En esta memoria se presenta un grupo de nuevos esquemas de análisis para TAG que, aunque mantienen las cotas de complejidad computacional mencionadas, mejoran el comportamiento práctico de los analizadores tabulares clásicos. Se emplean dos enfoques: el primero se basa en la aplicación de filtros *Left Corner* a los analizadores tipo Earley para TAG definidos en la literatura; el segundo consiste en la construcción de analizadores que funcionen de forma dinámica, comportándose como un analizador para TIG cuando analiza el núcleo de una gramática TAG que responde a las restricciones del formalismo TIG, y como un analizador TAG en el resto de los casos.

Dado que el número de analizadores definidos para TIG es escaso, presentamos un conjunto de esquemas de análisis para este formalismo con un doble objetivo: crear una red de analizadores para TIG similar a la existente para CFG, e integrarlos con los de TAG para construir analizadores conforme al segundo enfoque descrito.

Aprovechando las posibilidades que nos brindan los esquemas de análisis sintáctico,



probamos las relaciones formales entre los nuevos esquemas para TIG y TAG. También mostramos un método general de prueba de corrección para esquemas de análisis sintáctico de TAG, probando la corrección de los nuevos analizadores tipo *Left Corner* presentados en la memoria.

# Índice General

<b>I</b>	<b>Preliminares</b>	<b>1</b>
<b>1</b>	<b>Introducción</b>	<b>3</b>
1.1	Formalismos gramaticales . . . . .	3
1.2	Análisis sintáctico . . . . .	6
1.2.1	Análisis sintáctico de TAG . . . . .	9
1.3	Objetivos . . . . .	10
1.4	Estructura de la memoria y publicaciones asociadas . . . . .	12
<b>2</b>	<b>Enfoque deductivo en el análisis sintáctico</b>	<b>17</b>
2.1	Gramáticas incontextuales . . . . .	17
2.2	Sistemas deductivos . . . . .	20
2.2.1	Corrección de los sistemas deductivos . . . . .	22
2.3	Esquemas de análisis basados en ítems . . . . .	23
2.3.1	Corrección de los esquemas de análisis . . . . .	25
2.4	Ejemplos de esquemas de análisis . . . . .	26
2.5	Complejidad . . . . .	31
2.6	Relaciones entre esquemas de análisis . . . . .	33
<b>II</b>	<b>Gramáticas de adjunción de árboles</b>	<b>37</b>
<b>3</b>	<b>Gramáticas de adjunción de árboles</b>	<b>39</b>
3.1	Definición . . . . .	39
3.2	Propiedades lingüísticas y computacionales . . . . .	42
3.3	Formalismos relacionados . . . . .	43
<b>4</b>	<b>Algoritmos de análisis sintáctico para TAG</b>	<b>45</b>
4.1	Esquemas de análisis sintáctico y TAG . . . . .	46
4.2	Esquema tipo CYK . . . . .	50
4.3	Esquema tipo Earley ascendente . . . . .	52
4.4	Esquema tipo de Vreught y Honig . . . . .	54



4.5	Esquemas tipo Earley . . . . .	56
4.5.1	Esquema tipo Earley con la propiedad del prefijo válido . . . . .	59
<b>5</b>	<b>Algoritmos de análisis basados en Left Corner para TAG</b>	<b>63</b>
5.1	Esquema tipo Left Corner ascendente . . . . .	63
5.1.1	Resultados experimentales . . . . .	66
5.2	La relación Left Corner en TAG . . . . .	68
5.3	Esquemas tipo Left Corner con ítems predictivos . . . . .	70
5.3.1	Esquema tipo Left Corner con la propiedad del prefijo válido . . . . .	75
5.4	Esquema tipo Left Corner con ítems simplificados . . . . .	79
5.5	Esquema tipo Left Corner . . . . .	82
5.5.1	Mejorando la eficiencia del esquema tipo Left Corner . . . . .	83
5.5.2	Resultados experimentales . . . . .	86
<b>6</b>	<b>Ampliando la red de analizadores para TAG</b>	<b>89</b>
6.1	Relaciones entre esquemas tipo Left Corner para TAG . . . . .	89
6.2	Integrando los esquemas tipo Left Corner en la red de analizadores para TAG	104
<b>III</b>	<b>Gramáticas de inserción de árboles</b>	<b>113</b>
<b>7</b>	<b>Gramáticas de inserción de árboles</b>	<b>115</b>
7.1	Gramáticas de inserción de árboles . . . . .	115
7.1.1	Notación . . . . .	119
7.2	Relaciones entre TIG, CFG y TAG . . . . .	120
7.3	Un analizador tipo Earley para TIG . . . . .	123
7.4	Revisando el analizador tipo Earley para TIG . . . . .	127
7.4.1	Representación plana de los árboles elementales . . . . .	127
7.4.2	Analizador tipo Earley para TIG con notación plana . . . . .	129
<b>8</b>	<b>Algoritmos de análisis sintáctico para TIG</b>	<b>131</b>
8.1	Esquema tipo CYK . . . . .	131
8.2	Esquema tipo Earley ascendente . . . . .	133
8.2.1	Una variante del esquema tipo Earley ascendente . . . . .	135
8.3	Esquema tipo de Vreught y Honig . . . . .	137
8.4	Esquema tipo Earley . . . . .	140
8.5	Red de analizadores para TIG . . . . .	142
<b>9</b>	<b>Algoritmos de análisis basados en Left Corner para TIG</b>	<b>159</b>
9.1	Esquema tipo Left Corner ascendente . . . . .	159
9.2	La relación Left Corner en TIG . . . . .	162
9.3	Esquema tipo Left Corner con ítems predictivos . . . . .	162

9.4	Esquema tipo Left Corner con ítems simplificados . . . . .	167
9.5	Esquema tipo Left Corner . . . . .	171
9.6	Ampliando la red de analizadores para TIG . . . . .	173
9.6.1	Relaciones entre esquemas tipo Left Corner para TIG . . . . .	174
9.6.2	Integrando los esquemas tipo Left Corner en la red de analizadores para TIG . . . . .	185
<b>IV</b>	<b>Análisis sintáctico de TAG dirigido por el núcleo TIG</b>	<b>195</b>
<b>10</b>	<b>Integrando analizadores sintácticos para TIG y TAG</b>	<b>197</b>
10.1	Convenciones y notación . . . . .	198
10.2	Esquema tipo CYK . . . . .	200
10.2.1	Relación entre los esquemas tipo CYK para TAG . . . . .	203
10.3	Esquema tipo Earley ascendente . . . . .	208
10.4	Esquema tipo de Vreught y Honig . . . . .	211
10.5	Esquema tipo Earley . . . . .	214
10.5.1	Resultados experimentales . . . . .	218
<b>11</b>	<b>Conclusión</b>	<b>219</b>
<b>V</b>	<b>Apéndices</b>	<b>221</b>
<b>A</b>	<b>Pruebas de corrección de esquemas para TAG</b>	<b>223</b>
A.1	Prueba de corrección del esquema <b>buLC</b> . . . . .	223
A.2	Prueba de corrección del esquema <b>pLC</b> . . . . .	229
A.3	Prueba de corrección del esquema <b>sLC</b> . . . . .	241
<b>B</b>	<b>Adaptando los esquemas para TIG</b>	<b>243</b>
B.1	Esquema tipo CYK . . . . .	243
B.2	Esquema tipo Earley ascendente . . . . .	245
B.3	Esquema tipo de Vreught y Honig . . . . .	246
B.4	Esquema tipo Earley . . . . .	247
	<b>Bibliografía</b>	<b>249</b>



# Índice de Figuras

3.1	<i>Operación de adjunción</i>	41
3.2	<i>Operación de sustitución</i>	42
3.3	<i>Árbol de derivación</i>	43
3.4	<i>Dependencias cruzadas mediante operación de adjunción</i>	44
4.1	<i>Red de analizadores para TAG</i>	46
4.2	<i>Representación multicapa de una gramática TAG</i>	48
4.3	<i>Ítems del esquema CYK</i>	51
4.4	<i>Ítems del esquema buE</i>	53
4.5	<i>Ítems del esquema dVH</i>	55
4.6	<i>Ítems del esquema Nederhof</i>	59
5.1	<i>Comparativa entre los esquemas tipo Earley y Left Corner ascendentes (1)</i>	67
5.2	<i>Comparativa entre los esquemas tipo Earley y Left Corner ascendentes (2)</i>	68
5.3	<i>Conjunto de árboles elementales para experimentos</i>	69
5.4	<i>Ítems left corner del esquema pLC</i>	71
5.5	<i>Ítems left corner del esquema pLC<sup>VPP</sup></i>	76
5.6	<i>Comparativa entre los esquemas tipo Earley y Left Corner</i>	86
6.1	<i>Red ampliada de analizadores para TAG</i>	110
7.1	<i>Adjunción derecha</i>	117
7.2	<i>Adjunción izquierda</i>	118
7.3	<i>Adjunción simultánea</i>	119
7.4	<i>Adjunción wrapping</i>	120
7.5	<i>Adjunción de un auxiliar derecho en la espina de un auxiliar izquierdo</i>	121
7.6	<i>Complejidad de árboles y lenguajes en TIG, TAG, CFG y TSG</i>	123
8.1	<i>Red de analizadores para TIG</i>	158
9.1	<i>Red ampliada de analizadores para TIG</i>	193



10.1	<i>Distribución de los árboles auxiliares en XTAG</i>	198
10.2	<i>Adjunción de dos árboles auxiliares en un nodo</i>	200
10.3	<i>Comparativa entre analizadores tipo Earley para TAG</i>	218

Parte I

Preliminares



# Capítulo 1

## Introducción

*Lengua: "Sistema de comunicación verbal y casi siempre escrito, propio de una comunidad humana", "Sistema lingüístico considerado en su estructura".*

*(Diccionario de la lengua española, vigésima segunda edición)*

### 1.1 Formalismos gramaticales

Aunque el Diccionario de la lengua española incluye acepciones específicas en el contexto de la informática, el término *lenguaje* se utiliza en ámbitos muy diversos. Por ejemplo, los lenguajes formales se emplean en matemáticas, los lenguajes lógicos en lógica y los lenguajes de programación en informática. Para distinguirlos del resto, a los lenguajes usados como medio de comunicación oral y escrita entre humanos se les denomina lenguajes naturales, y la lingüística general es la disciplina que se encarga de su estudio.

La *sintaxis* constituye el núcleo fundamental de todos los campos de estudio cubiertos por la lingüística. Aunque no hay un consenso respecto al objeto de estudio de la sintaxis (visión descriptiva [34] vs visión prescriptiva [61]), parece generalmente aceptado que su función es establecer las reglas de construcción de sentencias para un lenguaje dado. De manera que un lenguaje se puede considerar como un conjunto cuyos elementos son especificables mediante reglas sintácticas. El conjunto de expresiones lingüísticas que se pueden componer en un lenguaje natural no es finito, lo que implica que todos los lenguajes naturales son conjuntos infinitos.

La *lingüística computacional* es un campo de estudio que se ha desarrollado de la intersección entre las disciplinas de lingüística general y el subconjunto de la informática denominado *inteligencia artificial*, y cuyo fin es el estudio de todos los aspectos del procesamiento del lenguaje natural mediante máquinas. Es necesario definir sistemas formales que definan los elementos de los conjuntos infinitos de expresiones lingüísticas y asignen estructuras a cada uno de los elementos de estos conjuntos. Estos sistemas formales son las *gramáticas*. En definitiva, una gramática es una forma simple de representar ciertos aspectos que se cono-

cen de un lenguaje que sea explícita y lo bastante formal para ser "comprendida" por una máquina.

El hito fundamental y punto de arranque para la lingüística computacional son los trabajos de Noam Chomsky [31, 32, 33], que establecen una jerarquía de cuatro tipos de lenguajes en función de su poder expresivo, y que, de menor a mayor expresividad, son los siguientes: lenguajes regulares (RL, Regular Languages), lenguajes independientes del contexto o incontextuales (CFL, Context Free Languages), lenguajes dependientes del contexto (CSL, Context Sensible Languages) y lenguajes sin restricciones o conjuntos recursivamente enumerables. Cada uno de estos tipos de lenguajes presentan formalismos gramaticales asociados que posibilitan la generación de los mismos.

Una importante cuestión que se plantea desde hace más de 40 años es cual es la potencia expresiva que se requiere a un formalismo gramatical para que sea capaz de describir lenguajes naturales. Después de todo este tiempo, lo que parece quedar claro es que los lenguajes incontextuales no cubren todas las construcciones lingüísticas en todos los lenguajes naturales. Un caso característico que pone de manifiesto esta afirmación es una formación básica en un dialecto del alemán hablado en Zürich (Suiza) presentada simultáneamente en [64, 118]. Esta construcción responde a un patrón de la forma

$$NP_a^m \quad NP_d^n \quad V_a^m \quad V_a^m ,$$

que evidentemente no puede ser descrito mediante CFL. En la literatura podemos encontrar otros ejemplos de casos similares, como lo son ciertas construcciones del holandés descritas en [67, 141].

Otra cuestión de gran interés es determinar que características debe poseer un formalismo gramatical para que sea adecuado para describir lenguajes naturales [119, 121]. Entre los criterios más relevantes, podemos citar los tres siguientes:

- *Naturaleza lingüística*

Las personas que deben usar el formalismo gramatical para diseñar las reglas gramaticales que describen un lenguaje natural, en general, son lingüistas. Por tanto, es lógico exigirle al formalismo que presente una notación que les permita expresar las descripciones lingüísticas de manera fácil y flexible.

- *Potencia generadora*

Tal como dijimos previamente, el formalismo debe disponer de la potencia expresiva suficiente para describir lenguajes naturales. En [142, 70] se establecen las propiedades formales necesarias para que un formalismo gramatical sea capaz de representar un lenguaje natural, y que son: (1) capacidad para generar CFL, (2) análisis sintáctico en tiempo polinomial, (3) captura de dependencias anidadas y ciertas dependencias cruzadas, (4) y propiedad de crecimiento constante. A los formalismos que cumplen estas propiedades se les incluye en un grupo situado entre los que generan CFL y

los que generan CSL, y que se denominan gramáticas suavemente dependientes del contexto (MCSG, *Mildly Context-Sensitive Grammars*).

- *Eficiencia computacional*

También es imprescindible que una máquina sea capaz de representar y emplear el formalismo con costes espaciales y temporales razonables.

Numerosos son los formalismos presentados a lo largo de la corta historia de la lingüística computacional y que, en mayor o menor medida, pretenden cumplir los tres criterios que acabamos de enumerar. Las gramáticas incontextuales (CFG, *Context Free Grammars*) [32] fue el primero y más extendido. Una detallada descripción de las propiedades de los CFL se puede encontrar en [4]. Entre sus virtudes podemos encontrar su eficiencia computacional, sin embargo, adolece de una importante falta de expresividad lingüística y potencia generadora. Para aumentar la capacidad de las CFG, se presentaron algunas variantes [101, 66, 130] en las que las reglas se acompañan de condiciones y acciones. Pero estas variantes presentan problemas para mantener y extender cuando se usan para describir lenguajes de amplia cobertura [120]. Otro formalismo muy extendido fue el de redes de transición ampliadas (ATN, *Augmented Transition Networks*) [143], en el que las gramáticas se expresan mediante grafos dirigidos. Al contrario que CFG, tiene una potencia expresiva adecuada, pero presenta el mismo problema de mantenimiento que tienen las variantes sobre CFG nombradas, además de un serio problema para definir algoritmos de análisis sintáctico con diversas estrategias [29].

Un grupo significativo de formalismos gramaticales son los basados en *unificación*, cuyo denominador común es que usan la unificación de *estructuras de rasgos* como operación básica [120]. Dentro de este grupo se presentan dos tendencias, por un lado se encuentran aquellos que surgen motivados por ciertas teorías lingüísticas y, por otro, los que se encuentran diseñados para su uso como herramientas por lingüistas computacionales. Entre los que se encuadran en la primera tendencia, podemos citar las gramáticas léxico funcionales (LFG, *Lexical Functional Grammars*) [73], las gramáticas de estructura de frases generalizadas (GPSG, *Generalized Phrase Structure Grammars*) [91, 53], las gramáticas de estructura de frases dirigidas por núcleo (HPSG, *Head Phrase Structure Grammars*) [92, 93] y las gramáticas categoriales de unificación (UCG, *Unification Categorical Grammars*) [147]. Sin embargo, todas presentan un problema común, la complejidad computacional en el caso peor para el análisis sintáctico puede ser intratable, tal como se pone de manifiesto en [11]. Entre los que siguen la segunda tendencia, podemos citar dos formalismos que se caracterizan por su flexibilidad y eficiencia para el análisis sintáctico y su dificultad para definir gramáticas de lenguajes naturales [117], como las gramáticas de cláusulas definidas (DCG, *Definite Clause Grammars*) [90], que nacen a la luz del lenguaje de programación Prolog, y las gramáticas de unificación funcionales (FUG, *Functional Unification Grammars*) [76].

Entre los formalismos pertenecientes al grupo de MCSG, un capítulo destacado merecen las gramáticas de adjunción de árboles (TAG, *Tree Adjoining Grammars*) [69], ya que se trata

de un formalismo que combina una significativa expresividad lingüística con la potencia generadora necesaria para describir lenguajes naturales, aunque el coste computacional teórico de su análisis sintáctico es elevado. Su relevancia queda de manifiesto por su extendido uso, aún en nuestros días, y los numerosos formalismos derivados de las mismas que se pueden encontrar en la literatura. Otros formalismos MCSG reseñables que tienen la misma capacidad generativa que las TAG son las gramáticas de núcleo (HG, *Head Grammars*) [91, 100], las gramáticas categoriales combinatorias (CCG, *Combinatory Categorical Grammars*) [128] y las gramáticas lineales de índices (LIG, *Linear Indexed Grammars*) [51, 52]. Estas últimas se derivan mediante una restricción sobre una extensión de las CFG, concretamente, las gramáticas de índices (IG, *Indexed Grammars*) [3]. Estos tres formalismos presentan los mismos problemas de coste computacional de análisis que las TAG, pero además padecen de una inferior naturaleza lingüística. En [138] se muestra la *equivalencia débil*<sup>1</sup> que existe entre estos cuatro formalismos.

## 1.2 Análisis sintáctico

Todos los tipos de gramáticas empleados en lingüística computacional, de una forma u otra, disponen de un lenguaje para especificar cada uno de los siguientes elementos [54]:

- *Categorías sintácticas*. Con mucha frecuencia también son denominadas POS (*parts of speech*).
- *Entradas léxicas*. Por tanto, debe tener acceso a un *lexicon*, diccionario o lista de palabras. Un *lexicon*, en su forma más elemental, es una lista de palabras que asocia cada palabra con sus propiedades sintácticas, la más importante de las cuales es su categoría sintáctica.
- *Reglas sintácticas o estructuras de composición básicas*. La forma de las reglas sintácticas pueden variar considerablemente en los distintos formalismos. Las más habituales son las denominadas *reglas de estructura de frase*, que informan de que puede estar compuesta (lado derecho de la regla) una categoría sintáctica (lado izquierdo de la regla). Por ejemplo,  $S \rightarrow NP VP$  indica que una sentencia puede estar formada por un sintagma nominal seguido de un sintagma verbal. Aunque esta forma de reglas es quizás la más extendida, no disponen de la expresividad lingüística adecuada, lo que ha provocado la aparición de formalismos que emplean estructuras de composición básicas más complejas y adecuadas para describir fenómenos lingüísticos. Por ejemplo, las TAG usan árboles como tipo de datos para representar reglas sintácticas.
- *Estructuras sintácticas*. Muestran los distintos componentes de una sentencia y como se combinan las reglas de la gramática para obtener dicha estructura. Todos los

---

<sup>1</sup>La equivalencia débil entre dos formalismos implica que son capaces de generar los mismos lenguajes, pero no las mismas estructuras sintácticas.

formalismos conocidos usan alguna forma de grafo acíclico dirigido como tipo de datos para representar estructuras sintácticas.

Un *analizador sintáctico* es un algoritmo que, dadas una gramática y una cadena (sentencia o palabra), intenta devolver una o más instancias de estructuras sintácticas. A cada una de las estructuras asociadas a una sentencia se le denomina *árbol derivado*. La función de un analizador sintáctico es doble, por un lado determina si una sentencia es gramatical, es decir, si pertenece al lenguaje descrito por las reglas sintácticas, y, por otro lado, devuelve los árboles derivados (más de uno si se trata de una cadena ambigua) correspondientes a la sentencia que se analiza. Cuando el algoritmo sólo realiza la primera función se le denomina *reconocedor*.

En función de la estrategia de construcción de las estructuras sintácticas, los analizadores sintácticos se dividen en ascendentes, descendentes y mixtos. Los ascendentes son los que crean los árboles derivados desde las hojas hasta la raíz, los descendentes desde la raíz a las hojas y los mixtos son los que usan una combinación de ambas estrategias. También en función de la forma de recorrer la cadena de entrada, los analizadores se pueden dividir en unidireccionales y bidireccionales. Los primeros recorren la cadena de entrada de izquierda a derecha y los segundos en ambas direcciones.

Por razones históricas, ya que son los más longevos, y de eficiencia computacional, puesto que se analizan con un coste espacial de  $\mathcal{O}(n^2)$  y temporal de  $\mathcal{O}(n^3)$  en el caso peor ( $n$  es la longitud de la cadena a analizar), el mayor número de analizadores sintácticos definidos en la literatura corresponden al formalismo CFG. Existen dos importantes clases de algoritmos de análisis para CFG:

- *Analizadores tabulares*

También denominados analizadores basados en *chart*, que son aquellos que emplean técnicas de *programación dinámica* [12]. Estos algoritmos almacenan análisis parciales de la cadena de entrada en una tabla llamada *chart* y usa la información recogida en la misma para construir análisis mayores. Los elementos almacenados en el *chart* se denominan *ítems*.

- *Analizadores LR*

Un analizador LR consta de una *tabla de análisis LR* generada a partir de las reglas de la gramática, una *estructura de control* y una *pila* [4]. En cada paso del algoritmo, los denominados *ítems LR* se apilan o desapilan en función de la cadena de entrada y la tabla LR. Como se puede observar, los analizadores LR son autómatas de pila y, por tanto, sólo pueden reconocer una subclase de las CFG. La mayor difusión de este tipo de analizadores fue a raíz de la definición en [130, 131] de un algoritmo *LR generalizado* que analiza cualquier clase de CFG.

El primer analizador tabular de cierto interés fue presentado por Cocke, Younger y Kasami [74, 145]. Se trata de un método ascendente puro con recorrido inidireccional de la



cadena de entrada, pero requiere que la CFG esté expresada en Forma Normal de Chomsky (CFN, *Chomsky Normal Form*) [63]. Earley introduce en [48, 49] un analizador tabular con estrategia mixta que trabaja con CFG no restringidas. Concretamente, se trata de un algoritmo ascendente con información descendente y recorrido unidireccional de la cadena de entrada. La importancia del algoritmo de Earley es tal, que se puede considerar el punto de referencia de muchos otros analizadores tabulares descritos en la literatura. Se denominan *predictivos* a los analizadores que emplean una estrategia igual a la del algoritmo de Earley. En [127] se presenta una versión ascendente pura del algoritmo de Earley, denominada buE (*bottom-up Earley*). En [56] se propone un reconocedor derivado del de Earley, obtenido mediante la eliminación de predicciones de la palabra vacía ( $\epsilon$ ) y la precompilación de cadenas de derivaciones.

Uno de los analizadores ascendentes de mayor interés práctico, derivado del algoritmo de Earley, es el denominado *Left Corner* (LC) (esquina izquierda), cuya primera versión determinista se presentó en [104]. Este analizador usa una relación entre símbolos de la gramática, denominada relación de *Left Corner*, para precompilar las predicciones, disminuyendo de este modo el número de *ítems* que se almacena en el *chart* durante el proceso de análisis y, por tanto, aumentando la prestaciones del algoritmo. Las primeras ideas de analizador LC generalizado fueron mostradas en [96]. Un analizador ascendente de estilo left corner en Prolog fue presentado en [82], sin embargo, presenta excesivos problemas debido al uso del formalismo DCG. Otra variante de LC generalizado [83] utiliza pilas estructuradas en grafos. En [127] se presenta una versión de LC que emplea una aproximación basada en *chart* similar al analizador ascendente dirigido de Kay [75]. En ese mismo trabajo Sikkell muestra la relación existente entre el analizador de Earley y los basados en LC.

Entre los algoritmos ascendentes con recorrido bidireccional de la cadena de entrada podemos destacar el definido por de Vreught y Honig en [139, 140] y los analizadores basados en *núcleo* (HC, *Head Corner*). La diferencia fundamental entre ambos es que mientras el primero puede trabajar con CFG sin restricciones, los segundos requieren que cada regla de una CFG disponga de un *núcleo*, que no es más que un símbolo destacado (por motivaciones lingüísticas o interés computacional) de la parte derecha de la regla, de forma que el análisis de cada regla comience por él. El uso de predicciones dirigidas por núcleo para mejorar la eficiencia fue sugerido por Kay en [77]. El primer analizador HC ascendente puro fue presentado en [106], y los mismos autores muestran una descripción detallada de los algoritmos HC en el entorno de algoritmos bidireccionales en [107]. Otros analizadores de tipo HC, basados en variantes de algoritmos de análisis LR, se presentan en [87].

Por último, también podemos citar el desarrollo de analizadores basados en técnicas programación paralela [5], cuyas aplicaciones prácticas están muy limitadas por las fuertes exigencias respecto a las arquitecturas del hardware sobre las que funcionan.

Las críticas más importantes a todos estos analizadores radican precisamente en lo que podría ser su gran virtud, su *completitud*. Es decir, para que estos algoritmos cumplan su función, que es devolver todos los árboles derivados asociados a una sentencia, es necesario

una definición completa de la gramática. Esto supone un serio problema en los lenguajes naturales, provocado fundamentalmente por su riqueza gramatical y su naturaleza ambigua. Por ello, recientemente existe un creciente interés por analizadores que usan técnicas estadísticas para [81]:

- acelerar el análisis sintáctico, usando probabilidades para limitar el número de estructuras sintácticas y, de esta manera, reducir los tiempos de ejecución [30].
- tareas de *desambiguación sintáctica*, mediante la selección del árbol de análisis más probable [15].

Determinados campos de aplicación del procesamiento del lenguaje natural, por ejemplo la extracción y recuperación de información [65], son terrenos abonados para el desarrollo de nuevas técnicas de análisis sintáctico. Estas aplicaciones habitualmente trabajan *on-line* y con textos no restringidos, por consiguiente, son cruciales tanto la obtención de un árbol de análisis (aunque sea incompleto) como el tiempo de respuesta. Para dar cobertura a estas necesidades, surgen dos grupos de algoritmos de análisis: los *analizadores superficiales* (*Shallow Parsing*), que buscan el análisis completo pero no en profundidad, y los *analizadores fragmentales* o agrupadores sintácticos (*Chunking*), que buscan el análisis en profundidad pero no necesariamente completo.

### 1.2.1 Análisis sintáctico de TAG

La mayoría de los algoritmos de análisis sintáctico para TAG descritos en la literatura están basados en otros ya definidos para CFG. Para poner de manifiesto el paralelismo existente entre la historia y los métodos de análisis sintáctico de ambos formalismos, vamos a citar algunos analizadores para TAG, encuadrándolos en los mismos grupos en los que dividimos los nombrados para CFG:

- *Analizadores tabulares*

Entre los ascendentes unidireccionales, podemos citar el primer analizador sintáctico definido para TAG [136] basado en el método CYK y la versión ascendente del método de Earley [7] basada en el buE de Sikkel. En el formalismo TAG, los analizadores predictivos basados en el algoritmo de Earley se dividen en dos clases, en función de si verifican [84] o no [71] la *propiedad del prefijo válido* o detección temprana de errores [108]. Entre los ascendentes bidireccionales, encontramos el definido en [41], que está basado en el algoritmo propuesto para CFG por de Vreught y Honig. Entre los ascendentes bidireccionales basados en HC, el más destacable es el analizador presentado por Van Noord en [88].

- *Analizadores LR*

En este grupo podemos destacar el trabajo pionero en este tipo de analizadores, presentado en [112] y basado en EPDA (*Embedded Push-Down Automata*), o los presentados en [85, 97], que están basados en PDA (*Push Down Automata*).

Aparte de estas adaptaciones de los métodos clásicos, se han empleado una serie de técnicas avanzadas para construir analizadores para TAG. Dado la gran cantidad de trabajos existentes y que no es nuestro propósito ser exhaustivos, solo citaremos algunos que puedan servir de referencia sobre diversas técnicas. Entre las cuales podemos citar el análisis de TAG en dos etapas [94, 95], el uso de estrategias paralelas [89] y el análisis de TAG mediante el análisis de CFG síncronas [57] o mediante la multiplicación de matrices booleanas [98].

Un hito importante en el campo del análisis sintáctico fue la aparición de los *esquemas de análisis sintáctico* [125] como un método general y formal para la descripción de algoritmos de análisis sintáctico. El trabajo de Sikkel se puede considerar como una formalización de estudios previos sobre analizadores deductivos [122]. Entre las ventajas del uso de esquemas de análisis para describir analizadores podemos destacar las siguientes:

- Al ser especificaciones, abstrae de los problemas de la implementación derivados de las estructuras de control y de datos.
- Al ser un sistema formal, permite establecer relaciones formales entre los algoritmos.
- Sus fundamentos matemáticos permiten demostrar la corrección de los algoritmos de una forma casi sistemática [126].

En [127] se especifican mediante esquemas de análisis los principales analizadores para CFG descritos en la literatura y se establecen las relaciones existentes entre ellos. Usando el modelo de Sikkel, Miguel Ángel Alonso [6] presenta un conjunto de esquemas que constituyen el núcleo de un red de analizadores tabulares para TAG similar a la definida por Sikkel para CFG, concretamente, presenta esquemas basados en los métodos CYK, buE y Earley. Para este último presenta versiones con y sin prefijo válido. En el mismo trabajo, el autor muestra las relaciones formales existentes entre ellos. Posteriormente, Víctor Jesús Díaz [35] amplía dicho conjunto con la introducción de un esquema bidireccional basado en el algoritmo de Vreught y Honig y algunas variantes sobre el mismo.

### 1.3 Objetivos

El principal problema que presentan los algoritmos de análisis sintáctico para TAG basados en los métodos clásicos de analizadores tabulares para CFG es su complejidad teórica, que se encuentra en las cotas de  $\mathcal{O}(n^6)$  en tiempo y  $\mathcal{O}(n^4)$  en espacio para el caso peor, siendo  $n$  la longitud de la cadena de entrada. Debido a esto, en los últimos años se está empleando un gran esfuerzo en el desarrollo de técnicas que permitan mejorar el comportamiento práctico de estos analizadores. En algunos casos se emplean autómatas de estado finito para compactar el tamaño de las gramáticas [50, 80, 59], mientras en otros se ha optado por la restricciones sobre el formalismo. Satta y Schuler [105] proponen un conjunto de restricciones sobre la definición de TAG que reducen la cota de complejidad temporal del análisis hasta  $\mathcal{O}(n^5)$ . Otros autores, sin embargo, basándose en el hecho de que la mayor parte

de los lenguajes naturales caen dentro de los CFL, presentan restricciones a las TAG que disminuyen su poder expresivo a CFL. A cambio, su análisis se puede realizar con el mismo coste que el análisis de CFG. En este grupo, podemos encontrar la propuesta de Rogers [102] y las gramáticas de inserción de árboles (TIG, *Tree Insertion Grammars*) [114]. Por último, también se encuentran algunas propuestas que usan la combinación de diversas técnicas avanzadas [58].

El trabajo desarrollado en esta memoria se encuadra dentro del diseño de analizadores tabulares para gramáticas de adjunción de árboles. Usaremos los esquemas de análisis sintáctico como método de especificación de los distintos algoritmos, lo que nos permitirá establecer de manera formal las relaciones entre los nuevos esquemas presentados y los ya existentes, así como demostrar su corrección.

Nuestro objetivo es obtener analizadores para TAG que, aunque presenten los mismos costes teóricos, muestren un comportamiento práctico mejor que otros ya definidos. Para ello vamos a aplicar dos técnicas:

- Extender el concepto de esquina izquierda (LC, *Left Corner*) de CFG a TAG. Definiremos *Left Corner* en una gramática TAG y la aplicaremos para obtener un nuevo esquema ascendente derivado de buE. A continuación definiremos *relación de Left Corner* en un *árbol elemental* de una TAG y la aplicaremos para obtener un nuevo esquema predictivo derivado de Earley. Usando implementaciones de los nuevos esquemas y un banco suficientemente significativo de gramáticas y sentencias de pruebas, comprobaremos las mejoras que aportan en las prestaciones.

Dado que todos los analizadores se especifican con esquemas de análisis, enriqueceremos la red de analizadores introducida en [6, 35] con estos algoritmos basados en LC, tanto en su versión ascendente como predictiva, estableciendo las relaciones con los esquemas allí definidos. También nos planteamos como objetivo adicional marcar las pautas generales para demostrar la corrección de esquemas de análisis para TAG, demostrando la corrección de los nuevos esquemas presentados.

- Dirigir el análisis por el núcleo TIG de la gramática. Una cuestión que se suele plantear es porqué emplear algoritmos con un coste teórico tan elevado para llevar a cabo el análisis sintáctico de lenguajes que en su mayor parte caen dentro de los CFL. Como respuesta a esta cuestión, proponemos una alternativa al analizador tabular clásico para TAG, consistente en determinar dentro de una gramática TAG el conjunto capaz de generar únicamente CFL, y desarrollar algoritmos que trabajen con un coste igual al de analizadores CFG cuando analicen dicho conjunto y con un coste igual al de analizadores TAG para el resto de la gramática.

Las gramáticas de inserción de árboles se caracterizan porque restringen el formalismo TAG hasta el punto que su capacidad generadora se reduce a CFL, sin embargo, se ha comprobado que la mayor parte de las estructuras usadas en las gramáticas de amplia



cobertura se adaptan a este formalismo. Por ello, la estrategia que emplearemos para construir este nuevo tipo de analizadores para TAG va a consistir en determinar el núcleo TIG dentro de una gramática TAG e integrar los algoritmos de análisis para TIG con los definidos para TAG.

En la literatura sólo se han presentado dos analizadores para TIG, uno predictivo basado en el algoritmo de Earley [114] y otro ascendente basado en CYK [115]. Ésto plantea la necesidad de definir una red de analizadores para TIG similar a la disponible para TAG para poder construir un conjunto de esquemas para TAG derivados a partir de esta estrategia. La importancia del formalismo TIG hace que la creación de esta red constituya por si sola un objetivo de nuestro trabajo, de manera que, además de definir los esquemas, estableceremos las relaciones formales entre ellos.

Mostraremos como los esquemas para TAG desarrollados con esta estrategia reducen la complejidad teórica media de los analizadores tabulares clásicos para TAG. Y, para terminar, compararemos el comportamiento práctico de dos esquemas con estrategia ascendente basado en el algoritmo de Earley para TAG, uno construido con esta técnica y otro derivado de forma directa, para lo que se usarán también implementaciones de ambos esquemas y un conjunto adecuado de gramáticas y sentencias de pruebas.

## 1.4 Estructura de la memoria y publicaciones asociadas

La memoria se encuentra estructurada en cinco partes claramente diferenciadas. La parte I la constituye este capítulo y otro que introduce los esquemas de análisis sintáctico. Las partes II y III están dedicadas a los esquemas de análisis para TAG y TIG, respectivamente. En la parte IV definimos un conjunto de esquemas de análisis combinados para TAG y TIG, usando la técnica descrita anteriormente. En los apéndices se presentan las pruebas de corrección de algunos de los analizadores para TAG definidos en la parte II y se definen un conjunto de esquemas auxiliares para TIG.

### Parte I. Preliminares

Capítulo 1.- Estado del arte de los formalismos gramaticales y algoritmos de análisis, objetivos de la memoria y distribución de los contenidos.

Capítulo 2.- Introducción de los conceptos teóricos necesarios para comprender y aplicar los esquemas de análisis sintáctico, los cuales se usan para definir todos los analizadores que presentemos en la memoria.

### Parte II. Gramáticas de adjunción de árboles

Capítulo 3.- Definición, propiedades fundamentales y formalismos relacionados con las gramáticas de adjunción de árboles.

Capítulo 4.- Se muestra como se adaptan los esquemas de análisis sintácticos para especificar analizadores para TAG. A continuación se da un repaso al conjunto de esquemas de análisis para TAG descritos en la literatura y que constituyen la red básica de analizadores cuyo núcleo es el algoritmo de Earley. Concretamente, los esquemas basados en CYK, buE, Earley (con y sin prefijo válido) y de Vreught y Honig. De este último solamente recogemos su definición, ya que una descripción detallada del mismo y su relación con otros esquemas fue incluida en [35]. También se muestra un diagrama con las relaciones formales entre los analizadores presentados en este capítulo.

El analizador ascendente bidireccional aquí mostrado ha sido publicado en:

- V. J. Díaz, V. Carrillo y M. A. Alonso. A bidirectional bottom-up parser for TAG. In Proc. of *6th International Workshop on Parsing Technologies (IWPT'2000)*, Trento, Italy. (2000)
- V. J. Díaz, M. A. Alonso y V. Carrillo. Bidirectional parsing of TAGs without heads. In Proc. of *5th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+5)*, Paris, France. (2000)

Capítulo 5.- En este capítulo se define el concepto de *Left Corner* en el contexto del formalismo TAG, el cual se aplica para obtener una versión ascendente y dos predictivas de algoritmos basados en LC para TAG. Adicionalmente, se presenta una variante predictiva que sirve como esquema intermedio para mostrar la relación existente entre los dos analizadores predictivos definidos. También se hace una comparativa de prestaciones entre las versiones ascendentes y predictivas del analizador basado en Earley del capítulo 4 y las basadas en LC descritas en este capítulo.

Los analizadores basados en LC, tanto la versión ascendente como las dos predictivas, y los resultados experimentales obtenidos en la comparativa han sido publicados en:

- V. Carrillo, V. J. Díaz y M. A. Alonso. Análisis sintáctico ascendente de TAGs guiado por la esquina izquierda. *Procesamiento del lenguaje Natural*, 27, pp 47-54. (2001)
- V. Carrillo, V. J. Díaz y M. A. Alonso. A predictive left corner parser for tree adjoining grammars. *Revista electrónica en Tecnologías del Habla*, 1. (2003)
- V. J. Díaz, M. A. Alonso y V. Carrillo. A Left Corner parser for TAGs. In Proc. of *6th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+6)*, Venice, Italy. (2002)

Capítulo 6.- Este capítulo demuestra las relaciones formales existentes entre los esquemas basados en LC que se introdujeron en el capítulo anterior, así como la relación de éstos con los esquemas mostrados en el capítulo 4.

Todas las demostraciones de este capítulo han sido publicadas en el siguiente informe interno:

- V. Carrillo. Esquemas basados en left corner sin prefijo válido para TAGs: Relaciones. Informe Interno LSI-2002-03, Departamento de Lenguajes y Sistemas Informáticos, Universidad de Sevilla. (2002)

### Parte III. Gramáticas de inserción de árboles

Capítulo 7.- En este capítulo se definen las gramáticas de inserción de árboles y se muestran las relaciones entre este formalismo, las CFG y las TAG. También se define el esquema de análisis sintáctico correspondiente al analizador basado en Earley para TIG descrito en [114], y se establece la equivalencia entre éste y el algoritmo de Earley para CFG mediante el empleo de una nueva forma para representar los árboles de la gramática. La nueva representación de los árboles de la gramática, denominada notación *plana*, y su aplicación para establecer la equivalencia entre los analizadores para TIG y CFG han sido publicadas en:

- V. J. Díaz, V. Carrillo y M. Toro. Elementary tree representation. In Proc. of 1st Workshop on Tabulation in Parsing and Deduction (TAPD'98), pp 10–15, Paris, France. (1998)
- V. J. Díaz, V. Carrillo y M. Toro. A review of Earley-based parser for TIG. *Methodology and Tools in Knowledge-Based Systems, Vol 1415 of Lecture Notes in Artificial Intelligence*, pp 732–738, Springer-Verlag. (1998)

Capítulo 8.- Se definen un conjunto de nuevos analizadores para TIG, concretamente, los basados en CYK, buE, Earley y de Vreught y Honig. Y se establecen las relaciones formales existentes entre ellos.

Los nuevos esquemas para TIG y las demostraciones de las relaciones han sido publicados en:

- V. Carrillo, V. J. Díaz y M. A. Alonso. Algoritmos de análisis para gramáticas de inserción de árboles. *Procesamiento del lenguaje Natural, 29*, pp 89–96. (2002)
- V. Carrillo. Algoritmos de análisis para gramáticas de inserción de árboles: Relaciones. Informe Interno LSI-2003-01, Departamento de Lenguajes y Sistemas Informáticos, Universidad de Sevilla. (2003)

Capítulo 9.- De forma similar a como se hizo en el capítulo 5 para las TAG, aquí se introduce el concepto de *Left Corner* en el contexto del formalismo TIG y se definen una versión ascendente y dos predictivas de algoritmos basados en LC para TIG. También se presenta una variante predictiva que sirve como esquema de enlace para mostrar la relación entre los dos esquemas predictivos. Por último, se integran estos esquemas basados en LC en la red de analizadores para TIG iniciada en el capítulo anterior.

Los analizadores basados en LC, tanto la versión ascendente como las dos predictivas, y las demostraciones de las relaciones formales han sido publicados en:

- V. Carrillo y V. J. Díaz. Left corner parser for tree insertion grammars. *Artificial Intelligence: Methodology, Systems, and Applications - AIMS'A'02, Vol 2443 of Lecture Notes in Artificial Intelligence*, pp 142–151, Springer-Verlag. (2002)
- V. Carrillo, M. A. Alonso y V. J. Díaz Análisis eficiente de gramáticas de inserción de árboles. *II Jornadas en Tecnologías del Habla*, Granada, España. (2002)
- V. Carrillo. Algoritmos de análisis para gramáticas de inserción de árboles: Relaciones. Informe Interno LSI-2003-01, Departamento de Lenguajes y Sistemas Informáticos, Universidad de Sevilla. (2003)

#### Parte IV. Análisis sintáctico de TAG dirigido por el núcleo TIG

Capítulo 10.- En este capítulo se definen un conjunto de analizadores para TAG dirigidos por el núcleo TIG de la gramática. Para ello se integran los esquemas de análisis para TAG basados en CYK, buE, Earley y de Vreught y Honig definidos en el capítulo 4 con variantes de sus homónimos para TIG definidos en el capítulo 8. Con objeto de no desviar la atención de los objetivos de este tema, las definiciones de estas variantes de los esquemas para TIG se trasladan al apéndice B. Se hace un estudio detallado de la complejidad teórica de los nuevos analizadores obtenidos para poner de manifiesto las mejoras que ofrecen en la complejidad teórica media. Además se realiza una comparativa experimental entre el nuevo esquema basado en Earley y el mostrado en el capítulo 4.

El esquema combinado basado en Earley y los resultados experimentales de la comparativa han sido publicados en:

- M. A. Alonso, V. Carrillo y V. J. Díaz. Mixed Parsing of Tree Insertion and Tree Adjoining Grammars. *Advances in Artificial Intelligence. IBERAMIA 2002, Vol 2527 of Lecture Notes in Artificial Intelligence*, pp 694–703, Springer-Verlag. (2002)
- M. A. Alonso, V. Carrillo y V. J. Díaz. Análisis sintáctico combinado de TAGs y TIGs. *Procesamiento del lenguaje Natural*, 29, pp 65–72. (2002)

#### Parte V. Apéndices

Apéndice A.- El objetivo de este apéndice es marcar las pautas para probar la corrección de esquemas para TAG. Para ello, demostraremos la corrección de los nuevos esquemas basados en LC que no verifican la propiedad del *prefijo válido* que introducimos en esta memoria.

Apéndice B.- Se definen un conjunto de esquemas para TIG similar al descrito en el capítulo 8, pero que, a diferencia de aquellos, sólo permiten una adjunción sobre un nodo. El interés práctico de estos esquemas se limita a que se pueden integrar con los esquemas para TAG para obtener los esquemas combinados definidos en el capítulo 10.



Resultados de investigación colaterales, de menor relevancia para la línea argumental de la tesis, fueron publicados en: [18], [19], [20], [21], [22], [23], [36], [39] y [40].

## Capítulo 2

# Enfoque deductivo en el análisis sintáctico

Este capítulo pretende recoger los conceptos necesarios para comprender y usar los esquemas de análisis sintáctico definidos por Sikkel [127] para especificar, analizar y relacionar algoritmos de análisis sintáctico para gramáticas incontextuales.

El capítulo se estructura de la siguiente forma. En la sección 1 se da un breve repaso a la gramáticas incontextuales y un conjunto de conceptos relacionados con ellas. En la sección 2 se definen los sistemas deductivos de forma genérica, así como algunas ampliaciones para poder definir la corrección de los mismos. En la sección 3 se definen un tipo particular de sistema deductivo pensado para especificar de manera sencilla algoritmos de análisis sintáctico: los esquemas de análisis sintáctico basados en ítems. Además, en esa misma sección se propone un método general para demostrar la corrección de los algoritmos de análisis sintáctico especificados mediante estos esquemas. En la sección 4 se muestran la especificación mediante esquemas de análisis sintáctico de algunos de los algoritmos de análisis clásicos para CFG. En la sección 5 se define el concepto de complejidad y se detalla cómo se calcula la complejidad de los algoritmos especificados con esquemas de análisis. En la sección 6 se definen los distintos tipos de relaciones formales que se pueden establecer entre los esquemas de análisis, así como las propiedades más destacadas de dichas relaciones.

### 2.1 Gramáticas incontextuales

Vamos a recordar algunos conceptos de teoría de lenguajes formales [62] que usaremos en el resto de este trabajo.

#### **Definición 2.1** Cadenas

Sea  $X$  un conjunto arbitrario. Denotamos  $X^+$  el conjunto de cadenas no vacías  $x_1 \dots x_k$ , ( $k \geq 1$ ) sobre  $X$ .

Denotamos  $X^*$  al conjunto de cadenas  $x_1 \dots x_k$ , ( $k \geq 0$ ) sobre  $X$ .

**Definición 2.2** *Gramática incontextual*

Una gramática independiente del contexto o gramática incontextual (CFG) es una 4-tupla  $G = (V_N, V_T, \mathcal{P}, S)$  que satisface:

- (i) el conjunto de símbolos terminales  $V_T$  y el conjunto de símbolos no terminales  $V_N$  son alfabetos tomados de alguna clase universal de símbolos  $Sym$ , y  $V_T \cap V_N = \emptyset$ ;
- (ii) el conjunto de producciones  $\mathcal{P}$  consta de un número finito de pares  $(A, \alpha)$  donde  $A \in V_N$  y  $\alpha \in (V_N \cup V_T)^*$ ;
- (iii) el símbolo de comienzo o axioma  $S$  es un símbolo no terminal de  $V_N$ .

Denotamos  $CFG$  a la clase de gramáticas independientes del contexto.

**Notación 1**

- $V = V_N \cup V_T$ .
- Las producciones  $(A, \alpha)$  se escribirán  $A \rightarrow \alpha$ .
- Usaremos
  - $A, B, C, \dots$  para variables definidas sobre  $V_N$ ;
  - $X, Y, \dots$  para variables definidas sobre  $V$ ;
  - $a, b, \dots$  para variables definidas sobre  $V_T$ ;
  - $v, w, x, \dots$  para variables definidas sobre  $V_T^*$ ;
  - $\alpha, \beta, \gamma, \delta, \dots$  para variables definidas sobre  $V^*$ ;
  - la cadena vacía se denota mediante  $\varepsilon$ ;
  - una cadena de entrada la escribiremos habitualmente como  $a_1 \dots a_n$ .

**Definición 2.3** *Relación de derivación  $\Rightarrow$* 

La relación  $\Rightarrow$  sobre  $V^* \times V^*$  se define de la siguiente manera:

- $\alpha \Rightarrow \beta$  si existen  $\alpha_1, \alpha_2, A, \gamma$  tales que
- $\alpha = \alpha_1 A \alpha_2, \beta = \alpha_1 \gamma \alpha_2$  y  $A \rightarrow \gamma \in \mathcal{P}$ .

La clausura transitiva se denota  $\stackrel{\pm}{\Rightarrow}$  y la reflexiva y transitiva  $\stackrel{*}{\Rightarrow}$ .

Podemos definir distintas subclases de  $CFG$ , las cuales generarán distintas clases de lenguajes incontextuales. Entre otras, dos subclases frecuentemente usadas son las CFG acíclicas y las CFG sin cadenas vacías  $\varepsilon$ . En este capítulo y para ilustrar algún analizador usaremos una subclase especial: las gramáticas en *Forma Normal de Chomsky*.

**Definición 2.4** *Forma Normal de Chomsky (CNF)*

Una gramática incontextual  $G$  está en Forma Normal de Chomsky (CNF) si  $\mathcal{P}$  solo contiene producciones de la forma  $A \rightarrow BC$  y  $A \rightarrow a$ .

Denotamos  $CNF$  a la clase de gramáticas en Forma Normal de Chomsky.

**Definición 2.5** *Árboles*

Sea  $\mathcal{U}$  la clase de árboles finitos en la cual los hijos de un nodo tienen un orden de izquierda a derecha y cada nodo se etiqueta con un símbolo de  $Sym$ . Para  $G = (V_N, V_T, \mathcal{P}, S) \in CFG$ , el conjunto  $Trees(G) \subset \mathcal{U}$  es el conjunto de árboles cuyos nodos están etiquetados con símbolos de  $V \cup \{\varepsilon\}$ , y cada nodo  $u$  satisface una de las siguientes condiciones:

- $u$  es una hoja;
- $u$  está etiquetado con  $A$ , los hijos de  $u$  están etiquetados  $X_1, \dots, X_n$  y existe una producción  $A \rightarrow X_1 \dots X_n \in \mathcal{P}$ ;
- $u$  está etiquetado con  $A$ ,  $u$  tiene un solo hijo etiquetado  $\epsilon$  y existe una producción  $A \rightarrow \epsilon \in \mathcal{P}$ .

Usaremos  $\tau, \sigma, \dots$  para representar variables de árboles.

### Definición 2.6 Raíz y cosecha

Para  $G \in \text{CFG}$  y  $\tau \in \text{Trees}(G)$  se definen:

$\text{root}(\tau)$  es la etiqueta de la raíz de  $\tau$ ;

$\text{yield}(\tau)$  es la cadena que se obtiene mediante la concatenación de las etiquetas de todas las hojas de  $\tau$  en orden de izquierda a derecha.

Las hojas de  $\tau$  se etiquetan con símbolos de  $V \cup \{\epsilon\}$ . La cosecha es una cadena en  $V^*$ , ya que los símbolos  $\epsilon$  desaparecen en la concatenación. Solo si todas las hojas están etiquetadas con  $\epsilon$  la cosecha es la cadena vacía.

### Definición 2.7 Árbol de análisis

Un árbol  $\tau \in \text{Trees}(G)$  se denomina árbol de análisis para una cadena de entrada  $a_1 \dots a_n$  si  $\text{root}(\tau) = S$  y  $\text{yield}(\tau) = a_1 \dots a_n$ .

Una cadena perteneciente a  $V_T^*$  se dice que es válida con respecto a  $G$  si tiene un árbol de análisis. Una cadena válida también se denomina sentencia.

Ahora vamos a presentar un conjunto de extensiones no estándar a la definición de gramáticas incontextuales. Estas ampliaciones se necesitan para la definición de los esquemas de análisis basados en ítems, que veremos más adelante.

También vamos a incluir una convención notacional que se emplea en muchos algoritmos de análisis: usar dos símbolos que no pertenecen al alfabeto  $V$  para marcar el comienzo y el final de la cadena de entrada. Así, el final de la sentencia se indicará mediante un *marcador de fin de sentencia*, usualmente  $\$$ , el cual se añade a la cadena  $a_1 \dots a_n$  como el símbolo de la posición  $(n+1)$ -ésima. Del mismo modo, el comienzo de una sentencia se indicará mediante un *marcador de comienzo de sentencia*, usualmente  $\#$ , que se añade a la cadena como el símbolo de la posición 0.

### Definición 2.8 Terminal marcado

Para cada  $G \in \text{CFG}$  un terminal marcado es un par  $(a, j) \in (V_T \cup \{\#, \$\}) \times \mathbb{N}$ .

Escribiremos  $\underline{a}_j$  en lugar de  $(a, j)$ . También escribiremos  $\underline{V}_T$  para representar el conjunto  $(V_T \cup \{\#, \$\}) \times \mathbb{N}$ .



Para cada palabra de la sentencia  $a_j$  creamos un árbol  $\langle a \rightarrow \underline{a}_j \rangle$ . De esta forma, una sentencia  $a_1 \dots a_n$  se puede representar por un conjunto de árboles definido por

$$\{a \rightarrow \underline{a}_j \mid a \text{ es la } j\text{-ésima palabra de la sentencia}\}$$

Como es obvio, a los símbolos marcadores de comienzo y fin de sentencia les asociamos los árboles  $\langle \# \rightarrow \underline{\#}_0 \rangle$  y  $\langle \$ \rightarrow \underline{\$}_{n+1} \rangle$ , respectivamente.

**Definición 2.9** *Extensiones a CFG*

- Una pseudo-producción es una producción de la forma  $a \rightarrow \underline{a}_j$ . Denotaremos con  $\underline{\mathcal{P}}$  el conjunto de pseudo-producciones para una gramática dada.
- Las variables  $\alpha, \beta, \gamma, \dots$  ahora se definen sobre  $V_N \cup V_T \cup \underline{V}_T$ .
- La definición de  $Trees(G)$  es modificada como sigue: (i) los nodos se etiquetan con símbolos de  $V_N \cup V_T \cup \underline{V}_T$ ; (ii) además de las tres alternativas originales de la definición, un nodo  $u$  puede estar etiquetado con un símbolo terminal  $a$  y tener un solo hijo etiquetado  $\underline{a}_j$  para algún  $j$ .
- Un árbol  $\tau \in Trees(G)$  se denomina árbol de análisis marcado para una sentencia si  $root(\tau) = S$  y  $yield(\tau) = \underline{a}_1 \dots \underline{a}_n$ .

**Definición 2.10** *Conjunto de árboles de análisis marcados*

El conjunto de árboles de análisis marcados  $\mathcal{P}_G^{(n)}$  para una gramática incontextual  $G$  y todas las cadenas de longitud  $n$  se define

$$\mathcal{P}_G^{(n)} = \{\tau \in Trees(G) \mid \exists a_1 \dots a_n \in V_T^* : root(\tau) = S \wedge yield(\tau) = \underline{a}_1 \dots \underline{a}_n\}$$

El conjunto de árboles de análisis marcados  $\mathcal{P}_G(a_1 \dots a_n)$  para una gramática incontextual  $G$  y una cadena  $a_1 \dots a_n$  se define

$$\mathcal{P}_G(a_1 \dots a_n) = \{\tau \in Trees(G) \mid root(\tau) = S \wedge yield(\tau) = \underline{a}_1 \dots \underline{a}_n\}$$

## 2.2 Sistemas deductivos

El concepto general de sistemas deductivos es bien conocido en la lógica matemática. Sin embargo, en esta sección vamos a presentar una definición de sistema deductivo, tal como se muestra en [127], adaptada a facilitar la descripción de algoritmos de análisis.

Un sistema deductivo contiene un conjunto arbitrario de objetos, denominados *entidades*. El objetivo de un sistema deductivo es establecer que entidades son *válidas* a partir de un conjunto de *hipótesis* y por medio de un conjunto de *pasos deductivos*. Las entidades, en sentido amplio, sólo son objetos identificables que dependerán del sistema deductivo concreto. Así en los esquemas deductivos que veremos en la siguiente sección, dichas entidades serán "ítems".

**Definición 2.11** Paso deductivo, antecedente y consecuente

Sea  $X$  un conjunto de entidades y  $H$  un conjunto de hipótesis. Un paso deductivo es un par  $(Y, x)$ , donde  $Y \subseteq H \cup X$  es un conjunto finito y  $x \in X$ . Por tanto, un paso deductivo  $(Y, x)$  es un elemento del conjunto  $\wp_{fin}(H \cup X) \times X$ , donde  $\wp_{fin}$  es el conjunto potencia restringido a conjuntos finitos.

En un paso deductivo  $(\{y_1, \dots, y_k\}, x)$ , a las entidades  $y_1, \dots, y_k$  se les denominan antecedentes y a  $x$  el consecuente.

Un paso deductivo puede no tener antecedentes y el consecuente puede ser deducido sin tener en cuenta el conjunto de hipótesis.

**Definición 2.12** Sistema deductivo

Un sistema deductivo  $\mathbb{D}$  es una tripleta  $\langle X, H, \mathcal{D} \rangle$ , donde:

$X$  es un conjunto de entidades, denominado el dominio de  $\mathbb{D}$ ;

$H$  es un conjunto de hipótesis, que no necesariamente tiene que ser un subconjunto de  $X$ ;

$\mathcal{D} \subseteq \wp_{fin}(H \cup X) \times X$  es un conjunto de pasos deductivos.

Supongamos que una entidad  $x$  se puede deducir desde otras dos  $y_1$  e  $y_2$  mediante el paso deductivo  $(\{y_1, y_2\}, x)$ , entonces también se puede inferir desde un superconjunto de los antecedentes, por ejemplo,  $y_1, y_2$  e  $y_3$ . Sin embargo, esto no garantiza que si  $(\{y_1, y_2\}, x) \in \mathcal{D}$  también  $(\{y_1, y_2, y_3\}, x) \in \mathcal{D}$ . Por ello, a continuación vamos a definir la relación de inferencia  $\vdash$ , que es la clausura de  $\mathcal{D}$  mediante la adición de antecedentes a una inferencia.

**Definición 2.13** Relación de inferencia  $\vdash$ 

Sea un sistema deductivo  $\mathbb{D} = \langle X, H, \mathcal{D} \rangle$ , la relación  $\vdash \subseteq \wp(H \cup X) \times X$  se define mediante

$$Y \vdash x \text{ si } (Y', x) \in \mathcal{D} \text{ para algún } Y' \subseteq Y.$$

**Definición 2.14** Secuencias deductivas

Sea  $\mathbb{D} = \langle X, H, \mathcal{D} \rangle$  un sistema deductivo. Denotamos como  $X^+$  al conjunto de secuencias finitas no vacías  $x_1, \dots, x_j$ , donde  $j \geq 1$  y  $x_i \in X$  para  $1 \leq i \leq j$ .

Una secuencia deductiva en  $\mathbb{D}$  es un par  $(Y; x_1, \dots, x_j) \in \wp_{fin}(H \cup X) \times X^+$ , tal que

$$Y \cup \{x_1, \dots, x_{i-1}\} \vdash x_i \text{ para } 1 \leq i \leq j.$$

Como notación informal práctica para una secuencia deductiva  $(Y; x_1, \dots, x_j)$  usaremos  $Y \vdash x_1 \vdash \dots \vdash x_j$ .

El conjunto de secuencias de deducción  $\Delta_{\mathbb{D}} \subseteq \wp(H \cup X) \times X^+$  para un sistema deductivo  $\mathbb{D}$  se define

$$\Delta_{\mathbb{D}} = \{(Y; x_1, \dots, x_j) \in \wp(H \cup X) \times X^+ \mid Y \vdash x_1 \vdash \dots \vdash x_j\}.$$



**Definición 2.15** *Relación de inferencia reflexiva y transitiva*  $\vdash^*$

Sea  $ID = \langle X, H, \mathcal{D} \rangle$  un sistema deductivo. Se define la relación  $\vdash^*$  sobre  $\wp_{fin}(H \cup X) \times X$  mediante

$$Y \vdash^* x \text{ si } x \in Y \text{ o } Y \vdash \dots \vdash x.$$

**Definición 2.16** *Entidades válidas*

Para un sistema deductivo  $ID = \langle X, H, \mathcal{D} \rangle$  el conjunto de entidades válidas viene definido por

$$\mathcal{V} = \{x \in X \mid H \vdash^* x\}.$$

### 2.2.1 Corrección de los sistemas deductivos

Puesto que vamos a usar los sistemas deductivos para describir algoritmos de análisis sintáctico, y para cualquier algoritmo una propiedad fundamental es su *corrección*, es necesario que hagamos una ampliación de la definición de sistema deductivo vista previamente para poder establecer la corrección de los mismos.

De manera informal, primero se establece un conjunto de *entidades finales*, el cual es un subconjunto del conjunto de entidades del sistema. Después se divide este subconjunto en *entidades finales correctas e incorrectas*. La corrección es conocida por definición. Por ejemplo, si las entidades fueran árboles de análisis sintáctico para cualquier sentencia: las entidades finales correctas serían aquellos árboles de análisis para la sentencia que se está analizando, y las entidades finales incorrectas aquellos árboles de análisis para otras sentencias distintas de la que se está analizando.

Un sistema deductivo es correcto si todas las entidades finales correctas son válidas y todas las entidades finales incorrectas no son válidas. Esto se formaliza mediante las siguientes definiciones.

**Definición 2.17** *Sistema deductivo ampliado*

Un sistema deductivo ampliado  $\mathbb{E}$  es una 5-tupla  $\langle X, H, F, C, \mathcal{D} \rangle$ , donde

$X$  es el conjunto de entidades,

$H$  es el conjunto de hipótesis,

$F \subseteq X$  es el conjunto de entidades finales,

$C \subseteq F$  es el conjunto de entidades finales correctas,

$\mathcal{D} \subseteq \wp_{fin}(H \cup X) \times X$  es el conjunto de pasos deductivos.

Las definiciones de entidades válidas y relaciones de inferencia vistas previamente son válidas para sistemas deductivos ampliados.

**Definición 2.18** *Corrección*

Sea  $\mathbb{E} = \langle X, H, F, C, \mathcal{D} \rangle$  un sistema deductivo ampliado.

- $\mathbb{E}$  es consistente si todas las entidades finales válidas son correctas, es decir,  $F \cap \mathcal{V} \subseteq C$ .

- $\mathbb{E}$  es completo si todas las entidades finales correctas son válidas, es decir,  $C \subseteq F \cap \mathcal{V}$ .
- $\mathbb{E}$  es correcto si es consistente y completo, es decir,  $C = F \cap \mathcal{V}$ .

## 2.3 Esquemas de análisis basados en ítems

Una vez establecidas las bases teóricas necesarias, vamos a proceder a definir los *esquemas de análisis basados en ítems*, que no son más que sistemas deductivos donde las entidades son *ítems*.

De manera simplificada, se puede considerar que un *ítem* representa un conjunto de restricciones sobre un árbol de análisis parcial o completo, es decir, es una especificación parcial de un árbol de análisis. Si hay un conjunto de árboles que cumplen dicha especificación entonces este conjunto constituye una clase de equivalencia en el dominio de los árboles de la gramática. A continuación formalizaremos el concepto de *ítem* y sus tipos.

### Definición 2.19

Una partición  $\Pi(X) \subset \wp(X)$  es una colección de subconjuntos disjuntos no vacíos de parejas de  $X$  tal que cada  $x \in X$  está contenido en algún  $\pi \in \Pi(X)$ . Cada partición  $\Pi$  define una relación de equivalencia  $\sim_\pi$  mediante

$$x \sim_\pi y \text{ si hay un } \pi \in \Pi(X) \text{ tal que } \{x, y\} \subseteq \pi.$$

### Definición 2.20 Conjunto de ítems

Sea  $Trees(G)$  el conjunto de árboles para una gramática incontextual  $G$ . Un conjunto  $\mathcal{I} \in \wp(Trees(G))$  se denomina un conjunto de ítems si hay una partición  $\Pi$  de  $Trees(G)$  tal que  $\mathcal{I} \subseteq \Pi(Trees(G)) \cup \{\emptyset\}$ .

Según el tipo de árboles que especifican, los ítems se pueden clasificar en: *completos* si especifican árboles de análisis completos, *intermedios* si especifican árboles de análisis parciales y *mixtos* que son una combinación de los dos anteriores.

### Definición 2.21 Tipos de ítems

Sea  $\mathcal{I}$  un conjunto de ítems.

- Un ítem  $\iota \in \mathcal{I}$  se dice vacío si  $\iota = \emptyset$ .
- Un ítem no vacío  $\iota \in \mathcal{I}$  se dice completo si todo  $\tau \in \iota$  es un árbol de análisis marcado para alguna sentencia.
- Un ítem no vacío  $\iota \in \mathcal{I}$  se dice intermedio si todo  $\tau \in \iota$  no es un árbol de análisis marcado para cualquier sentencia.
- Un ítem  $\iota \in \mathcal{I}$  se dice mixto si existen  $\sigma, \tau \in \iota$  tales que  $\sigma$  es un árbol de análisis marcado y  $\tau$  no es un árbol de análisis marcado.

**Definición 2.22** *Conjunto de ítems regular y semiregular*

Un conjunto de ítems regular es aquel que no contiene ítems mixtos ni vacíos.

Un conjunto de ítems semiregular es aquel que no contiene ítems mixtos.

**Definición 2.23** *Ítems finales e ítems finales correctos*

Sea  $\Pi$  una partición de  $Trees(G)$  para una gramática incontextual  $G$ .

El conjunto de ítems finales  $\mathcal{F}_{G,\Pi}^{(n)}$  para una cadena de entrada de longitud  $n$  se define

$$\mathcal{F}_{G,\Pi}^{(n)} = \{\iota \in \Pi(Trees(G)) \mid \exists \tau \in \iota : \tau \in \mathcal{P}_G^{(n)}\}.$$

El conjunto de ítems finales correctos  $\mathcal{C}_{G,\Pi}$  para una cadena de entrada  $a_1 \dots a_n$  se define

$$\mathcal{C}_{G,\Pi}(a_1 \dots a_n) = \{\iota \in \Pi(Trees(G)) \mid \exists \tau \in \iota : \tau \in \mathcal{P}_G(a_1 \dots a_n)\}.$$

**Definición 2.24** *Sistema de análisis y sistemas de análisis correcto*

Un sistema deductivo  $\langle \mathcal{I}, H, \mathcal{D} \rangle$  es un sistema de análisis para una gramática incontextual  $G$  y una cadena de entrada  $a_1 \dots a_n$  si se cumple:

(i)  $\mathcal{I} = \mathcal{I}(G, \Pi)$  es un conjunto de ítems al que denominamos dominio,

(ii)  $\mathcal{F}_{G,\Pi}^{(n)} \subset \mathcal{I}$ ,

(iii)  $[a \rightarrow \underline{a}_i] \in H$  para cada  $a_i$ ,  $1 \leq i \leq n$ .

Un sistema de análisis  $\langle \mathcal{I}, H, \mathcal{D} \rangle$  para una gramática incontextual  $G$  y una cadena de entrada  $a_1 \dots a_n$  es correcto si el sistema deductivo ampliado

$$\langle \mathcal{I}, H, \mathcal{F}_{G,\Pi}^{(n)}, \mathcal{C}_{G,\Pi}(a_1 \dots a_n), \mathcal{D} \rangle$$

es correcto.

**Definición 2.25** *Sistema de análisis no instanciado y sistema de análisis no instanciado correcto*

Un sistema de análisis no instanciado para una gramática  $G$  es una tripleta  $\langle \mathcal{I}, \mathcal{H}, \mathcal{D} \rangle$ , donde

$\mathcal{H}$  es una función que asigna un conjunto de hipótesis a cada cadena de entrada  $a_1 \dots a_n$  y se define

$$\mathcal{H}(a_1 \dots a_n) = \{[a \rightarrow \underline{a}_i] \mid a = a_i\} \cup \{[\# \rightarrow \underline{\#}_0], [\$ \rightarrow \underline{\$}_{n+1}]\};$$

y  $\langle \mathcal{I}, \mathcal{H}(a_1 \dots a_n), \mathcal{D} \rangle$  es un sistema de análisis para cada  $a_1 \dots a_n \in V_T^*$ .

Un sistema de análisis no instanciado  $\langle \mathcal{I}, \mathcal{H}, \mathcal{D} \rangle$  es correcto si  $\langle \mathcal{I}, \mathcal{H}(a_1 \dots a_n), \mathcal{D} \rangle$  es correcto para toda  $a_1 \dots a_n \in V_T^*$ .

**Definición 2.26** *Esquema de análisis y esquema de análisis correcto*

Un esquema de análisis  $\mathbf{P}$  para una subclase de gramáticas incontextuales  $\mathcal{CG} \subseteq \mathcal{CFG}$  es una función que asigna un sistema de análisis no instanciado a cada gramática  $G \in \mathcal{CG}$ .

Un esquema de análisis  $\mathbf{P}$  es correcto si para cada  $G \in \mathcal{CG}$  el sistema de análisis no instanciado  $\mathbf{P}(G)$  es correcto.

**Definición 2.27** *Semiregularidad*

Un sistema de análisis  $\langle \mathcal{I}, H, \mathcal{D} \rangle$  es semiregular si  $\mathcal{I}$  es un conjunto de ítems semiregular y  $\emptyset \notin \mathcal{V}$ .

Un esquema de análisis  $\mathbf{P}$  para una subclase de gramáticas incontextuales  $\mathcal{CG}$  es semiregular si  $\mathbf{P}(G)(a_1 \dots a_n)$  es semiregular para todo  $G \in \mathcal{CG}$  y toda  $a_1 \dots a_n \in V_T^*$ .

**Notación 2**

- Como notación más convencional, los ítems hipótesis los escribiremos  $[a, i - 1, i]$  en lugar de  $[a \rightarrow \underline{a}_i]$ .
- Por simplicidad, el conjunto de ítems finales  $\mathcal{F}_{G, \Pi}^{(n)}$  lo escribiremos como  $\mathcal{F}$  y el conjunto de ítems finales correctos  $\mathcal{C}_{G, \Pi}(a_1 \dots a_n)$  como  $\mathcal{C}$ .
- No haremos distinción entre sistemas de análisis y sistemas de análisis no instanciados, y denotaremos ambos como  $\mathbf{IP}$ .
- Subrayaremos los nombres de los analizadores para CFG con el objeto de distinguirlos de los análogos para TAG y TIG que veremos en posteriores capítulos.

**2.3.1 Corrección de los esquemas de análisis**

Dado un sistema de análisis  $\mathbf{IP}$ , con su conjunto de ítems válidos  $\mathcal{V}$ , es relativamente sencillo demostrar  $\mathcal{F} \cap \mathcal{V} = \mathcal{C}$ , es decir, que el sistema es correcto. El problema fundamental es determinar el conjunto de ítems válidos. En esta sección describimos el método general propuesto en [126] para establecer el conjunto  $\mathcal{V}$ .

El método consta de los siguientes pasos:

1. Definir un conjunto de *ítems viables*  $\mathcal{W} \subseteq \mathcal{I}$  que deben ser reconocidos por un sistema de análisis  $\mathbf{IP}$ . Si mostramos que  $\mathcal{V} = \mathcal{W}$ , entonces ya tenemos el conjunto de ítems válidos.
2. Probar que  $\mathcal{V} \subseteq \mathcal{W}$ , es decir, la consistencia, por inducción sobre  $\vdash$ .
3. Probar que  $\mathcal{W} \subseteq \mathcal{V}$ , es decir, la completitud. Para esta demostración es necesario definir una *función de longitud de derivación* (*dlf*)  $d$  sobre  $\mathcal{W}$ , y a continuación probar la completitud por inducción sobre  $d$ .

La elección de una *dlf* adecuada es fundamental. Podríamos estar tentados de elegir una función tal que si  $H \vdash^m \xi$  entonces  $d(\xi) = m$ , sin embargo, en la definición de nuestra función de longitud de derivación no es posible usar la relación de inferencia  $\vdash^*$ , ya que lo que hay que probar precisamente es que para cualquier  $\xi \in \mathcal{V}$  se cumple que  $H \vdash^* \xi$ . Por ello lo habitual es definir esta función en base a la longitud de derivación gramatical, es decir, asigna a cada ítem el número de arcos atravesados en el recorrido de árbol de análisis que se necesita para derivarlo.



**Definición 2.28** *Función de longitud de derivación (dlf)*

Dado un subconjunto de ítems  $\mathcal{W} \subseteq \mathcal{I}$  de un sistema de análisis  $\mathbb{P}$ , una función  $d : H \cup \mathcal{W} \rightarrow \mathbb{N}$  es una dlf si cumple:

1.  $d(h) = 0$  si  $h \in H$ ;
2. para cada  $\xi \in \mathcal{W}$  hay algún  $\eta_1, \dots, \eta_k \vdash \xi \in \mathcal{D}$  tal que  $\{\eta_1, \dots, \eta_k\} \subseteq \mathcal{W}$  y  $d(\eta_i) < d(\xi)$  para  $1 \leq i \leq k$ .

**Proposición 2.1**

Dado un subconjunto de ítems  $\mathcal{W} \subseteq \mathcal{I}$  de un sistema de análisis  $\mathbb{P}$ .

- Si para todo  $\eta_1, \dots, \eta_k \vdash \xi \in \mathcal{D}$  donde  $\eta_i \in H \cup \mathcal{W}$ ,  $1 \leq i \leq k$ , se mantiene que  $\xi \in \mathcal{W}$  entonces  $\mathcal{V} \subseteq \mathcal{W}$ .
- Si existe una función de longitud de derivación  $d : H \cup \mathcal{W} \rightarrow \mathbb{N}$  entonces  $\mathcal{W} \subseteq \mathcal{V}$ .

En este capítulo no vamos a ver ningún ejemplo de cómo determinar la corrección de un sistema (esquema) de análisis sintáctico, ya que este método quedará suficientemente ilustrado cuando acometamos las demostraciones de corrección de los nuevos analizadores para TAG que presentamos en el apéndice A.

## 2.4 Ejemplos de esquemas de análisis

Después de introducir todos los conceptos necesarios, vamos a usar los esquemas de análisis para describir algunos algoritmos de análisis sintáctico clásicos para gramáticas incontextuales.

En la literatura podemos encontrar multitud de analizadores para CFG, los cuales podemos agrupar en dos categorías básicas en función de su estrategia de análisis: ascendentes y descendentes. Nuestro planteamiento en esta sección no es ser exhaustivos. Por cuestiones obvias de espacio no vamos a presentar todos los analizadores para CFG conocidos, sino aquellos de especial relevancia para el desarrollo argumental de la memoria.

Empezaremos por un analizador que presenta una estrategia de análisis ascendente, pero que presenta un serio inconveniente a la hora de definir gramáticas para el procesamiento del lenguaje natural, ya que sólo se puede aplicar a gramáticas en Forma Normal de Chomsky. Este analizador recibe el nombre de CYK y fue presentado por Cocke, Younger y Kasami en [145, 74].

**Ejemplo 2.1** *CYK*

Dadas una gramática  $G = (V_N, V_T, \mathcal{P}, S) \in \mathcal{CNF}$  y una cadena  $a_1 \dots a_n$ , un sistema de análisis para el algoritmo CYK se define:

$$\mathbb{P}_{\text{CYK}} = \langle \mathcal{I}_{\text{CYK}}, H, \mathcal{D}_{\text{CYK}} \rangle$$

El conjunto de hipótesis  $H$  es:

$$H = \{[a, i - 1, i] \mid a = a_i, 1 \leq i \leq n\} \cup \{[\#, -1, 0], [ \$, n, n + 1]\}$$

Como se puede comprobar el conjunto de hipótesis se obtiene aplicando la función  $\mathcal{H}$  del sistema no instanciado a la cadena de entrada. De ahora en adelante omitiremos la especificación del conjunto de hipótesis de los sistemas de análisis cuando se aplique la función por defecto  $\mathcal{H}(a_1 \dots a_n)$ .

El dominio de ítems del sistema viene dado por:

$$\mathcal{I}_{\text{CYK}} = \{[A, i, j] \mid A \in V_N, 0 \leq i \leq j\}$$

La semántica de un ítem de este conjunto es que el símbolo no terminal  $A$  reconoce el segmento de la cadena de entrada comprendido entre las posiciones  $i + 1$  y  $j$ , es decir,  $A \xrightarrow{*} a_{i+1} \dots a_j$ .

La estrategia ascendente que sigue el método es la siguiente: se inicia el reconocimiento con los símbolos terminales de la cadena de entrada (recogidos en el conjunto de hipótesis) y que constituyen las hojas de los árboles de análisis, a continuación, el algoritmo va construyendo los árboles de forma ascendente hasta que se alcanza el símbolo axioma.

Se puede especificar una regla de inferencia mediante un conjunto de pasos deductivos que cubran todas las instancias de inferencias. Un conjunto de reglas de inferencia, entonces, se denota mediante la unión de los correspondientes conjuntos de pasos deductivos. Así, el conjunto de pasos deductivos del sistema  $\text{IP}_{\text{CYK}}$  es:

$$\mathcal{D}_{\text{CYK}} = \mathcal{D}_{\text{CYK}}^{\text{Ini}} \cup \mathcal{D}_{\text{CYK}}^{\varepsilon} \cup \mathcal{D}_{\text{CYK}}^{\text{Comp}}$$

El paso deductivo  $\mathcal{D}_{\text{CYK}}^{\text{Ini}}$  es el que inicia el reconocimiento ascendente. Tiene como antecedente los ítems hipótesis y como consecuente los ítems relacionados con las producciones de la forma  $A \rightarrow a \in \mathcal{P}$  cuyo símbolo  $a$  sea igual al del antecedente:

$$\mathcal{D}_{\text{CYK}}^{\text{Ini}} = \frac{[a, i - 1, i]}{[A, i - 1, i]} \quad A \rightarrow a \in \mathcal{P}$$

El analizador también debería reconocer los símbolos  $\varepsilon$  entre todas las posibles posiciones de la cadena de entrada:

$$\mathcal{D}_{\text{CYK}}^{\varepsilon} = \frac{}{[A, i, i]} \quad A \rightarrow \varepsilon \in \mathcal{P}$$

Supongamos que existe una producción  $A \rightarrow BC \in \mathcal{P}$  y que los símbolos  $B$  y  $C$  ya han sido reconocidos, dominando  $B$  un fragmento de la cadena de entrada adjunto al que domina  $C$ . En este caso el algoritmo CYK, mediante el paso deductivo  $\mathcal{D}_{\text{CYK}}^{\text{Comp}}$ , combina el reconocimiento de ambos símbolos sobre un ítem asociado al símbolo  $A$  y que abarca el fragmento de la cadena de entrada dominado por  $B$  y  $C$ :

$$\mathcal{D}_{\text{CYK}}^{\text{Comp}} = \frac{\begin{array}{l} [B, i, j] \\ [C, j, k] \end{array}}{[A, i, k]} \quad A \rightarrow BC \in \mathcal{P}$$

El conjunto de ítems válidos es:

$$\mathcal{V}_{\text{CYK}} = \{[A, i, j] \mid A \xRightarrow{*} a_{i+1} \dots a_j\}$$

La demostración de que  $\mathcal{V}_{\text{CYK}}$  es el conjunto de ítems válidos de  $\mathbb{P}_{\text{CYK}}$  se encuentra en [126].

El conjunto de ítems finales y el subconjunto de ítems finales correctos son:

$$\mathcal{F}_{\text{CYK}} = \{[S, 0, n]\},$$

$$\mathcal{C}_{\text{CYK}} = \{[S, 0, n] \mid S \xRightarrow{*} a_1 \dots a_n\}.$$

El siguiente ejemplo es otro analizador ascendente, pero que elimina la restricción de CYK respecto a la forma de la gramática. Antes de pasar a describir el analizador, vamos a definir lo que es un ítem Earley.

**Definición 2.29** *Ítem Earley*

Dada una gramática  $G \in \text{CFG}$ , un ítem Earley tiene la forma

$$[A \rightarrow \alpha \bullet \beta, i, j] \text{ con } A \rightarrow \alpha\beta \in \mathcal{P}.$$

De manera informal, podemos decir que un ítem Earley representa el reconocimiento parcial de una producción. La parte de la producción situada a la izquierda del punto ( $\alpha$ ) es la parte reconocida, la cual domina el fragmento de la cadena de entrada entre las posiciones  $i$  y  $j$ . Mientras la parte situada a la derecha ( $\beta$ ) es lo que queda por reconocer de la producción.

**Ejemplo 2.2** *Bottom-up Earley*

El analizador bottom-up Earley (buE) reconoce el conjunto de ítems

$$\{[A \rightarrow \alpha \bullet \beta, i, j] \mid \alpha \xRightarrow{*} a_{i+1} \dots a_j\}$$

para una  $G \in \text{CFG}$  y  $a_1 \dots a_n \in V_T^*$ .

El lado derecho de la producción se analiza moviendo el punto paso a paso a la derecha mediante el reconocimiento del símbolo que se encuentra inmediatamente a la derecha del punto. Si el punto se encuentra en la posición derecha de la producción ( $\beta = \varepsilon$ ), se observa que el ítem sería equivalente a  $[A, i, j]$  en el analizador CYK.

Dadas una gramática  $G = (V_N, V_T, P, S) \in \text{CFG}$  y una cadena  $a_1 \dots a_n$ , un sistema de análisis bottom-up Earley se define:

$$\mathbb{P}_{\text{buE}} = \langle \mathcal{I}_{\text{buE}}, H, \mathcal{D}_{\text{buE}} \rangle$$

donde las hipótesis se definen de la forma habitual y el conjunto de ítems es:

$$\mathcal{I}_{\text{buE}} = \{[A \rightarrow \alpha \bullet \beta, i, j] \mid A \rightarrow \alpha\beta \in \mathcal{P}, 0 \leq i \leq j\}$$

El conjunto de pasos deductivos viene dado por:

$$\mathcal{D}_{\text{buE}} = \mathcal{D}_{\text{buE}}^{\text{Ini}} \cup \mathcal{D}_{\text{buE}}^{\text{Scan}} \cup \mathcal{D}_{\text{buE}}^{\varepsilon} \cup \mathcal{D}_{\text{buE}}^{\text{Comp}}$$

El paso  $\mathcal{D}_{\text{buE}}^{\text{Ini}}$  es el que inicia el reconocimiento ascendente. Al no existir restricción en la forma de la gramática, los símbolos terminales, por los que iniciaba el reconocimiento el algoritmo CYK mediante las hipótesis, se encuentran dentro de las producciones. Esto obliga al analizador a lanzar el reconocimiento de todas las producciones de la gramática desde todas las posiciones posibles de la cadena de entrada:

$$\mathcal{D}_{\text{buE}}^{\text{Ini}} = \overline{[A \rightarrow \bullet \nu, i, i]}$$

Mediante el paso deductivo  $\mathcal{D}_{\text{buE}}^{\text{Scan}}$  se efectúa la lectura de un terminal de la cadena de entrada cuando el reconocimiento lo ha alcanzado. Si dicho terminal es la cadena vacía, entonces se aplica el paso  $\mathcal{D}_{\text{buE}}^{\varepsilon}$ :

$$\mathcal{D}_{\text{buE}}^{\text{Scan}} = \frac{\begin{array}{l} [A \rightarrow \alpha \bullet a\beta, i, j] \\ [a, j, j + 1] \end{array}}{[A \rightarrow \alpha a \bullet \beta, i, j + 1]}$$

$$\mathcal{D}_{\text{buE}}^{\varepsilon} = \frac{[A \rightarrow \alpha \bullet \varepsilon\beta, i, j]}{[A \rightarrow \alpha \varepsilon \bullet \beta, i, j]}$$

Una vez que se ha completado el reconocimiento de un subárbol, se debe continuar el reconocimiento ascendente. El algoritmo obliga a que el reconocimiento hasta dicho subárbol se haya completado, ya que este analizador efectúa un recorrido unidireccional de izquierda a derecha a través de la cadena de entrada. El paso  $\mathcal{D}_{\text{buE}}^{\text{Comp}}$  lleva a cabo esta tarea:

$$\mathcal{D}_{\text{buE}}^{\text{Comp}} = \frac{\begin{array}{l} [A \rightarrow \alpha \bullet B\beta, i, j] \\ [B \rightarrow \nu \bullet, j, k] \end{array}}{[A \rightarrow \alpha B \bullet \beta, i, k]}$$

El conjunto de ítems válidos es:

$$\mathcal{V}_{\text{buE}} = \{[A \rightarrow \alpha \bullet \beta, i, j] \mid \alpha \xrightarrow{*} a_{i+1} \dots a_j\}$$

La demostración de que  $\mathcal{V}_{\text{buE}}$  es el conjunto de ítems válidos de  $\text{IP}_{\text{buE}}$  se encuentra en [126].

El conjunto de ítems finales y el subconjunto de ítems finales correctos son:

$$\mathcal{F}_{\text{buE}} = \{[S \rightarrow \nu \bullet, 0, n]\},$$

$$\mathcal{C}_{\text{buE}} = \{[S \rightarrow \nu \bullet, 0, n] \mid S \xrightarrow{*} a_1 \dots a_n\}.$$

El analizador buE presenta un serio problema de eficiencia, debido fundamentalmente a la gran cantidad de ítems que se lanzan mediante el paso deductivo  $\mathcal{D}_{\text{buE}}^{\text{Ini}}$ . Este paso lanza la predicción de todos los posibles subárboles que se pueden construir con la estrategia ascendente, la mayoría de los cuales no conducen a un árbol de análisis correcto.



El siguiente ejemplo es la especificación en forma de esquema de análisis del analizador más conocido para CFG: el algoritmo de Earley [48, 49]. Se trata de un analizador ascendente con reconocimiento de la cadena de entrada de izquierda a derecha, al igual que el ejemplo anterior. Sin embargo, este algoritmo incorpora información predictiva con objeto de filtrar el gran número de ítems lanzados por el paso  $\mathcal{D}_{\text{buE}}^{\text{Ini}}$ .

### Ejemplo 2.3 Earley

Dadas una gramática  $G = (V_N, V_T, P, S) \in \text{CFG}$  y una cadena  $a_1 \dots a_n$ , un sistema de análisis Earley se define:

$$\mathbb{P}_{\text{Earley}} = \langle \mathcal{I}_{\text{Earley}}, H, \mathcal{D}_{\text{Earley}} \rangle$$

donde  $H$  se define de la forma habitual y el dominio del sistema viene dado por:

$$\mathcal{I}_{\text{Earley}} = \{[A \rightarrow \alpha \bullet \beta, i, j] \mid A \rightarrow \alpha\beta \in \mathcal{P} \wedge 0 \leq i \leq j\}$$

El conjunto de pasos deductivos es:

$$\mathcal{D}_{\text{Earley}} = \mathcal{D}_{\text{Earley}}^{\text{Ini}} \cup \mathcal{D}_{\text{Earley}}^{\text{Scan}} \cup \mathcal{D}_{\text{Earley}}^{\varepsilon} \cup \mathcal{D}_{\text{Earley}}^{\text{Pred}} \cup \mathcal{D}_{\text{Earley}}^{\text{Comp}}$$

El paso  $\mathcal{D}_{\text{Earley}}^{\text{Ini}}$  añade el axioma que se necesita para iniciar el análisis, añadiéndolo a las hipótesis obtenidas para la cadena de entrada:

$$\mathcal{D}_{\text{Earley}}^{\text{Ini}} = \overline{[S \rightarrow \bullet \nu, 0, 0]}$$

El paso  $\mathcal{D}_{\text{Earley}}^{\text{Pred}}$  realiza la predicción de análisis de un subárbol. Es, junto a  $\mathcal{D}_{\text{Earley}}^{\text{Ini}}$ , el que introduce información predictiva (descendente) para guiar la construcción ascendente del árbol de análisis. Para ello, una vez que el reconocimiento llega al nodo raíz de un subárbol, esta regla lanza los ítems asociados a las producciones cuyo lado izquierdo sea dicho nodo.

$$\mathcal{D}_{\text{Earley}}^{\text{Pred}} = \frac{[A \rightarrow \alpha \bullet B\beta, i, j]}{[B \rightarrow \bullet \nu, j, j]}$$

Las pasos deductivos restantes son iguales a los del sistema  $\mathbb{P}_{\text{buE}}$ :

$$\mathcal{D}_{\text{Earley}}^{\text{Scan}} = \mathcal{D}_{\text{buE}}^{\text{Scan}}$$

$$\mathcal{D}_{\text{Earley}}^{\varepsilon} = \mathcal{D}_{\text{buE}}^{\varepsilon}$$

$$\mathcal{D}_{\text{Earley}}^{\text{Comp}} = \mathcal{D}_{\text{buE}}^{\text{Comp}}$$

El conjunto de ítems válidos es:

$$\mathcal{V}_{\text{Earley}} = \{[A \rightarrow \alpha \bullet \beta, i, j] \mid \alpha \xrightarrow{*} a_{i+1} \dots a_j, S \xrightarrow{*} a_1 \dots a_i A \nu\}$$

La demostración de que  $\mathcal{V}_{\text{Earley}}$  es el conjunto de ítems válidos de  $\mathbb{P}_{\text{Earley}}$  se encuentra en [126].

Los conjuntos de ítems finales e ítems finales correctos son iguales a los del sistema anterior:

$$\mathcal{F}_{\text{Earley}} = \mathcal{F}_{\text{buE}}$$

$$\mathcal{C}_{\text{Earley}} = \mathcal{C}_{\text{buE}}$$

## 2.5 Complejidad

Como ya hemos dicho, un sistema de análisis no es más que la especificación abstracta de un algoritmo de análisis sintáctico. Por tanto, como en cualquier algoritmo, es un factor esencial determinar su complejidad computacional con objeto de poder elegir un analizador lo más eficiente posible.

De manera informal, podemos decir que la complejidad computacional de un analizador es el tiempo y espacio (de memoria) necesarios para que analice una cadena de entrada. Aunque en la literatura no se hace esta distinción, nosotros vamos a diferenciar dos tipos de complejidad:

### Complejidad teórica .

Nos da la medida en el mejor, peor y caso medio de la complejidad de un algoritmo y se obtiene aplicando modelos matemáticos. La complejidad teórica viene especificada por dos medidas: (i) la de la complejidad temporal, o tiempo necesario para realizar el análisis en base a una determinada unidad y, (ii) la de la complejidad espacial, o espacio de memoria de computador necesaria para el análisis. Es una práctica habitual, indicar solo la complejidad en el caso peor, y es lo que nosotros haremos a lo largo del texto. De este tipo de complejidad nos vamos a ocupar en esta sección.

### Complejidad empírica .

Nos da la medida en casos prácticos. No existe una forma estándar para calcularla, pero la manera más común consiste en llevar a cabo una implementación del algoritmo y aplicarle una batería de cadenas de entrada para distintas gramáticas. La complejidad empírica no es una medida formal y, por tanto, va a depender de factores de implementación, como el hardware sobre el que se ejecuta, el lenguaje de programación e implementación elegidos o la batería de pruebas seleccionadas. A pesar de todo, es de gran importancia conocer el comportamiento de un analizador en casos prácticos, ya que la complejidad teórica no nos da la medida para una aplicación práctica determinada. Por todo esto, hemos dedicado algunas secciones de este trabajo a evaluar el comportamiento de algunos de los algoritmos de análisis que proponemos en él.

De todos es conocido que el cálculo de la complejidad teórica, tanto temporal como espacial, no es una tarea en absoluto trivial cuando tenemos un analizador descrito en forma algorítmica [14]. En la mayoría de las ocasiones este cálculo es cuestión de juicio, intuición y experiencia. Sin embargo, la especificación abstracta de los algoritmos de análisis mediante los esquemas de análisis facilitan enormemente la determinación de la complejidad de los mismos. Ésto es debido a dos razones fundamentales: la primera es que el uso de esquemas nos abstrae de las estructuras de control y datos que se emplearán en la implementación, y la segunda, la propia forma de los ítems.

En los ejemplos anteriores hemos visto la forma que tienen los ítems en los esquemas de análisis. Concretamente, un ítem consta de: (i) un componente que indica la restricción sobre los árboles de gramática que representa dicha clase de equivalencia, y que normalmente es una regla de producción punteada de la gramática; (ii) una lista de expresiones enteras, llamadas índices, que dan alguna información sobre posiciones en la cadena de entrada. Se puede observar que, mientras el primer componente depende exclusivamente de la gramática, el segundo solo depende de la cadena de entrada. Teniendo en cuenta todo lo dicho, es relativamente sencillo determinar la complejidad de un esquema de análisis.

Respecto al espacio de memoria necesario para almacenar un análisis, en el peor de los casos necesitaríamos espacio para almacenar todos los distintos ítems que se pueden generar. Dada una gramática  $G$  y una cadena de entrada de longitud  $n$ , como cada índice puede tomar valores entre 0 y  $n$ , los índices se pueden combinar de  $n^k$  formas, donde  $k$  es el número de índices que presenta un ítem. Por tanto, la complejidad espacial será  $\mathcal{O}(|G|n^k)$ , donde el factor multiplicativo  $|G|$  lo introduce el primer componente del ítem y depende del tamaño de la gramática. Este factor  $|G|$  no se puede despreciar, ya que en gramáticas de lenguajes naturales presenta valores muy grandes. Aún así, en la mayoría de las ocasiones, vamos a obviarlo cuando estudiemos la complejidad de los analizadores.

La complejidad temporal requiere un estudio más detallado. La complejidad en el caso peor viene dada por la complejidad del paso deductivo de mayor coste temporal. Para calcular el coste de un paso deductivo es necesario establecer cómo se combinan los índices en la misma. Dada una gramática  $G$  y una cadena de entrada de longitud  $n$ , una primera aproximación podría ser  $\mathcal{O}(|G|^p n^q)$ , donde  $p$  es el número de antecedentes de la regla deductiva y  $q$  es el número de valores repetidos en los índices de los antecedentes. Por ejemplo, para el paso deductivo:

$$\mathcal{D}_{\text{buE}}^{\text{Comp}} = \frac{\begin{array}{l} [A \rightarrow \alpha \bullet B\beta, i, j] \\ [B \rightarrow \nu \bullet, j, k] \end{array}}{[A \rightarrow \alpha B \bullet \beta, i, k]}$$

su complejidad temporal es  $\mathcal{O}(n^3)$ , ya que presenta repetidos los índices  $i, j$  y  $k$ . De nuevo omitimos el factor inducido por el tamaño de la gramática.

**Ejemplo 2.4** *Complejidad de los algoritmos CYK, buE y Earley*

*La complejidad espacial en los tres casos es  $\mathcal{O}(n^2)$ , debido a que todos tienen dos índices en los ítems.*

*La complejidad temporal en los tres casos es de  $\mathcal{O}(n^3)$ , debido a que cada paso de compleción incorpora tres valores repetidos.*

Bajo determinadas circunstancias es posible reducir la complejidad temporal de un paso deductivo mediante la denominada *aplicación parcial*, que consiste básicamente en la división del mismo en una secuencia de pasos intermedios. Las únicas condiciones exigibles son: (1) que todos los antecedentes y condiciones del paso original deben aparecer al menos en uno de los pasos intermedios y (2) que el consecuente del paso original debe coincidir con el

del último paso intermedio. Veremos algún ejemplo de *aplicación parcial* cuando definamos esquemas de análisis sintáctico para TAG.

## 2.6 Relaciones entre esquemas de análisis

El uso de sistemas de análisis para la especificación de analizadores sintácticos nos permite explotar todas las propiedades de los sistemas deductivos, entre otras, la posibilidad de establecer relaciones entre distintos sistemas. El establecer las relaciones formales entre los analizadores tiene una serie de ventajas según el uso que hagamos de ellas:

- **Método descriptivo**  
Nos permite crear una red de analizadores donde se puede comprobar, desde un alto nivel de abstracción, cómo se relacionan distintos analizadores, que en principio, pueden parecer que no poseen nada en común.
- **Método generativo**  
Una vez definido un analizador, podemos crear un nuevo analizador aplicándole alguna de las relaciones. En este caso, entendemos las relaciones como métodos de transformación y, por tanto, como un método sistemático para la creación de nuevos analizadores a partir de otros conocidos.
- **Método para determinar la corrección**  
Una de las ventajas de formalizar las relaciones entre esquemas es que nos permite establecer qué propiedades respecto a la corrección preserva cada tipo de relación. De manera que si un analizador se relaciona con otro, cuya corrección está probada, podemos deducir ciertas propiedades del primero a partir de la relación que mantiene con el segundo.

En esta sección introduciremos los conceptos teóricos necesarios para definir las relaciones que se pueden establecer entre esquemas de análisis y las propiedades que cumplen.

### Notación 3

- Consideraremos los siguientes esquemas:  $\mathbf{P}_1 = (\mathcal{I}_1, \mathcal{H}, \mathcal{D}_1)$  y  $\mathbf{P}_2 = (\mathcal{I}_2, \mathcal{H}, \mathcal{D}_2)$ ;
- $\vdash_1$  y  $\vdash_2$  son las relaciones de inferencia definidas sobre los esquemas  $\mathbf{P}_1$  y  $\mathbf{P}_2$ ;
- $\mathcal{V}_1$  y  $\mathcal{V}_2$  son los conjuntos de ítems válidos de los esquemas  $\mathbf{P}_1$  y  $\mathbf{P}_2$ .

### Definición 2.30 Función regular entre ítems

Una función  $f : \mathcal{I}_1 \rightarrow \mathcal{I}_2$  es una función regular entre ítems si para todo ítem  $\iota \in \mathcal{I}_1$  y para todo árbol  $t \in \iota$ , se verifica que  $t \in f(\iota)$ .

Vamos a generalizar esta definición para que sea aplicable a conjuntos de ítems, pasos deductivos y secuencias deductivas.

**Definición 2.31** *Función entre conjunto ítems*

Dado un conjunto  $Y \subseteq \mathcal{I}_1$ , una función entre conjunto ítems se define mediante:

$$f(Y) = \{\xi \in \mathcal{I}_2 \mid \exists \eta \in Y : f(\eta) = \xi\}$$

**Definición 2.32** *Función entre pasos deductivos*

Dado un paso deductivo  $\eta_1 \dots \eta_k \vdash \xi \in \mathcal{D}_1$ , una función entre pasos deductivos se define mediante:

$$f(\eta_1 \dots \eta_k \vdash \xi) = f(\eta_1) \dots f(\eta_k) \vdash f(\xi)$$

Se asume que: (1) el conjunto de hipótesis es disjunto con respecto al conjunto de los ítems en  $\mathcal{I}_1$  y  $\mathcal{I}_2$ , y (2)  $f(h) = h$  para todo  $h \in \mathcal{H}$ .

**Definición 2.33** *Función entre secuencias deductivas*

Si tenemos en cuenta que  $\Delta_{\mathbf{P}_2} = Y_2 \vdash_2 x_1 \vdash_2 \dots \vdash_2 x_j$  y  $\Delta_{\mathbf{P}_1} = Y_1 \vdash_1 x'_1 \vdash_1 \dots \vdash_1 x'_j$ , una función  $f(\Delta_{\mathbf{P}_1}) = \Delta_{\mathbf{P}_2}$  es una función entre secuencias deductivas si y solo si se cumple  $Y_1 \in \wp(\mathcal{H} \cup \mathcal{I}_1)$  con  $f(Y_1) = Y_2$  y  $x'_1, \dots, x'_j \in \mathcal{I}_1$  con  $f(x'_i) = x_i$ .

Vamos a describir a continuación de manera informal los distintos tipos de relaciones que se pueden establecer entre dos esquemas, las cuales se pueden dividir en dos grupos: generalizaciones y filtros.

**Generalizaciones**

Un esquema es la generalización de otro cuando es fruto de un refinamiento y/o una extensión. Por tanto, las generalizaciones incluyen los refinamientos y la extensión.

- Refinamientos

Introducen más detalles en el analizador con objeto de obtener mejoras cualitativas en el mismo. En este grupo se sitúan los siguientes tipos de relaciones:

- Refinamiento de ítems

Cuando los ítems de un esquema se dividen en varios ítems para obtener otro esquema. Este cambio en el conjunto de ítems puede requerir una modificación del conjunto de pasos deductivos del nuevo esquema para adaptarlo al nuevo dominio.

La relación inversa al refinamiento de ítems se denomina *contracción de ítems*, y consiste en agrupar en un sólo ítem de un esquema varios ítems de otro.

- Refinamiento de pasos deductivos

Cuando un paso deductivo de un esquema de análisis es descompuesto en varios pasos para obtener otro esquema. Este cambio en el conjunto de pasos deductivos puede requerir una modificación del dominio del nuevo esquema.

- Extensión

Cuando un esquema se obtiene ampliando la clase de gramáticas sobre la que está definido otro esquema.

### Filtros

Introducen mejoras cuantitativas en el analizador mediante la eliminación de ítems de su dominio o la reducción de las secuencias deductivas. Existen tres tipos:

- Filtro estático

Cuando se eliminan ítems y/o pasos deductivos redundantes de un esquema para obtener un nuevo esquema. Produce una optimización en *tiempo de compilación* y son independientes de la cadena de entrada.

- Filtro dinámico

Cuando se introduce información contextual en un esquema, mediante la adición de nuevos antecedentes en los pasos deductivos. De esta forma el reconocimiento de ítems durante el proceso deductivo se puede hacer dependiendo de la existencia de otros ítems. Este tipo de filtro, a diferencia del anterior, genera optimizaciones en *tiempo de ejecución* y van a depender de la cadena de entrada.

- Contracción de secuencias deductivas

Cuando una secuencia deductiva de un esquema se sustituye por otra de menor longitud. Se trata de la relación inversa al refinamiento de pasos deductivos.

Pasemos ahora a definir cada una de estas relaciones, así como las propiedades de interés que se derivan de ellas.

#### Definición 2.34 Refinamiento de ítems

El esquema  $P_2$  es un refinamiento de los ítems del esquema  $P_1$ , y lo denotamos como  $P_1 \xrightarrow{\text{ir}} P_2$ , si existe una función regular entre ítems  $f: \mathcal{I}_2 \rightarrow \mathcal{I}_1$  tal que:

1.  $\mathcal{I}_1 = f(\mathcal{I}_2)$

2.  $\Delta_{P_1} = f(\Delta_{P_2})$

#### Corolario 2.1

La relación  $\xrightarrow{\text{ir}}$  es reflexiva y transitiva.

Además si  $P_1 \xrightarrow{\text{ir}} P_2$ , entonces la corrección del esquema  $P_2$  implica la corrección de  $P_1$ .

La relación inversa al refinamiento de ítems se denomina contracción de ítems y la denotamos como  $\xrightarrow{\text{ic}}$ . Por tanto, si se verifica  $P_2 \xrightarrow{\text{ic}} P_1$ , entonces  $P_1 \xrightarrow{\text{ir}} P_2$ .

#### Corolario 2.2

La relación  $\xrightarrow{\text{ic}}$  es reflexiva, transitiva y preserva la corrección.



**Definición 2.35** *Refinamiento de pasos deductivos*

El esquema  $P_2$  es un refinamiento de los pasos deductivos del esquema  $P_1$ , y lo denotamos como  $P_1 \xrightarrow{sr} P_2$ , si:

1.  $\mathcal{I}_1 \subseteq \mathcal{I}_2$
2.  $\vdash_1^* \subseteq \vdash_2^*$

**Corolario 2.3**

La relación  $\xrightarrow{sr}$  es reflexiva, transitiva y preserva la completitud.

**Definición 2.36** *Extensión*

Sea  $P_1$  un esquema definido sobre una clase de gramáticas  $CG_1$  y  $P_2$  un esquema definido sobre una clase de gramáticas  $CG_2$ , decimos que  $P_2$  es una extensión del esquema  $P_1$ , y lo denotamos como  $P_1 \xrightarrow{ext} P_2$ , si:

1.  $CG_1 \subseteq CG_2$
2.  $P_1(G)(a_1 \dots a_n) = P_2(G)(a_1 \dots a_n)$  para toda  $G \in CG_1$  y cadena de entrada  $a_1 \dots a_n$ .

**Corolario 2.4**

La relación  $\xrightarrow{ext}$  es reflexiva y transitiva.

**Definición 2.37** *Filtro estático*

El esquema  $P_2$  es un filtro estático del esquema  $P_1$ , y lo denotamos como  $P_1 \xrightarrow{sf} P_2$ , si:

1.  $\mathcal{I}_1 \supseteq \mathcal{I}_2$
2.  $\mathcal{D}_1 \supseteq \mathcal{D}_2$

**Definición 2.38** *Filtro dinámico*

El esquema  $P_2$  es un filtro dinámico del esquema  $P_1$ , y lo denotamos como  $P_1 \xrightarrow{df} P_2$ , si:

1.  $\mathcal{I}_1 \supseteq \mathcal{I}_2$
2.  $\vdash_1 \supseteq \vdash_2$

**Definición 2.39** *Contracción de secuencias deductivas*

El esquema  $P_2$  es una contracción de secuencias deductivas del esquema  $P_1$ , y lo denotamos como  $P_1 \xrightarrow{sc} P_2$ , si:

1.  $\mathcal{I}_1 \supseteq \mathcal{I}_2$
2.  $\vdash_1^* \supseteq \vdash_2^*$

**Corolario 2.5**

Los filtros  $\xrightarrow{sf}$ ,  $\xrightarrow{df}$  y  $\xrightarrow{sc}$  son relaciones reflexivas, transitivas y preservan la consistencia. Y mantienen las siguientes relaciones:

$$\xrightarrow{sf} \subseteq \xrightarrow{df} \subseteq \xrightarrow{sc}$$

## Parte II

# Gramáticas de adjunción de árboles

## Capítulo 3

# Gramáticas de adjunción de árboles

En este capítulo vamos a hacer un breve repaso a las características fundamentales de las gramáticas de adjunción de árboles. En primer lugar mostraremos su definición básica y algunas extensiones, a continuación presentaremos sus propiedades fundamentales, y concluiremos citando un conjunto de formalismos relacionados con ellas.

### 3.1 Definición

Las gramáticas de adjunción de árboles (TAG, *Tree Adjoining Grammar*) es un formalismo gramatical cuya característica fundamental radica en el uso de árboles como estructuras elementales, a diferencia de otros formalismos, como las gramáticas incontextuales, que usan reglas de producción. La operación básica de composición de árboles que emplea las TAG se denomina *adjunción*. El uso de árboles como estructuras elementales aumenta de manera considerable el conjunto de propiedades lingüísticas que presenta el formalismo a la hora de definir gramáticas de lenguajes naturales, mientras el empleo de la operación de adjunción aporta a las TAGs la *sensibilidad al contexto* necesaria para reconocer construcciones lingüísticas de lenguajes naturales que no se pueden reconocer mediante CFG.

La definición original de las gramáticas de adjunción de árboles fue introducida por Joshi, Levy y Takahashi en [69], estableciéndose la definición actual en [68].

**Definición 3.1** *Gramática de adjunción de árboles TAG*

*Una gramática de adjunción de árboles es una quintupla  $(V_T, V_N, I, A, S)$ , donde:*

- $V_T$  es un conjunto finito de símbolos terminales,
- $V_N$  es un conjunto finito de símbolos no terminales,
- $I$  es un conjunto finito de árboles finitos denominado árboles iniciales,



- $A$  es un conjunto finito de árboles finitos denominado árboles auxiliares,
- $S \in V_N$  es el axioma de la gramática.

Al conjunto  $I \cup A$  se le denomina *árboles elementales*. Las restricciones que presenta el formalismo sobre cada tipo de árbol elemental son las siguientes:

- Los nodos interiores de los árboles elementales se etiquetan con símbolos no terminales.
- Las raíces de los árboles iniciales se etiquetan con el axioma y sus hojas con símbolos no terminales o la cadena vacía, es decir, si un árbol  $\alpha \in I$  entonces  $root(\alpha) = S$  y  $yield(\alpha) = V_T^*$ .
- Las raíces de los árboles auxiliares se pueden etiquetar con cualquier símbolo no terminal y sus hojas se etiquetan con símbolos no terminales o la cadena vacía, excepto el un nodo, denominado *nodo pie*, cuya etiqueta coincide con la de la raíz. Al camino de la raíz al pie de un árbol auxiliar se le denomina *espina*.

Según el tipo de la cosecha que produce un árbol elemental, se pueden clasificar en nulos y propios. Se denomina *árbol inicial nulo* a todo árbol inicial cuya cosecha es la palabra vacía. Se denomina *árbol auxiliar nulo* a todo árbol auxiliar cuya cosecha, exceptuando el nodo pie, es la palabra vacía. Se denomina *árbol inicial propio* a todo árbol inicial cuya cosecha contiene algún símbolo terminal. Se denomina *árbol auxiliar propio* a todo árbol auxiliar cuya cosecha contiene algún símbolo terminal. Para evitar gramáticas infinitamente ambiguas, las gramáticas de adjunción de árboles no permiten el uso de los árboles auxiliares nulos.

La *adjunción* es la operación básica que incorpora el formalismo TAG para componer *árboles derivados* a partir de los árboles elementales de la gramática. Dado un nodo de un árbol  $\gamma$  (elemental o derivado), al que se denomina *nodo de adjunción*, y un árbol auxiliar  $\beta$  cuya raíz está etiquetada con el mismo símbolo que dicho nodo, la operación de adjunción construye un nuevo árbol derivado de la siguiente manera (ver figura 3.1): (1) se poda el subárbol de  $\gamma$  que pende del nodo de adjunción, (2) se cuelga el árbol auxiliar  $\beta$  del nodo de adjunción, identificando la raíz de  $\beta$  con dicho nodo, (3) se cuelga del nodo pie de  $\beta$  el subárbol escindido de  $\gamma$ , identificando la raíz de este subárbol con el nodo pie de  $\beta$ .

La definición de adjunción sólo requiere que el nodo de adjunción presente la misma etiqueta que la raíz de un árbol auxiliar para que se pueda llevar a cabo la adjunción. Sin embargo, se puede limitar el conjunto de árboles auxiliares que se pueden adjuntar en un nodo mediante las denominadas *restricciones de adjunción*. Las restricciones de adjunción que se pueden establecer sobre un nodo son de tres tipos: *selectiva* (SA) si se especifica el conjunto de árboles auxiliares que se pueden adjuntar (la adjunción no es obligatoria), *nula* (NA) si no se permite la adjunción, y *obligatoria* (OA) si se especifica el conjunto de árboles auxiliares que se pueden adjuntar (la adjunción de uno de ellos es obligatoria).

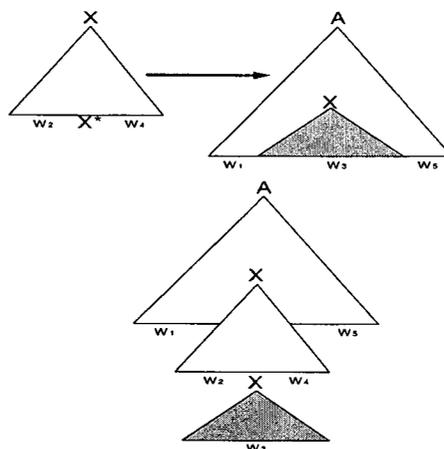


Figura 3.1: Operación de adjunción

Una TAG se dice que está lexicalizada (LTAG, *Lexicalized Tree Adjoining Grammar*) [108] si todo árbol elemental presenta al menos un símbolo terminal destacado en su frontera, al cual se denomina *ancla*. Entre las ventajas que se obtienen de la lexicalización de las TAG, podemos citar la introducción de técnicas basadas en *super-etiquetas* [72] para mejorar la eficiencia de los analizadores sintácticos.

En el contexto de las LTAG, además de la adjunción, se suele introducir la *sustitución* como operación de composición adicional. La sustitución en las LTAG funciona de forma análoga a la sustitución en las CFG, pero teniendo en cuenta que los elementos que se combinan no son reglas sino árboles. Dado un árbol inicial  $\alpha$  cuya raíz presenta la misma etiqueta que un nodo que se encuentra en la frontera de un árbol elemental  $\gamma$ , al cual se denomina *nodo de sustitución* y se denota gráficamente acompañándolo de un símbolo  $\downarrow$ , la operación de sustitución (ver figura 3.2) consiste en colgar el árbol  $\alpha$  de dicho nodo, identificándolo con la raíz de  $\alpha$ .

Las diferencias entre la definición de una gramática TAG lexicalizada con la operación de sustitución y la definición anteriormente presentada de una TAG son básicamente tres: no se permiten árboles elementales nulos, las raíces de los árboles iniciales pueden estar etiquetadas con cualquier símbolo no terminal de la gramática y en la frontera de los árboles elementales pueden existir símbolos no terminales (además del nodo pie), aunque éstos deben estar marcados para sustitución. Además todos los nodos marcados para sustitución presentan restricción de sustitución obligatoria y de adjunción nula.

Un *árbol derivado* es el que se obtiene mediante la aplicación de sucesivas operaciones de adjunción y sustitución, si se trata de una LTAG, sobre un árbol elemental. Dada una gramática de adjunción de árboles  $G$ , si denotamos como  $T(G)$  el conjunto de árboles derivados a partir de los árboles iniciales cuyas raíces están etiquetadas con el axioma de la gramática y que cumplen todas las restricciones de adjunción y sustitución, el conjunto de



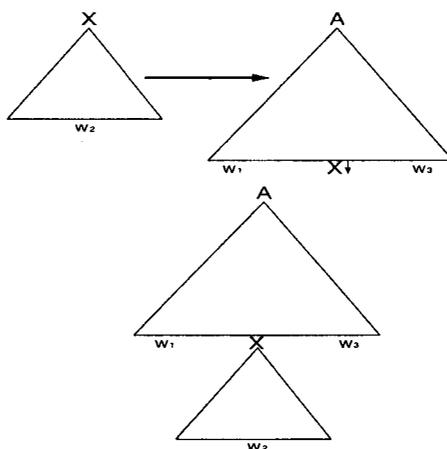


Figura 3.2: Operación de sustitución

cosechas de  $T(G)$  constituye el lenguaje definido por la gramática.

Mientras que en las CFG los árboles derivados contienen la información suficiente para determinar como se construyeron, en las TAG es necesario almacenar dicha información en un nuevo tipo de árbol, denominado *árbol de derivación* [103]. Un árbol de derivación mantiene la historia de las adjunciones que se han efectuado para obtener un árbol derivado. Para ello los nodos de estos árboles se etiquetan con denominaciones de árboles elementales, de modo que su nodo raíz siempre corresponde a un árbol inicial y el resto a árboles auxiliares<sup>1</sup>. Cada nodo del árbol de derivación se acompaña del nodo, habitualmente con notación de Görn [55], sobre el que se efectuó la adjunción. Véase un ejemplo en la figura 3.3.

## 3.2 Propiedades lingüísticas y computacionales

Aquí haremos un somero repaso a las propiedades más relevantes de las gramáticas de adjunción de árboles. Un estudio más detallado se puede encontrar en [71].

Entre las principales aportaciones que hace este formalismo desde el punto de vista lingüístico [79] podemos destacar dos: *dominio de localidad extendido* y *factorización de la recursión del dominio de dependencias*. La primera es consecuencia directa del uso de árboles como estructuras básicas de composición, lo que posibilita establecer dependencias de mayor distancia entre elementos gramaticales de la misma estructura que las que permiten otros formalismos basados en reglas. Mientras la segunda se deriva del uso de la adjunción como operación básica de composición, la cual hace posible transformar una dependencia local entre dos elementos de un árbol elemental en una dependencia de larga distancia, mediante la inclusión de árboles auxiliares entre ellos [78].

Y entre las propiedades computacionales citaremos las siguientes:

<sup>1</sup>También pueden ser árboles iniciales cuando se incluye la operación de sustitución.

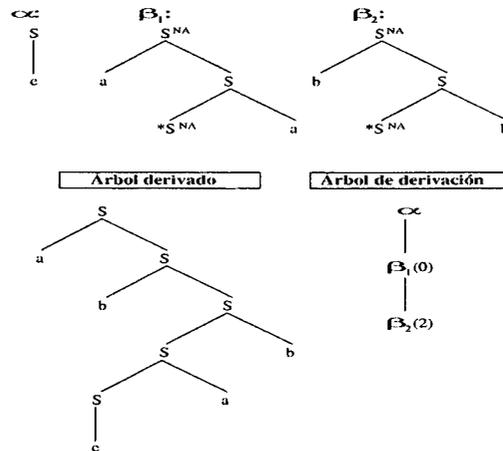


Figura 3.3: Árbol de derivación

- Los lenguajes de adjunción de árboles forman una familia abstracta de lenguajes completa (*Full AFL*) [136].
- Los lenguajes incontextuales están estrictamente contenidos en los lenguajes de adjunción de árboles ( $CFL \subset TAG$ ), aunque las estructuras sintácticas que se generan con TAG son imposible de conseguir con CFG [68].
- El análisis sintáctico de TAG se puede realizar en tiempo polinomial.
- Los lenguajes de adjunción árboles permiten capturar dependencias anidadas y ciertas dependencias cruzadas (ver figura 3.4).
- Los lenguajes de adjunción de árboles son semilineales [142].

Por cumplir las cuatro últimas propiedades se puede concluir que las TAG forman parte de los formalismos suavemente dependientes del contexto.

### 3.3 Formalismos relacionados

Además de las LTAG, muchos son los formalismos que han sido derivados a partir de las gramáticas de adjunción de árboles. En esta sección, sin ánimo de ser exhaustivos, vamos a citar los que consideramos más significativos.

#### TAG basadas en estructuras de rasgos (FTAG, *Feature structures based TAG*)

Presentadas en [137]. Asocian *estructuras de rasgos* de tamaño finito a los nodos de los árboles elementales, de manera que las operaciones de adjunción y sustitución deben ir acompañadas de un proceso de unificación de dichos rasgos. En la literatura podemos encontrar gramáticas FTAG de amplia cobertura para el inglés [47, 144] y el francés [1],



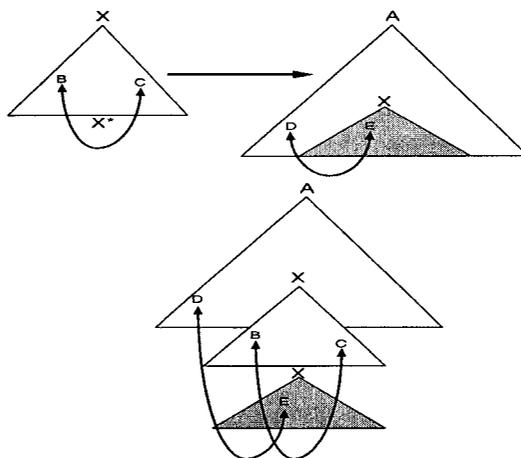


Figura 3.4: Dependencias cruzadas mediante operación de adjunción

así como gramáticas parciales de varios idiomas, como la que presentamos en [18, 19] para describir los fenómenos lingüísticos asociados a los adjetivos del español.

**LTAG estocásticas (SLTAG, *Stochastic Lexicalized Tree Adjoining Grammar*)**

Presentadas en [110]. Introducen la probabilidad de que una adjunción o sustitución ocurra en cada nodo de las LTAG con objeto de aumentar la eficiencia del análisis.

**TAG multicomponente (MCTAG, *Multicomponent Tree Adjoining Grammar*)**

Presentadas en [69]. Permiten adjuntar un conjunto de árboles auxiliares sobre un nodo o un conjunto de nodos. Las MCTAG se clasifican en cuatro grupos en función de la forma en que definen la operación de adjunción [142]. Un completo estudio sobre las MCTAG y su análisis sintáctico se puede encontrar en [13].

**TAG sincronas (*Synchronous Tree Adjoining Grammar*)**

Presentadas en [123]. Establecen correspondencias entre pares de lenguajes mediante la aplicación sincronizada de las operaciones de composición sobre las dos gramáticas TAG que los describen. Entre sus campos de aplicación se encuentra la generación de lenguajes [124], aunque su uso está más extendido en la traducción automática [2].

**Gramáticas de descripción de árboles (DTG, *D-Tree Grammar*)**

Presentadas en [99]. Se trata de un formalismo lingüísticamente motivado y con una concepción similar a las TAG [133], ya que emplea árboles (*árboles de descripción, D-tree*) como estructuras básicas de composición, aunque los lenguajes que generan no son comparables a los de las TAG [134].

A las gramáticas de inserción de árboles, dado el especial interés que revisten para el trabajo desarrollado en esta memoria, le dedicaremos un capítulo completo.

## Capítulo 4

# Algoritmos de análisis sintáctico para TAG

Ya en la introducción hicimos una panorámica de la historia del análisis sintáctico de TAG. En este capítulo vamos a detallar, usando los esquemas de análisis sintáctico, los esquemas que constituyen el esqueleto de la red de analizadores para TAG cuyo núcleo fundamental es el algoritmo Earley.

Los esquemas de análisis sintáctico que hemos definido están basados en ítems, lo que hace imprescindible establecer la forma que tendrán los mismos cuando se definan esquemas para TAG. Así que antes de presentar los algoritmos de análisis para TAG, definiremos las características generales de los distintos tipos de ítems con los que trabajarán cada esquema, así como otras convenciones y notación que se emplearán a lo largo del texto.

Comenzaremos describiendo un analizador basado en el conocido algoritmo CYK para CFG, que será el punto de partida para construir la mencionada red de esquemas para TAG. El interés práctico de este analizador es limitado, debido a las fuertes restricciones que exigen a la gramática. El siguiente esquema que mostramos también presenta una estrategia ascendente, pero elimina las restricciones sobre la gramática impuestas por el algoritmo CYK. El tercer esquema ascendente que se define se caracteriza por el recorrido bidireccional de la cadena de entrada y está basado en el algoritmo para gramáticas independientes del contexto definido por de Vreught y Honig. Aunque este esquema no parezca mantener relación con el resto de los analizadores mostrados, en [35] se demuestra que dicha relación sí existe. Por último, se definirá un esquema para TAG con estrategia ascendente e información predictiva inspirado en el conocido algoritmo de Earley para CFG, tanto la versión que verifica la *propiedad de prefijo válido* como la que no lo hace. En la figura 4.1 se muestran las relaciones formales entre ellos<sup>1</sup>.

---

<sup>1</sup>En la figura se enmarcan los analizadores de interés, el resto son esquemas intermedios definidos para establecer las relaciones formales. En [6] se demuestran todas estas relaciones, salvo la existente entre los esquemas **buE** y **dvH**, que se puede encontrar en [35].



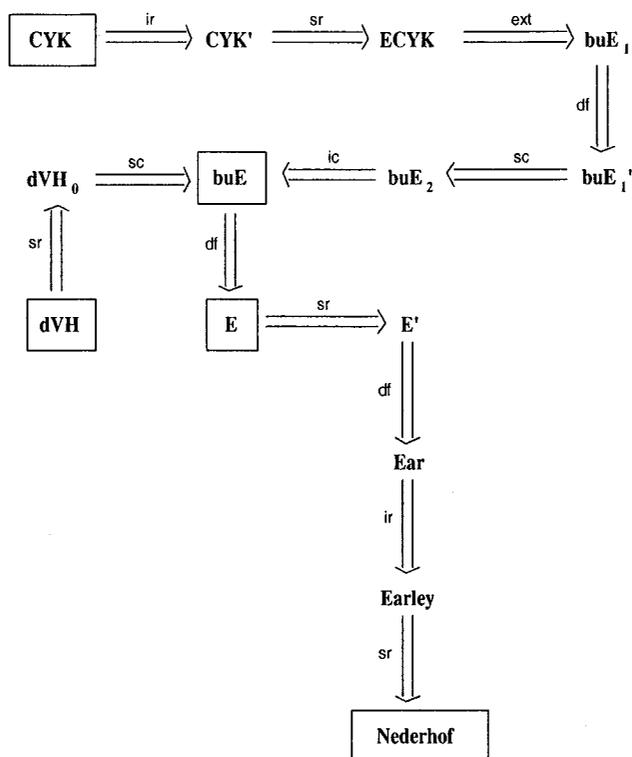


Figura 4.1: Red de analizadores para TAG

## 4.1 Esquemas de análisis sintáctico y TAG

En primer lugar introduciremos la notación y convenciones específicas que se emplean en la definiciones de los esquemas de análisis para TAG.

### Notación 4

- Usamos:
  - $\alpha, \beta, \gamma$  para representar árboles elementales.
  - $N^\gamma, M^\gamma, \dots$  para identificar los nodos de un árbol elemental  $\gamma$ .
  - $\delta, \nu, \omega, \dots$  para representar secuencias de nodos.
  - $\mathbf{R}^\gamma$  para denotar el nodo raíz de un árbol elemental  $\gamma$ .
  - $\mathbf{F}^\beta$  para denotar el nodo pie de un árbol auxiliar  $\beta$ .
- $\text{label}(N^\gamma)$  es la etiqueta asociada al nodo  $N^\gamma$ .
- $\text{adj}(M^\gamma)$  es el conjunto de árboles auxiliares que se pueden adjuntar sobre el nodo  $M^\gamma$ . Entonces, si un árbol auxiliar  $\beta$  se puede adjuntar sobre el nodo  $M^\gamma$  lo denotamos como  $\beta \in \text{adj}(M^\gamma)$ .
- Si un nodo  $M^\gamma$  no presenta una restricción de adjunción obligatoria, lo denotamos como  $\text{nil} \in \text{adj}(M^\gamma)$ , donde  $\text{nil}$  es un símbolo vacío.

- Dados dos pares de variables enteras,  $(p, q)$  y  $(i, j)$ , decimos que  $(p, q) \leq (i, j)$  si se cumple que  $i \leq p \leq q \leq j$ .
- Dadas dos variables enteras,  $p$  y  $q$ , definimos  $p \cup q$  como  $p$  si  $q$  está indefinido y como  $q$  si  $p$  está indefinido, estando indefinido en cualquier otro caso. Usaremos el símbolo  $-$  para indicar que una variable entera permanece indefinida.
- Tal como se especifica en [84], incluimos dos nuevos nodos: el nodo top, etiquetado como  $\top$ , que dominará el nodo raíz de todo árbol elemental, y el nodo bottom, etiquetado como  $\perp$ , que será el único descendiente del nodo pie de todo árbol auxiliar. Con objeto de no modificar la potencia expresiva de la gramática, ambos nodos presentan restricción de adjunción nula.

Nuestro objetivo en esta sección es definir las pautas generales para adaptar los esquemas de análisis sintáctico basados en ítems, presentados por Sikkel para CFG, al formalismo TAG. Las dos características distintivas de las TAG, los árboles como estructuras elementales y la adjunción como operación de composición básica, son las que establecerán la estructura básica de los ítems para estos analizadores.

El primer problema que se plantea es cómo representar los árboles en forma de reglas al estilo de las empleadas en los esquemas de análisis sintáctico para CFG mostrados. Existen dos alternativas para solventar este problema [38]:

- Asociar una sola regla a cada árbol elemental de una gramática TAG, usando para ello el mecanismo de *aplanamiento* que se detalla en el capítulo 7. A esta forma de representar los árboles elementales la denominamos *notación plana*. La ventaja fundamental de este método de representación es que se evita la navegación a través de los distintos niveles de un árbol y, por tanto, la definición de los analizadores se puede simplificar. Sin embargo, los inconvenientes que presenta, como la introducción de nuevos símbolos en la gramática y la notación engorrosa, la hacen poco funcional.
- Asociar un conjunto de reglas a cada árbol elemental de una gramática TAG. A esta forma de representar los árboles elementales la denominamos *representación multicapa*, y consiste en asociar a cada nodo interior  $M^\gamma$  de un árbol elemental  $\gamma$  una regla de la forma

$$M^\gamma \rightarrow M_1^\gamma \dots M_n^\gamma,$$

donde  $M_1^\gamma \dots M_n^\gamma$  es la secuencia ordenada de los descendientes directos de  $M^\gamma$ . El conjunto de todas las reglas asociadas a un árbol elemental  $\gamma$  se denota como  $\mathcal{P}(\gamma)$ , y el de todas las reglas asociadas a una gramática  $G$  como  $\mathcal{P}(G)$ . En la figura 4.2 se muestra la gramática TAG de la figura 3.3 representada con notación multicapa y con sus restricciones de adjunción.

El conjunto de reglas de producción que se obtienen con la representación multicapa de una gramática TAG es similar a las de una CFG, sin embargo, la definición de derivación en CFG no se puede aplicar sin más sobre este conjunto de producciones. Ésto es debido a la forma de composición de árboles que introduce la operación de adjunción, la cual hace que se

**Representación multicapa:**

$$\begin{aligned}\mathcal{P}(\alpha) &= \{S^0 \rightarrow e\} \\ \mathcal{P}(\beta_1) &= \{S^1 \rightarrow aS^2, S^2 \rightarrow S^3a\} \\ \mathcal{P}(\beta_2) &= \{S^4 \rightarrow bS^5, S^5 \rightarrow S^6b\}\end{aligned}$$

**Restricciones de adjunción:**

$$\begin{aligned}adj(S^0) &= \{\mathbf{nil}, \beta_1, \beta_2\} \\ adj(S^1) &= \{\mathbf{nil}\} & adj(S^2) &= \{\mathbf{nil}, \beta_1, \beta_2\} \\ adj(S^3) &= \{\mathbf{nil}\} & adj(S^5) &= \{\mathbf{nil}, \beta_1, \beta_2\} \\ adj(S^4) &= \{\mathbf{nil}\} & adj(S^6) &= \{\mathbf{nil}\}\end{aligned}$$

Figura 4.2: Representación multicapa de una gramática TAG

produzcan saltos entre árboles durante el proceso de derivación. Se plantea la necesidad de revisar el concepto de derivación en las gramáticas CFG para que recoja esta circunstancia y sea aplicable al conjunto de producciones de una gramática TAG.

**Definición 4.1** Relación de derivación  $\overset{*}{\Rightarrow}$  en  $\mathcal{P}(G)$

La relación  $\overset{*}{\Rightarrow}$  en  $\mathcal{P}(G)$  sobre  $V^* \times V^*$  se define como la clausura reflexiva y transitiva con los siguientes dos casos base, en función de si se ha efectuado o no una adjunción sobre un nodo:

- $\delta' M \gamma \delta'' \overset{*}{\Rightarrow} \delta' \nu \delta''$  si  $M \gamma \rightarrow \nu, \mathbf{nil} \in \text{adj}(M \gamma)$ .
- $\delta' M \gamma \delta'' \overset{*}{\Rightarrow} \delta' \omega' \nu \omega'' \delta''$  si  $M \gamma \rightarrow \nu, \beta \in \text{adj}(M \gamma)$ ,  
 $A^\beta \overset{*}{\Rightarrow} \omega' P^\beta \omega'', \text{label}(A^\beta) = \top, \text{label}(P^\beta) = \perp$ .

El primer caso define el recorrido en el interior de un árbol elemental, mientras el segundo caso define la posibilidad de introducir en la derivación un árbol auxiliar en el que la espina haya sido completamente reconocida. No obstante, la definición anterior es insuficiente para expresar correctamente las invariantes de ciertos ítems<sup>2</sup>, por lo que se requieren formas adicionales de derivación para TAG.

**Definición 4.2** Relación de derivación  $\overset{*}{\Rightarrow}_r$  en  $\mathcal{P}(G)$

La relación  $\overset{*}{\Rightarrow}_r$  en  $\mathcal{P}(G)$  sobre  $V^* \times V^*$  se define como  $\Rightarrow_r \cup \overset{*}{\Rightarrow}$  con el siguiente caso base:  $\delta' M \gamma \delta'' \Rightarrow_r \delta' P^\beta \delta''$  si  $\beta \in \text{adj}(M \gamma), \text{label}(P^\beta) = \top$ .

**Definición 4.3** Relación de derivación  $\overset{*}{\Rightarrow}_f$  en  $\mathcal{P}(G)$

La relación  $\overset{*}{\Rightarrow}_f$  en  $\mathcal{P}(G)$  sobre  $V^* \times V^*$  se define como  $\Rightarrow_f \cup \overset{*}{\Rightarrow}$  con el siguiente caso base:  $\delta' P^\beta \delta'' \Rightarrow_f \delta' \nu \delta''$  si  $M \gamma \rightarrow \nu \in \mathcal{P}(\gamma), \beta \in \text{adj}(M \gamma), \text{label}(P^\beta) = \perp$ .

<sup>2</sup>En comunicación privada de G. Satta a M.A. Alonso se muestran las limitaciones que presentan la definición de derivación para TAG, no discutida hasta ese momento en ningún foro, para expresar correctamente invariantes de ciertos tipos de ítems cuando aparecen adjunciones anidadas en la espina.

La relación  $\overset{*}{\Rightarrow}_r$  se usa para expresar el paso de un nodo de adjunción al nodo raíz de un árbol auxiliar adjuntable sobre él, y la relación  $\overset{*}{\Rightarrow}_f$  para expresar el paso de un nodo pie a un posible nodo de adjunción.

También se introducen otras dos formas auxiliares para distinguir derivaciones con o sin adjunción sobre un nodo, y que son especialmente útiles para expresar invariantes en analizadores de tipo CYK. Así  $M^\gamma \overset{*}{\Rightarrow}_t \delta$  si  $M^\gamma \overset{*}{\Rightarrow} \delta$  y  $M^\gamma \overset{*}{\Rightarrow}_b \delta$  si  $\nu \overset{*}{\Rightarrow} \delta$ , con  $M^\gamma \rightarrow \nu$ .

Ya estamos en disposición de establecer la forma que tendrán los ítems en los analizadores para TAG que vamos a presentar en la memoria. Generalizando, podemos decir que la forma de los ítems en los esquemas para TAG son extensiones a los ítems de los esquemas homónimos para CFG y que, al igual que en éstos, se estructuran alrededor de dos secciones básicas: la que representa el reconocimiento total o parcial de un árbol dominado por un nodo de la gramática y la tupla de números naturales que hacen referencia a segmentos reconocidos de la cadena de entrada.

La primera sección va a depender del esquema que se defina, encontrándonos las siguientes posibilidades:

- Un solo nodo, que expresa que el árbol dominado por él ha sido completamente reconocido. Esta forma sólo se emplea en el esquema basado en CYK.
- Un regla de  $\mathcal{P}(G)$  con un punto que separa el fragmento reconocido del árbol que representa dicha regla, situado a la izquierda del punto, del que queda aún por reconocer. Esta forma se emplea en todos los analizadores con recorrido unidireccional de la cadena de entrada.
- Un regla de  $\mathcal{P}(G)$  con dos puntos que delimitan el fragmento reconocido del árbol que representa dicha regla, situado entre ambos puntos. Esta forma se emplea en el esquema ascendente bidireccional basado en el algoritmo de Vreught y Honig.

La tupla de números naturales que hacen referencia a posiciones en la cadena de entrada la vamos a dividir en dos pares. El primer par delimita el fragmento en la cadena de entrada que domina la parte reconocida y especificada en la primera sección del ítem. El segundo par delimita el fragmento en la cadena de entrada que domina el nodo pie de un árbol auxiliar, si éste se encuentra en la parte reconocida y especificada en la primera sección del ítem. Este segundo par permite transmitir desde el pie hasta la raíz de un árbol auxiliar la información del segmento de la cadena de entrada reconocido por el nodo pie, información necesaria para completar la adjunción del árbol auxiliar sobre un nodo. La forma en la que se transmite dicha información depende del esquema [35], mientras en los esquemas con recorrido unidireccional de la cadena de entrada la transmisión se realiza a través de los nodos situados en la espina y los hermanos derechos, en el esquema basado en CYK se realiza a través de los nodos de la espina.

Aparte de las dos secciones básicas, en algunos esquemas se añaden componentes que introducen restricciones sobre los componentes básicos. En los analizadores que verifican

la *propiedad del prefijo válido* es necesario incluir un índice adicional, además de los dos pares de índices descritos, que indique la posición en la cadena de entrada donde comenzó el reconocimiento del árbol elemental al que pertenece la regla o nodo de la primera sección del ítem.

## 4.2 Esquema tipo CYK

El primer esquema que veremos, denominado **CYK**, presenta una estrategia de análisis ascendente y se trata de una extensión del algoritmo CYK definido originalmente para gramáticas independientes del contexto. **CYK** fue introducido en [7, 8] basándose, con algunas modificaciones, en el algoritmo presentado en [136, 132]. Posteriormente en [35] se modifica el mismo mediante la introducción de la operación de unión sobre índices descrita en la sección anterior.

El esquema **CYK** sólo es aplicable a gramáticas de adjunción de árboles cuyos árboles elementales presentan las siguientes restricciones: (i) un nodo interno, salvo el nodo pie, dominará directamente un máximo de dos nodos y (ii) los nodos etiquetados con símbolos terminales o la palabra vacía no tendrán nodos hermanos. Estas restricciones son una trasposición de  $\mathcal{CNF}$  para gramáticas de adjunción de árboles.

El conjunto de ítems del esquema  $\mathcal{I}_{\text{CYK}}$  (ver figura 4.3) se define mediante :

$$\mathcal{I}_{\text{CYK}} = \{[N^\gamma, i, j \mid p, q \mid \text{adj}]\}$$

donde  $\text{label}(N^\gamma) \in V_N$  y  $\gamma \in I \cup A$ . Los índices  $0 \leq i \leq j$  establecen las posiciones dentro de la cadena de entrada que delimitan el fragmento reconocido por  $N^\gamma$ . Si  $p$  y  $q$  presentan un valor conocido, entonces  $N^\gamma$  pertenece a la espina de un árbol auxiliar y se cumple  $(p, q) \leq (i, j)$ . El componente  $\text{adj} \in \{\text{true}, \text{false}\}$  se añade a la definición propuesta en [132, 136], y permite al analizador determinar si se ha llevado a cabo una adjunción sobre el nodo  $N^\gamma$  e impedir que se realicen más de una adjunción sobre dicho nodo.

Las reglas deductivas son las siguientes:

$$\mathcal{D}_{\text{CYK}} = \mathcal{D}_{\text{CYK}}^{\text{Scan}} \cup \mathcal{D}_{\text{CYK}}^\varepsilon \cup \mathcal{D}_{\text{CYK}}^{\text{Foot}} \cup \mathcal{D}_{\text{CYK}}^{\text{CompUna}} \cup \mathcal{D}_{\text{CYK}}^{\text{CompBin}} \cup \mathcal{D}_{\text{CYK}}^{\text{Adj}}$$

El reconocimiento ascendente se inicia con los pasos deductivos  $\mathcal{D}_{\text{CYK}}^{\text{Scan}}$  y  $\mathcal{D}_{\text{CYK}}^\varepsilon$ .  $\mathcal{D}_{\text{CYK}}^{\text{Scan}}$  introduce en el análisis los subárboles que dominan directamente nodos etiquetados con símbolos terminales que coincida con algún símbolo de la cadena de entrada. Y  $\mathcal{D}_{\text{CYK}}^\varepsilon$  introduce aquellos subárboles que dominan directamente nodos etiquetados con la palabra vacía. La estrategia ascendente obliga a que  $\mathcal{D}_{\text{CYK}}^\varepsilon$  sitúe el reconocimiento sobre cualquier posición de la cadena de entrada.

$$\mathcal{D}_{\text{CYK}}^{\text{Scan}} = \frac{[a, j, j+1]}{[N^\gamma, j, j+1 \mid -, - \mid \text{false}]} \quad N^\gamma \rightarrow a \in \mathcal{P}(\gamma)$$

$$\mathcal{D}_{\text{CYK}}^\varepsilon = \frac{}{[N^\gamma, j, j \mid -, - \mid \text{false}]} \quad N^\gamma \rightarrow \varepsilon \in \mathcal{P}(\gamma)$$

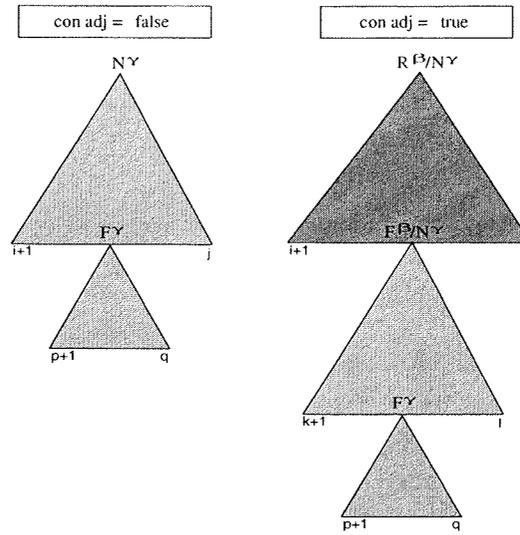


Figura 4.3: Ítems del esquema CYK

De nuevo la estrategia ascendente del esquema, al desconocer el subárbol escindido por una adjunción, obliga al analizador a predecir todos los posibles fragmentos de la cadena de entrada que puede cubrir el nodo pie. Lo que hace mediante el paso deductivo  $\mathcal{D}_{\text{CYK}}^{\text{Foot}}$ . Este paso es el que asigna los valores de los índices que mantienen información del fragmento de la cadena de entrada que domina el nodo pie, y que posteriormente será transmitida a través de los nodos de la espina.

$$\mathcal{D}_{\text{CYK}}^{\text{Foot}} = \frac{}{[P^\beta, k, l \mid k, l \mid false]} \text{label}(P^\beta) = \perp$$

Una vez que el reconocimiento de un subárbol dominado por un nodo  $M^\gamma$  se ha completado, se debe continuar con el reconocimiento del superárbol de dicho nodo. Así, mientras el paso deductivo  $\mathcal{D}_{\text{CYK}}^{\text{CompUna}}$  se aplica a nodos que dominan directamente a un solo nodo, el paso  $\mathcal{D}_{\text{CYK}}^{\text{CompBin}}$  se aplica a los que dominan directamente a dos nodos.

$$\mathcal{D}_{\text{CYK}}^{\text{CompUna}} = \frac{[M^\gamma, i, j \mid p, q \mid adj]}{[N^\gamma, i, j \mid p, q \mid false]} \quad \begin{array}{l} N^\gamma \rightarrow M^\gamma \in \mathcal{P}(\gamma) \\ adj = false \text{ sii } \mathbf{nil} \in \text{adj}(M^\gamma) \end{array}$$

$$\mathcal{D}_{\text{CYK}}^{\text{CompBin}} = \frac{[M^\gamma, i, k \mid p, q \mid adj][O^\gamma, k, j \mid p', q' \mid adj']}{[N^\gamma, i, j \mid p \cup p' \mid q \cup q' \mid false]} \quad \begin{array}{l} N^\gamma \rightarrow M^\gamma O^\gamma \in \mathcal{P}(\gamma) \\ adj = false \text{ sii } \mathbf{nil} \in \text{adj}(M^\gamma) \\ adj' = false \text{ sii } \mathbf{nil} \in \text{adj}(O^\gamma) \end{array}$$

Hay que indicar que  $adj$  (resp.  $adj'$ ) sólo puede tomar el valor  $false$  cuando el nodo  $M^\gamma$  (resp.  $O^\gamma$ ) no presente restricción de adjunción obligatoria. Obsérvese también el uso del operador unión sobre los índices para comprimir la definición de tres pasos deductivos en éste.



El paso deductivo  $\mathcal{D}_{\text{CYK}}^{\text{Adj}}$  continúa el reconocimiento del superárbol respecto del nodo  $M^\gamma$  sobre el que se ha efectuado una adjunción del árbol auxiliar  $\beta$ , una vez que se ha completado el reconocimiento de éste.

$$\mathcal{D}_{\text{CYK}}^{\text{Adj}} = \frac{[P^\beta, j, m \mid k, l \mid \text{false}] [M^\gamma, k, l \mid p, q \mid \text{false}]}{[M^\gamma, j, m \mid p, q \mid \text{true}]} \quad \begin{array}{l} \text{label}(P^\beta) = \top \\ \beta \in \text{adj}(M^\gamma) \end{array}$$

Se establece el valor *true* al componente *adj* del consecuente para impedir que se lleve a cabo más de una adjunción sobre el nodo  $M^\gamma$ .

El conjunto de ítems finales se define mediante:

$$\mathcal{F}_{\text{CYK}} = \{[P^\alpha, 0, n \mid -, - \mid \text{false}] \mid \alpha \in \mathbf{I}, \text{label}(P^\alpha) = \top\}$$

### 4.3 Esquema tipo Earley ascendente

El esquema **buE** que describimos ahora, presentado en [7, 8], se obtiene a partir de una generalización del esquema **CYK** mostrado en la sección anterior. El interés de este esquema radica en que se trata de un reconocedor con estrategia ascendente que elimina la restricción impuesta por el esquema **CYK** sobre la forma que deben tener los árboles elementales. Como se observa en los resultados experimentales mostrados en [35, 24], el comportamiento de **buE** es, en general, peor que otros analizadores que usan estrategias ascendentes o con algún tipo de filtro.

El dominio del esquema  $\mathcal{I}_{\text{buE}}$  (ver figura 4.4) se define mediante:

$$\mathcal{I}_{\text{buE}} = \{[N^\gamma \rightarrow \nu \bullet \omega, i, j \mid p, q]\}$$

donde  $N^\gamma \rightarrow \nu \omega \in \mathcal{P}(\gamma)$  y  $\gamma \in \mathbf{I} \cup \mathbf{A}$ . Los índices  $0 \leq i \leq j$  establecen las posiciones de la cadena de entrada que delimitan el fragmento reconocido por  $\nu$ . Si  $p$  y  $q$  presentan un valor definido, entonces  $N^\gamma$  pertenece a la espina de un árbol auxiliar,  $\nu$  domina el nodo pie de  $\gamma$  y se cumple  $(p, q) \leq (i, j)$ .

Los pasos deductivos del esquema vienen dados por:

$$\mathcal{D}_{\text{buE}} = \mathcal{D}_{\text{buE}}^{\text{Ini}} \cup \mathcal{D}_{\text{buE}}^{\text{Foot}} \cup \mathcal{D}_{\text{buE}}^{\text{Scan}} \cup \mathcal{D}_{\text{buE}}^{\text{e}} \cup \mathcal{D}_{\text{buE}}^{\text{Comp}} \cup \mathcal{D}_{\text{buE}}^{\text{AdjComp}}$$

El paso deductivo  $\mathcal{D}_{\text{buE}}^{\text{Ini}}$  establece de partida la predicción de todos los subárboles participantes en los árboles elementales. Debido a la estrategia ascendente pura del esquema, esta regla comienza el reconocimiento de los subárboles sobre cualquier posición de la cadena de entrada. Esto va a generar una gran cantidad de ítems espurios que con una adecuada técnica de filtrado podemos suprimir parcialmente, como veremos cuando definamos el esquema **buLC**.

$$\mathcal{D}_{\text{buE}}^{\text{Ini}} = \overline{[N^\gamma \rightarrow \bullet \delta, i, i \mid -, -]}$$

El paso deductivo  $\mathcal{D}_{\text{buE}}^{\text{Foot}}$  completa el reconocimiento de los subárboles que dominan los nodos pie de los árboles auxiliares. Al igual que el paso anterior, el análisis ascendente obliga

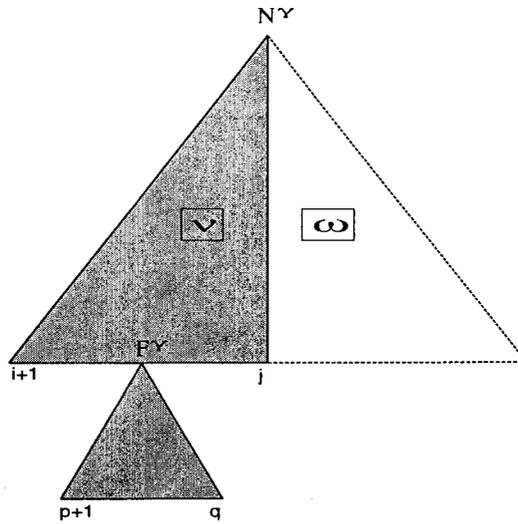


Figura 4.4: Ítems del esquema buE

a suponer que el nodo pie dominará cualquier subcadena válida de la cadena de entrada. Es en este paso donde se asignan valores para los dos últimos índices, ya que éstos establecen el fragmento de la cadena de entrada que domina el subárbol que pende del nodo pie del árbol auxiliar.

$$D_{\text{buE}}^{\text{Foot}} = \frac{}{[\mathbf{F}^\beta \rightarrow \perp \bullet, k, l \mid k, l]}$$

Los pasos de reconocimiento  $D_{\text{buE}}^{\text{Scan}}$  y  $D_{\text{buE}}^\varepsilon$  se aplican cuando el reconocimiento alcanza un nodo etiquetado con un símbolo terminal o cuando la cadena vacía que se corresponde con el símbolo actual de la cadena de entrada.

$$D_{\text{buE}}^{\text{Scan}} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q], [a, j, j+1]}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, j+1 \mid p, q]} \text{ label}(M^\gamma) = a$$

$$D_{\text{buE}}^\varepsilon = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, j \mid p, q]} \text{ label}(M^\gamma) = \varepsilon$$

El paso deductivo  $D_{\text{buE}}^{\text{Comp}}$  continúa el reconocimiento del superárbol respecto a  $M^\gamma$ , una vez que el subárbol dominado por él ha sido completamente reconocido. Este paso es el equivalente a una operación de completación en gramáticas incontextuales, por tanto, solo se puede aplicar cuando la adjunción no sea obligatoria en  $M^\gamma$ .

$$D_{\text{buE}}^{\text{Comp}} = \frac{[M^\gamma \rightarrow \delta \bullet, j, k \mid p, q], [N^\gamma \rightarrow \nu \bullet M^\gamma \omega, i, j \mid p', q']}{[N^\gamma \rightarrow \nu M^\gamma \bullet \omega, i, k \mid p \cup p', q \cup q']} \text{ nil} \in \text{adj}(M^\gamma)$$



El paso deductivo  $\mathcal{D}_{\text{buE}}^{\text{AdjComp}}$  continúa el reconocimiento del superárbol respecto a  $M^\gamma$  donde se ha efectuado la adjunción del árbol auxiliar  $\beta$ , una vez que éste ha sido completamente reconocido. Para evitar la posibilidad de múltiples operaciones de adjunción sobre un nodo, el punto se sitúa detrás del mismo en el consecuente.

$$\mathcal{D}_{\text{buE}}^{\text{AdjComp}} = \frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\beta \bullet, j, m \mid k, l] \\ [M^\gamma \rightarrow \delta \bullet, k, l \mid p, q] \\ [N^\gamma \rightarrow \nu \bullet M^\gamma \omega, i, j \mid p', q'] \end{array}}{[N^\gamma \rightarrow \nu M^\gamma \bullet \omega, i, m \mid p \cup p', q \cup q']} \quad \beta \in \text{adj}(M^\gamma)$$

El conjunto de ítems finales viene dado por:

$$\mathcal{F}_{\text{buE}} = \{[\top \rightarrow \mathbf{R}^\alpha \bullet, 0, n \mid -, -] \mid \alpha \in \mathbf{I}\}$$

## 4.4 Esquema tipo de Vreught y Honig

Definimos ahora el esquema **dVH** presentado en [41, 44, 42]. Se trata de un esquema con estrategia ascendente y recorrido bidireccional de la cadena de entrada que está basado en el algoritmo para gramáticas independientes del contexto definido por De Vreught y Honig. Al tratarse de un analizador ascendente sin ningún tipo de predicción, puede producir un mayor número de ítems durante el proceso de análisis, sin embargo, los estudios preliminares mostrados en [45] demuestran que su comportamiento en casos prácticos es similar al de analizadores que incorporan predicción. Las ventajas fundamentales que ofrece este esquema son dos: sus pasos deductivos requieren un máximo de dos antecedentes y permite obtener mayor información parcial para entradas incorrectas. Así, mientras lo primero disminuye el espacio de búsqueda del algoritmo y aumenta su velocidad, lo segundo hace de este algoritmo una opción adecuada para realizar análisis sintáctico fragmental.

Puesto que se trata de un algoritmo bidireccional, el reconocimiento puede comenzar en cualquier posición de las reglas y expandirse en ambas direcciones. Esto nos obliga a modificar la forma general de regla punteada usada en los esquemas de tipo Earley, de manera que introducimos un punto adicional para delimitar la parte reconocida dentro de una regla. Entonces el dominio del esquema **dVH** se define mediante ítems de la forma (ver figura 4.5):

$$\mathcal{I}_{\text{dVH}} = \{[N^\gamma \rightarrow \nu \bullet \delta \bullet \omega, i, j \mid p, q]\}$$

donde  $N^\gamma \rightarrow \nu \delta \omega \in \mathcal{P}(\gamma)$  y  $\gamma \in \mathbf{I} \cup \mathbf{A}$ . Los índices  $0 \leq i \leq j$  establecen las posiciones de la cadena de entrada que delimitan el fragmento reconocido por  $\delta$ . Si  $p$  y  $q$  presentan un valor definido, entonces  $N^\gamma$  pertenece a la espina de un árbol auxiliar,  $\delta$  domina el nodo pie de  $\gamma$  y se cumple  $(p, q) \leq (i, j)$ .

El conjunto de pasos deductivos viene dado por:

$$\mathcal{D}_{\text{dVH}} = \mathcal{D}_{\text{dVH}}^{\text{Scan}} \cup \mathcal{D}_{\text{dVH}}^{\text{e}} \cup \mathcal{D}_{\text{dVH}}^{\text{Comp}} \cup \mathcal{D}_{\text{dVH}}^{\text{Con}} \cup \mathcal{D}_{\text{dVH}}^{\text{Foot}} \cup \mathcal{D}_{\text{dVH}}^{\text{AdjComp}}$$

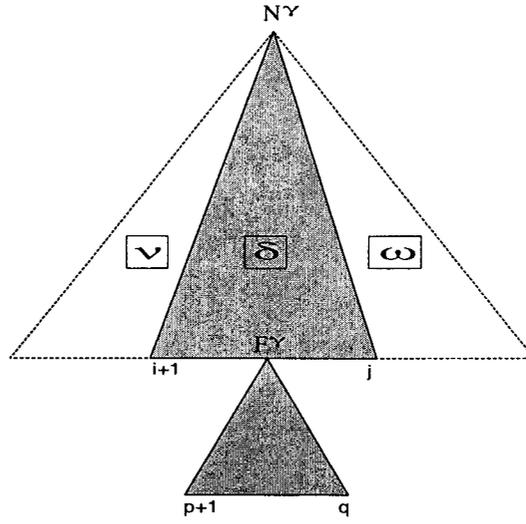


Figura 4.5: Ítems del esquema dVH

Al igual que el esquema **CYK**, el reconocimiento ascendente se inicia con los pasos deductivos  $\mathcal{D}_{dVH}^{Scan}$  y  $\mathcal{D}_{dVH}^{\varepsilon}$ , los cuales introducen en el análisis los subárboles que dominan directamente nodos etiquetados con símbolos terminales que coincida con algún símbolo de la cadena de entrada ( $\mathcal{D}_{dVH}^{Scan}$ ) o nodos etiquetados con la palabra vacía ( $\mathcal{D}_{dVH}^{\varepsilon}$ ).

$$\mathcal{D}_{dVH}^{Scan} = \frac{[a, j, j+1]}{[N^{\gamma} \rightarrow \nu \bullet M^{\gamma} \bullet \omega, j, j+1 \mid -, -]} \quad label(M^{\gamma}) = a$$

$$\mathcal{D}_{dVH}^{\varepsilon} = \frac{[ ]}{[N^{\gamma} \rightarrow \nu \bullet M^{\gamma} \bullet \omega, j, j \mid -, -]} \quad label(M^{\gamma}) = \varepsilon$$

La función del paso  $\mathcal{D}_{dVH}^{Comp}$  es igual a la del paso de completación del esquema **buE**, es decir, continúa el reconocimiento del superárbol respecto a un nodo una vez que el subárbol dominado por él ha sido completamente reconocido, siempre que la adjunción no sea obligatoria en dicho nodo.

$$\mathcal{D}_{dVH}^{Comp} = \frac{[M^{\gamma} \rightarrow \bullet \delta \bullet, i, j \mid p, q]}{[N^{\gamma} \rightarrow \nu \bullet M^{\gamma} \bullet \omega, i, j \mid p, q]} \quad nil \in adj(M^{\gamma})$$

Cuando dos fragmentos adyacentes en una regla hayan reconocido segmentos colindantes en la cadena de entrada, el paso  $\mathcal{D}_{dVH}^{Con}$  se encarga de concatenarlos.

$$\mathcal{D}_{dVH}^{Con} = \frac{\begin{array}{l} [N^{\gamma} \rightarrow \nu \bullet \delta \bullet \delta' \omega, i, j' \mid p, q] \\ [N^{\gamma} \rightarrow \nu \delta \bullet \delta' \bullet \omega, j', j \mid p', q'] \end{array}}{[N^{\gamma} \rightarrow \nu \bullet \delta \delta' \bullet \omega, i, j \mid p \cup p', q \cup q']}$$

El paso deductivo  $\mathcal{D}_{dVH}^{Foot}$  lleva a cabo el reconocimiento del nodo pie de los árboles auxiliares entre todas las posibles posiciones de la cadena de entrada. Evidentemente, este paso

se puede filtrar incluyendo un antecedente que refleje el hecho de que el subárbol que pende del nodo pie ya haya sido reconocido. Sin embargo, no consideraremos esta mejora, incluida en la definición original de [41], para establecer con más nitidez la relación existente entre este esquema y el **buE**.

$$\mathcal{D}_{\text{dVH}}^{\text{Foot}} = \frac{}{[\mathbf{F}^\beta \rightarrow \bullet \perp \bullet, k, l \mid k, l]} \quad \beta \in \text{adj}(M^\gamma)$$

El paso deductivo  $\mathcal{D}_{\text{dVH}}^{\text{AdjComp}}$  continúa el reconocimiento del superárbol respecto a  $M^\gamma$  donde se ha efectuado la adjunción del árbol auxiliar  $\beta$ , una vez que éste ha sido completamente reconocido.

$$\mathcal{D}_{\text{dVH}}^{\text{AdjComp}} = \frac{[\top \rightarrow \bullet \mathbf{R}^\beta \bullet, j, m \mid k, l] \quad [M^\gamma \rightarrow \bullet \delta \bullet, k, l \mid p, q]}{[N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, j, m \mid p, q]} \quad \beta \in \text{adj}(M^\gamma)$$

Como se puede observar, todos los pasos de este esquema presentan la misma funcionalidad que sus homólogos en el esquema **buE**. Las diferencias fundamentales entre ambos vienen dadas por la característica bidireccional del esquema **dVH**, la cual hace que el fragmento reconocido de un subárbol se pueda situar en cualquier posición de una regla de producción. En el esquema **buE**, sin embargo, el reconocimiento de la cadena de entrada es de izquierda a derecha, lo que hace que el comienzo del fragmento reconocido en una regla de producción siempre sea el comienzo de la propia regla. Esta diferencia genera que: (1) el paso de inicio existente en **buE** ya no sea necesario, ya que el reconocimiento de una regla no se tiene que realizar de izquierda a derecha; (2) se requiera un paso de concatenación para unir dos fragmentos reconocidos adyacentes en la misma regla, paso no necesario en **buE** debido a que las compleciones de subárbol y adjunción siempre se realizan sobre nodos cuyos contextos izquierdos ya han sido reconocidos; (3) se requiera una reformulación de todos los pasos para su adaptación a la forma de ítems con dos puntos.

El conjunto de ítems finales se define mediante:

$$\mathcal{F}_{\text{dVH}} = \{[\top \rightarrow \bullet \mathbf{R}^\alpha \bullet, 0, n \mid -, -] \mid \alpha \in I\}$$

Los analizadores ascendentes, en general, y el caso que nos ocupa, en particular, presentan un serio problema: los pasos de inicio del reconocimiento ascendente sobregeneran una gran cantidad de ítems. Con objeto de disminuir esta sobregeneración, en [35] se propone un esquema alternativo, al que denomina **dVH(1, 1)**, y que se obtiene mediante la aplicación de un filtro dinámico sobre **dVH**. Este esquema introduce en los ítems información de los contextos derecho e izquierdo de los nodos no terminales con objeto de incluir en el análisis solamente los ítems que sean consistentes con la cadena de entrada.

## 4.5 Esquemas tipo Earley

Vamos a ver a continuación el esquema **E** para TAG, que es esencialmente igual al definido en [71, 8]. Este analizador usa estrategia ascendente con información predictiva, al estilo del

algoritmo de Earley para CFG, pero no cumple la propiedad del prefijo válido.

Un analizador que verifica la *propiedad del prefijo válido* garantiza la detección temprana de errores. Esto no quiere decir que los analizadores que no cumplan esta propiedad no detecten los errores, sino que lo hacen más tarde. Formalmente, un analizador sintáctico satisface la propiedad del prefijo válido (VPP) si al leer la subcadena prefija  $a_1 \dots a_k$  de la cadena de entrada  $a_1 \dots a_k a_{k+1} \dots a_n$ , garantiza que existe una cadena  $b_i \dots b_m$ , donde  $b_i$  no tiene que pertenecer a la cadena de entrada de entrada, tal que  $a_1 \dots a_k b_i \dots b_m$  es una cadena válida del lenguaje.

El conjunto de ítems del esquema  $\mathcal{I}_E$  es igual al definido para el esquema **buE**:

$$\mathcal{I}_E = \mathcal{I}_{\text{buE}}$$

El conjunto de pasos deductivos del esquema  $\mathcal{D}_E$  se define mediante:

$$\begin{aligned} \mathcal{D}_E = & \mathcal{D}_E^{\text{Ini}} \cup \mathcal{D}_E^{\text{Scan}} \cup \mathcal{D}_E^\epsilon \cup \mathcal{D}_E^{\text{Pred}} \cup \mathcal{D}_E^{\text{Comp}} \cup \\ & \mathcal{D}_E^{\text{AdjPred}} \cup \mathcal{D}_E^{\text{FootPred}} \cup \mathcal{D}_E^{\text{FootComp}} \cup \mathcal{D}_E^{\text{AdjComp}} \end{aligned}$$

### Inicio

El reconocimiento comienza con la predicción de todos los árboles iniciales.

$$\mathcal{D}_E^{\text{Ini}} = \frac{[\top \rightarrow \bullet \mathbf{R}^\alpha, 0, 0 \mid -, -]}{[\top \rightarrow \bullet \mathbf{R}^\alpha, 0, 0 \mid -, -]} \quad \alpha \in I$$

### Reconocimiento

Los pasos deductivos de reconocimiento son iguales a los del esquema **buE**.

$$\mathcal{D}_E^{\text{Scan}} = \mathcal{D}_{\text{buE}}^{\text{Scan}}$$

$$\mathcal{D}_E^\epsilon = \mathcal{D}_{\text{buE}}^\epsilon$$

Los pasos deductivos que establecen la estrategia ascendente predictiva del analizador Earley para CFG son los correspondientes a predicciones y compleciones. Para el caso de las TAG, vamos a distinguir tres tipos de predicciones con sus correspondientes pasos de compleción asociados: subárbol, adjunción y pie.

### Predicción de subárbol

Este paso deductivo es similar al paso predictivo del analizador Earley para CFG. De manera que si se alcanza un nodo que no presenta adjunción obligatoria, el análisis debe continuar el reconocimiento descendente del subárbol dominado por dicho nodo.

$$\mathcal{D}_E^{\text{Pred}} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[M^\gamma \rightarrow \bullet v, j, j \mid -, -]} \quad \text{nil} \in \text{adj}(M^\gamma)$$



### Compleción de subárbol

Este paso deductivo de completación de subárbol es igual al del esquema **buE**.

$$\mathcal{D}_E^{\text{Comp}} = \mathcal{D}_{\text{buE}}^{\text{Comp}}$$

### Predicción de adjunción

Cuando el reconocimiento alcanza un nodo adjuntable, el análisis debe lanzar el reconocimiento de todos los árboles auxiliares que se pueden adjuntar en dicho nodo.

$$\mathcal{D}_E^{\text{AdjPred}} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[\top \rightarrow \bullet \mathbf{R}^\beta, j, j \mid -, -]} \quad \beta \in \text{adj}(M^\gamma)$$

### Predicción de pie

Cuando el reconocimiento alcanza el nodo pie de un árbol auxiliar  $\beta$ , se debe continuar con el subárbol escindido por la operación de adjunción. Ni en los ítems ni en el chart existe información suficiente para determinar sobre qué subárbol se debe continuar el reconocimiento, y precisamente esta carencia es la que provoca que el analizador no cumpla la propiedad del prefijo válido. Por ello, la operación  $\mathcal{D}_E^{\text{FootPred}}$  se ve obligada a lanzar todos los subárboles dominados por nodos donde se pueda adjuntar  $\beta$ .

$$\mathcal{D}_E^{\text{FootPred}} = \frac{[\mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -]}{[M^\gamma \rightarrow \bullet \delta, k, k \mid -, -]} \quad \beta \in \text{adj}(M^\gamma)$$

### Compleción de pie

Cuando se completa el reconocimiento de un subárbol escindido por una adjunción, el análisis debe continuar con el contexto derecho del árbol auxiliar adjuntado.

$$\mathcal{D}_E^{\text{FootComp}} = \frac{[M^\gamma \rightarrow \delta \bullet, k, l \mid p, q], [\mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -]}{[\mathbf{F}^\beta \rightarrow \perp \bullet, k, l \mid k, l]} \quad \beta \in \text{adj}(M^\gamma)$$

### Compleción de adjunción

Una vez que se ha completado el reconocimiento de un árbol auxiliar, debemos continuar el reconocimiento del árbol donde se ha efectuado la adjunción.

$$\mathcal{D}_E^{\text{AdjComp}} = \frac{[\top \rightarrow \mathbf{R}^\beta \bullet, j, m \mid k, l], [M^\gamma \rightarrow \nu \bullet, k, l \mid p, q], [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p', q']}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, m \mid p \cup p', q \cup q']} \quad \beta \in \text{adj}(M^\gamma)$$

El conjunto de ítems finales del esquema es igual al del esquema **buE**:

$$\mathcal{F}_E = \mathcal{F}_{\text{buE}}$$

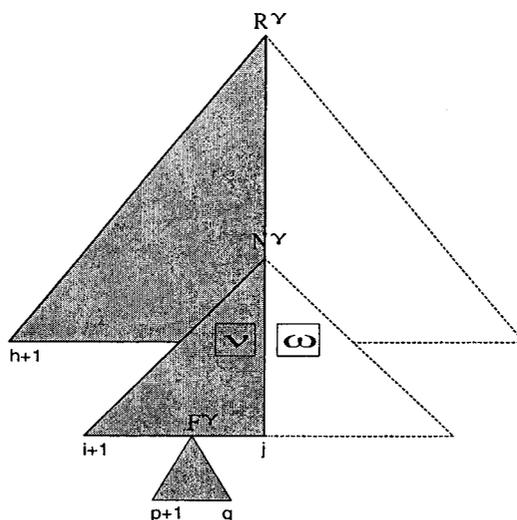


Figura 4.6: Ítems del esquema Nederhof

#### 4.5.1 Esquema tipo Earley con la propiedad del prefijo válido

En esta sección vamos a mostrar un esquema basado en el algoritmo de Earley que verifica la propiedad del prefijo válido. Este esquema fue definido por Nederhof en [84, 86], y su aportación fundamental radica en que es un algoritmo que verifica VPP con un coste temporal de  $\mathcal{O}(n^6)$ , igual al de analizadores que no verifican dicha propiedad, poniendo de manifiesto que no eran acertadas anteriores conjeturas sobre el coste adicional que supone para un analizador el mantener la VPP.

El conjunto de ítems del esquema se define mediante:

$$\mathcal{I}_{\text{Nederhof}} = \mathcal{I}_{\text{Nederhof}}^{(i)} \cup \mathcal{I}_{\text{Nederhof}}^{(ii)}$$

Para verificar la VPP es necesario conocer la posición en la cadena de entrada donde comienza cada árbol elemental que se introduce en el proceso de análisis, lo que obliga a añadir un índice más en la estructura de los ítems usados en el esquema **E**, dando lugar al siguiente tipo de ítems (ver figura 4.6):

$$\mathcal{I}_{\text{Nederhof}}^{(i)} = \{[N^\gamma \rightarrow \nu \bullet \omega, h, i, j \mid p, q]\}$$

donde  $N^\gamma \rightarrow \nu \omega \in \mathcal{P}(\gamma)$  y  $\gamma \in \mathbf{I} \cup \mathbf{A}$ . Los índices  $0 \leq i \leq j$  establecen las posiciones de la cadena de entrada que delimitan el fragmento reconocido por  $\nu$ . El índice  $h$  ( $0 \leq h \leq i$ ) indica la posición en la cadena de entrada donde se inicia el reconocimiento de  $\gamma$ . Si  $p$  y  $q$  presentan un valor definido, entonces  $N^\gamma$  pertenece a la espina de un árbol auxiliar,  $\nu$  domina el nodo pie de  $\gamma$  y se cumple  $(p, q) \leq (i, j)$ .

La introducción del nuevo índice provoca que la complejidad espacial en el caso peor aumente hasta  $\mathcal{O}(n^5)$  y, en una primera aproximación, también generaría un aumento de

la complejidad temporal. Con objeto de mantener la cota de complejidad temporal de los esquemas que no verifican la VPP, Nederhof propone la división en dos pasos de la operación de completión de adjunción, para lo que necesita definir un tipo de ítems intermedio de la siguiente forma:

$$\mathcal{I}_{\text{Nederhof}}^{(ii)} = \{[N^\gamma \rightarrow \nu \bullet, i, j \mid p, q]\}$$

donde  $N^\gamma \rightarrow \nu \in \mathcal{P}(\gamma)$  y  $\gamma \in \mathbf{I} \cup \mathbf{A}$ . Los índices  $0 \leq i \leq j$  establecen las posiciones de la cadena de entrada que delimitan el fragmento reconocido por  $N^\gamma$ . Si  $p$  y  $q$  presentan un valor definido, entonces  $N^\gamma$  pertenece a la espina de un árbol auxiliar,  $\nu$  domina el nodo pie de  $\gamma$  y se cumple  $(p, q) \leq (i, j)$ .

El conjunto de pasos deductivos del esquema  $\mathcal{D}_{\mathbb{E}}$  se define mediante:

$$\mathcal{D}_{\text{Nederhof}} = \mathcal{D}_{\text{Nederhof}}^{\text{Ini}} \cup \mathcal{D}_{\text{Nederhof}}^{\text{Scan}} \cup \mathcal{D}_{\text{Nederhof}}^{\varepsilon} \cup \mathcal{D}_{\text{Nederhof}}^{\text{Pred}} \cup \mathcal{D}_{\text{Nederhof}}^{\text{Comp}} \cup$$

$$\mathcal{D}_{\text{Nederhof}}^{\text{AdjPred}} \cup \mathcal{D}_{\text{Nederhof}}^{\text{FootPred}} \cup \mathcal{D}_{\text{Nederhof}}^{\text{FootComp}} \cup \mathcal{D}_{\text{Nederhof}}^{\text{AdjComp}^0} \cup \mathcal{D}_{\text{Nederhof}}^{\text{AdjComp}^1}$$

$$\mathcal{D}_{\text{Nederhof}}^{\text{Ini}} = \frac{[ \top \rightarrow \bullet \mathbf{R}^\alpha, 0, 0, 0 \mid -, - ]}{\alpha \in \mathbf{I}}$$

$$\mathcal{D}_{\text{Nederhof}}^{\text{Scan}} = \frac{\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, h, i, j - 1 \mid p, q]}{[a, j - 1, j]}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, h, i, j \mid p, q]} \quad \text{label}(M^\gamma) = a$$

$$\mathcal{D}_{\text{Nederhof}}^{\varepsilon} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, h, i, j \mid p, q]}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, h, i, j \mid p, q]} \quad \text{label}(M^\gamma) = \varepsilon$$

$$\mathcal{D}_{\text{Nederhof}}^{\text{Pred}} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, h, i, j \mid p, q]}{[M^\gamma \rightarrow \bullet \omega, h, j, j \mid -, -]} \quad \text{nil} \in \text{adj}(M^\gamma)$$

$$\mathcal{D}_{\text{Nederhof}}^{\text{Comp}} = \frac{\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, h, i, k \mid p, q]}{[M^\gamma \rightarrow \omega \bullet, h, k, j \mid p', q']}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, h, i, j \mid p \cup p', q \cup q']} \quad \text{nil} \in \text{adj}(M^\gamma)$$

$$\mathcal{D}_{\text{Nederhof}}^{\text{AdjPred}} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, h, i, j \mid p, q]}{[\top \rightarrow \bullet \mathbf{R}^\beta, j, j, j \mid -, -]} \quad \beta \in \text{adj}(M^\gamma)$$

$$\mathcal{D}_{\text{Nederhof}}^{\text{FootPred}} = \frac{\frac{[\mathbf{F}^\beta \rightarrow \bullet \perp, j, k, k \mid -, -]}{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, h, i, j \mid p, q]}}{[M^\gamma \rightarrow \bullet \omega, h, k, k \mid -, -]} \quad \beta \in \text{adj}(M^\gamma)$$

$$\mathcal{D}_{\text{Nederhof}}^{\text{FootComp}} = \frac{\begin{array}{l} [M^\gamma \rightarrow \omega \bullet, h, k, l \mid p, q] \\ [\mathbf{F}^\beta \rightarrow \bullet \perp, j, k, k \mid -, -] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, h, i, j \mid p', q'] \end{array}}{[\mathbf{F}^\beta \rightarrow \perp \bullet, j, k, l \mid k, l]} \quad \beta \in \text{adj}(M^\gamma)$$

$$\mathcal{D}_{\text{Nederhof}}^{\text{AdjComp}^0} = \frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\beta \bullet, j, j, m \mid k, l] \\ [M^\gamma \rightarrow \omega \bullet, h, k, l \mid p, q] \end{array}}{[[M^\gamma \rightarrow \omega \bullet, j, m \mid p, q]]} \quad \beta \in \text{adj}(M^\gamma)$$

$$\mathcal{D}_{\text{Nederhof}}^{\text{AdjComp}^1} = \frac{\begin{array}{l} [[M^\gamma \rightarrow \omega \bullet, j, m \mid p, q] \\ [M^\gamma \rightarrow \omega \bullet, h, k, l \mid p, q] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, h, i, j \mid p', q'] \end{array}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, h, i, m \mid p \cup p', q \cup q']}$$

El conjunto de ítems finales se define como:

$$\mathcal{F}_{\text{Nederhof}} = \{[\top \rightarrow \mathbf{R}^\alpha \bullet, 0, 0, n \mid -, -] \mid \alpha \in I\}$$

Las funcionalidades de todos los pasos deductivos son idénticas a las de sus homónimos del esquema tipo Earley sin VPP. El nuevo índice, en la mayoría de los pasos, se limita a ser transmitido al consecuente, aunque su introducción genera diferencias destacables en la forma en la que el analizador efectúa la predicción y compleción del pie. Diferencias que hacen que este analizador verifique la propiedad del prefijo válido.

Cuando el reconocimiento alcanza el nodo pie de un árbol auxiliar, el paso  $\mathcal{D}_{\text{Nederhof}}^{\text{FootPred}}$ , al igual que su homónimo en el esquema **E**, se encarga de continuar con el subárbol escindido por la operación de adjunción, pero, a diferencia de aquel, la información adicional que aporta el nuevo índice respecto a la posición de la cadena de entrada donde empezó a reconocerse el árbol auxiliar, permite determinar sobre qué subárbol se debe continuar el reconocimiento. De forma análoga, cuando se completa el reconocimiento de un subárbol escindido por una adjunción, el paso  $\mathcal{D}_{\text{Nederhof}}^{\text{FootComp}}$ , al igual que su homónimo en el esquema **E**, se encarga de establecer el fragmento de la cadena de entrada reconocido por el nodo pie del árbol auxiliar adjuntado e iniciar el análisis del contexto derecho del mismo, pero, a diferencia de aquel, la información adicional que aporta el nuevo índice respecto a la posición de la cadena de entrada donde empezaron a reconocerse el árbol auxiliar y el árbol al que pertenece el nodo adjuntado, permite asociar el subárbol escindido a los nodos pie de los árboles auxiliares adecuados.

Como dijimos previamente, este esquema también incluye una diferencia significativa en la forma de realizar la compleción de adjunción con respecto al esquema sin VPP. Si nos limitásemos a introducir el nuevo índice en el paso de compleción de adjunción de esquema

E, obtendríamos lo siguiente:

$$\mathcal{D}_{\text{Nederhof}}^{\text{AdjComp}} = \frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\beta \bullet, j, j, m \mid k, l], \\ [M^\gamma \rightarrow \omega \bullet, h, k, l \mid p, q], \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, h, i, j \mid p', q'] \end{array}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, h, i, m \mid p \cup p', q \cup q']} \quad \beta \in \text{adj}(M^\gamma)$$

Aparentemente la complejidad de este paso se sitúa en la cota de  $\mathcal{O}(n^8)$ , provocada por la aparición de 8 índices repetidos en los antecedentes. Sin embargo, podemos observar que los índices  $k$  y  $l$  sólo se emplean para establecer la relación entre los dos primeros antecedentes. Esta circunstancia la aprovecha Nederhof para dividir el paso mediante dos aplicaciones parciales. La primera, recogida en el paso  $\mathcal{D}_{\text{Nederhof}}^{\text{AdjComp}^0}$ , combina los dos primeros antecedentes y comprime en un ítem intermedio la información necesaria para la siguiente aplicación parcial. La segunda, recogida en el paso  $\mathcal{D}_{\text{Nederhof}}^{\text{AdjComp}^1}$ , combina el ítem intermedio resultado de la primera aplicación con el tercer antecedente del paso original, obteniendo el consecuente del paso original. En  $\mathcal{D}_{\text{Nederhof}}^{\text{AdjComp}^1}$  es necesario añadir también el segundo antecedente del paso original debido a que el ítem intermedio no transmite la información del índice  $h$ , el cual se requiere para establecer la relación entre los antecedentes segundo y tercero del paso original. Los dos nuevos pasos obtenidos presentan una complejidad temporal de  $\mathcal{O}(n^6)$ , siendo ésta la cota superior de todos los pasos y, por tanto, la complejidad del algoritmo.

La propuesta original de Nederhof incluye los siguientes tres pasos para la completación de adjunción:

$$\mathcal{D}_{\text{Nederhof}}^{\text{AdjComp}^0} = \frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\beta \bullet, j, j, m \mid k, l] \\ [M^\gamma \rightarrow \omega \bullet, h, k, l \mid p, q] \end{array}}{[[M^\gamma \rightarrow \omega \bullet, j, m \mid p, q]]} \quad \beta \in \text{adj}(M^\gamma)$$

$$\mathcal{D}_{\text{Nederhof}}^{\text{AdjComp}^1} = \frac{\begin{array}{l} [[M^\gamma \rightarrow \omega \bullet, j, m \mid p, q] \\ [\mathbf{F}^\gamma \rightarrow \perp \bullet, h, p, q \mid p, q] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, h, i, j \mid -, -] \end{array}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, h, i, m \mid p, q]}$$

$$\mathcal{D}_{\text{Nederhof}}^{\text{AdjComp}^2} = \frac{\begin{array}{l} [[M^\gamma \rightarrow \omega \bullet, j, m \mid -, -] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, h, i, j \mid p', q'] \end{array}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, h, i, m \mid p', q']}$$

Ésto es debido a que hace una distinción en la segunda aplicación parcial, en función de si el nodo  $M^\gamma$  pertenece ( $\mathcal{D}_{\text{Nederhof}}^{\text{AdjComp}^1}$ ) o no ( $\mathcal{D}_{\text{Nederhof}}^{\text{AdjComp}^2}$ ) a la espina de un árbol auxiliar. En el esquema presentado aquí se comprimen estos dos pasos en uno, mediante el uso de la operación de unión sobre índices y la eliminación del ítem redundante  $[\mathbf{F}^\gamma \rightarrow \perp \bullet, h, p, q \mid p, q]$ .

## Capítulo 5

# Algoritmos de análisis basados en Left Corner para TAG

Ahora presentaremos un conjunto de analizadores para TAG contruidos mediante la aplicación de filtros basados en la relación de Left Corner para TAG a los esquemas tipo Earley descritos en el capítulo anterior, tanto en su versión ascendente como predictiva. La relación de Left Corner para TAG se definirá como una extensión de la relación del mismo nombre para CFG.

El capítulo se estructura de la siguiente forma. En sección 1 se define un esquema tipo Left Corner ascendente, obtenido mediante la aplicación de un filtro basado en la esquina izquierda al analizador tipo Earley ascendente. También se presentan un conjunto de resultados experimentales que demuestran las mejoras que aporta este nuevo algoritmo en situaciones prácticas. En la sección 2 se presenta el concepto de relación de Left Corner en el contexto del formalismo TAG. En la sección 3 se define un esquema tipo Left Corner que hace uso de la relación de Left Corner para TAG definida en la sección 2, tanto la versión que verifica la VPP como la que no lo hace. En la sección 4 se presenta un esquema tipo Left Corner cuya único interés es el de servir de enlace para establecer las relaciones formales entre los esquemas de las secciones 3 y 5. En la sección 5 definimos un analizador de tipo Left Corner sin VPP que se obtiene aplicando un filtro basado en la relación de Left Corner para TAG al algoritmo tipo Earley sin VPP. Además, en la sección 5 se incluye una versión optimizada del esquema tipo Left Corner sin VPP y una comparativa, mediante resultados experimentales, de éste con el analizador tipo Earley sin VPP.

### 5.1 Esquema tipo Left Corner ascendente

El esquema **buLC** (*bottom-up Left Corner*) para TAG, que presentamos en [24], se obtiene mediante la aplicación de un filtro al esquema **buE**. El objetivo de este analizador es mejorar el comportamiento práctico de **buE**, disminuyendo el número de ítems que se generan

durante el proceso de análisis. Para alcanzar dicha mejora, vamos a eliminar los ítems que en el esquema **buE** no aportan nada significativo en el proceso constructivo, y que son aquellos de la forma:

$$[N^\gamma \rightarrow \bullet M^\gamma \omega, j, j, -, -]$$

Puesto que en la representación multicapa a cada nodo de los árboles elementales le asociamos una producción, y considerando cada producción como la representación de un subárbol de tan sólo un nivel, vamos a definir el concepto de *esquina izquierda*<sup>1</sup> de una producción y, por consiguiente, de su subárbol de un nivel asociado.

**Definición 5.1** *Esquina izquierda de una producción TAG*

Dada una producción  $N^\gamma \rightarrow P^\gamma \nu \in \mathcal{P}(\gamma)$ , denominaremos *esquina izquierda de esta producción al nodo  $P^\gamma$* .

El dominio del esquema  $\mathcal{I}_{\text{buLC}}$  se define mediante:

$$\mathcal{I}_{\text{buLC}} = \{[N^\gamma \rightarrow P^\gamma \nu \bullet \omega, i, j \mid p, q]\}$$

donde  $N^\gamma \rightarrow P^\gamma \nu \omega \in \mathcal{P}(\gamma)$  y  $\gamma \in \mathbf{I} \cup \mathbf{A}$ . Los índices  $0 \leq i \leq j$  establecen las posiciones de la cadena de entrada que delimitan el fragmento reconocido por  $P^\gamma \nu$ . Si  $p$  y  $q$  presentan un valor definido, entonces  $N^\gamma$  pertenece a la espina de un árbol auxiliar,  $P^\gamma \nu$  domina el nodo pie de  $\gamma$  y se cumple  $(p, q) \leq (i, j)$ .

Para eliminar los ítems con el punto delante de la esquina izquierda, vamos a sustituir el paso  $\mathcal{D}_{\text{buE}}^{\text{ini}}$ , que es el que introduce ítems de esta forma, por los cuatro siguientes:  $\mathcal{D}_{\text{buLC}}^{\text{LC}_t}$ ,  $\mathcal{D}_{\text{buLC}}^{\text{LC}_\varepsilon}$ ,  $\mathcal{D}_{\text{buLC}}^{\text{LC}_n}$  y  $\mathcal{D}_{\text{buLC}}^{\text{LC}_{\text{cad}}}$ . Los cuales llevarán a cabo el reconocimiento de la esquina izquierda de cada subárbol.

Los pasos deductivos del esquema vienen dados por:

$$\begin{aligned} \mathcal{D}_{\text{buLC}} = & \mathcal{D}_{\text{buLC}}^{\text{LC}_t} \cup \mathcal{D}_{\text{buLC}}^{\text{LC}_\varepsilon} \cup \mathcal{D}_{\text{buLC}}^{\text{LC}_n} \cup \mathcal{D}_{\text{buLC}}^{\text{LC}_{\text{cad}}} \cup \mathcal{D}_{\text{buLC}}^{\text{Foot}} \cup \\ & \mathcal{D}_{\text{buLC}}^{\text{Scan}} \cup \mathcal{D}_{\text{buLC}}^\varepsilon \cup \mathcal{D}_{\text{buLC}}^{\text{Comp}} \cup \mathcal{D}_{\text{buLC}}^{\text{AdjComp}} \end{aligned}$$

El paso deductivo  $\mathcal{D}_{\text{buLC}}^{\text{LC}_t}$  lanza el reconocimiento de los subárboles cuya esquina izquierda sea un símbolo terminal, entre todas las posiciones de la cadena de entrada donde aparece dicho símbolo. De esta forma eliminamos la posibilidad de que se generen producciones de la forma  $N^\gamma \rightarrow \bullet P^\gamma \omega$  cuando  $\text{label}(P^\gamma) \in V_T$ .  $\mathcal{D}_{\text{buLC}}^{\text{LC}_t}$  sustituye al paso  $\mathcal{D}_{\text{buLC}}^{\text{Scan}}$  en los símbolos de la esquina izquierda.

$$\mathcal{D}_{\text{buLC}}^{\text{LC}_t} = \frac{[a, j, j + 1]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, j + 1 \mid -, -]} \quad \text{label}(P^\gamma) = a$$

El paso  $\mathcal{D}_{\text{buLC}}^{\text{LC}_\varepsilon}$  funciona de forma análoga al anterior cuando la esquina izquierda es la cadena vacía.

$$\mathcal{D}_{\text{buLC}}^{\text{LC}_\varepsilon} = \frac{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, j \mid -, -]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, j \mid -, -]} \quad \text{label}(P^\gamma) = \varepsilon$$

<sup>1</sup>No hay que confundir este concepto con el de Left Corner que veremos más adelante.

En principio, el paso  $\mathcal{D}_{\text{buLC}}^{\text{Foot}}$  sería igual a  $\mathcal{D}_{\text{buE}}^{\text{Foot}}$ , sin embargo, podemos aprovechar la estrategia ascendente del reconocimiento para filtrar de forma dinámica los ítems que se generan mediante  $\mathcal{D}_{\text{buLC}}^{\text{Foot}}$ . Puesto que este paso completa el reconocimiento de los subárboles que dominan los nodos pie de los árboles auxiliares, su lanzamiento puede estar limitado a que previamente se haya reconocido el subárbol que cuelga de un nodo  $O^\gamma$  donde sea adjuntable un árbol auxiliar  $\beta$ . Con este planteamiento, el paso deductivo  $\mathcal{D}_{\text{buLC}}^{\text{Foot}}$  quedaría de la siguiente forma:

$$\mathcal{D}_{\text{buLC}}^{\text{Foot}} = \frac{[O^\gamma \rightarrow \nu \bullet, k, l \mid p, q]}{[\mathbf{F}^\beta \rightarrow \perp \bullet, k, l \mid k, l]} \quad \beta \in \text{adj}(O^\gamma)$$

El paso  $\mathcal{D}_{\text{buLC}}^{\text{LC}_n}$  inicia el reconocimiento de los subárboles que cumplen que: (1) su esquina izquierda sea un símbolo no terminal y (2) el subárbol que domina dicho símbolo haya sido completamente reconocido. Es decir, continúa el reconocimiento una vez completo el subárbol que domina la esquina izquierda.  $\mathcal{D}_{\text{buLC}}^{\text{LC}_n}$  es equivalente a la operación de completación en gramáticas incontextuales, por tanto, sólo se puede aplicar cuando la adjunción no sea obligatoria en la esquina izquierda.

$$\mathcal{D}_{\text{buLC}}^{\text{LC}_n} = \frac{[O^\gamma \rightarrow \nu \bullet, j, k \mid p, q]}{[Q^\gamma \rightarrow O^\gamma \bullet \omega, j, k \mid p, q]} \quad \text{nil} \in \text{adj}(O^\gamma)$$

Cuando la esquina izquierda de un subárbol sea un nodo adjuntable se aplica el paso deductivo  $\mathcal{D}_{\text{buLC}}^{\text{LC}_{\text{cad}}}$ , el cual requiere que se hayan reconocido completamente tanto el árbol auxiliar que se va a adjuntar como el subárbol que domina el nodo pie de dicho árbol. El paso  $\mathcal{D}_{\text{buLC}}^{\text{LC}_{\text{cad}}}$  sustituye a  $\mathcal{D}_{\text{buLC}}^{\text{AdjComp}}$  en los símbolos de la esquina izquierda.

$$\mathcal{D}_{\text{buLC}}^{\text{LC}_{\text{cad}}} = \frac{[\top \rightarrow \mathbf{R}^\beta \bullet, j, m \mid k, l]}{[O^\gamma \rightarrow \nu \bullet, k, l \mid p, q]} \quad \beta \in \text{adj}(O^\gamma)$$

El paso deductivo  $\mathcal{D}_{\text{buLC}}^{\text{Scan}}$  tiene la misma función que  $\mathcal{D}_{\text{buE}}^{\text{Scan}}$ , pero sólo se aplica a símbolos terminales que no sean esquinas izquierdas en sus subárboles.

$$\mathcal{D}_{\text{buLC}}^{\text{Scan}} = \frac{[a, j, j+1]}{[N^\gamma \rightarrow P^\gamma \nu \bullet M^\gamma \omega, i, j \mid p, q]} \quad \text{label}(M^\gamma) = a$$

El paso deductivo  $\mathcal{D}_{\text{buLC}}^\varepsilon$  tiene la misma función que  $\mathcal{D}_{\text{buE}}^\varepsilon$ , pero sólo se aplica a cadenas vacías que no sean esquinas izquierdas en sus subárboles.

$$\mathcal{D}_{\text{buLC}}^\varepsilon = \frac{[N^\gamma \rightarrow P^\gamma \nu \bullet M^\gamma \omega, i, j \mid p, q]}{[N^\gamma \rightarrow P^\gamma \nu M^\gamma \bullet \omega, i, j \mid p, q]} \quad \text{label}(M^\gamma) = \varepsilon$$

El paso  $\mathcal{D}_{\text{buLC}}^{\text{Comp}}$  tiene la misma función que  $\mathcal{D}_{\text{buE}}^{\text{Comp}}$ , pero sólo se aplica a símbolos no terminales sin restricción de adjunción obligatoria que no sean esquinas izquierdas en sus



subárboles.

$$\mathcal{D}_{\text{buLC}}^{\text{Comp}} = \frac{\begin{array}{l} [M^\gamma \rightarrow \delta \bullet, j, k \mid p, q] \\ [N^\gamma \rightarrow P^\gamma \nu \bullet M^\gamma \omega, i, j \mid p', q'] \end{array}}{[N^\gamma \rightarrow P^\gamma \nu M^\gamma \bullet \omega, i, k \mid p \cup p', q \cup q']} \quad \text{nil} \in \text{adj}(M^\gamma)$$

La función del paso deductivo  $\mathcal{D}_{\text{buLC}}^{\text{AdjComp}}$  es la misma que la de  $\mathcal{D}_{\text{buE}}^{\text{AdjComp}}$ , pero sólo se aplica a símbolos no terminales adjuntables que no sean esquinas izquierdas en sus subárboles.

$$\mathcal{D}_{\text{buLC}}^{\text{AdjComp}} = \frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\beta \bullet, j, m \mid k, l] \\ [M^\gamma \rightarrow \delta \bullet, k, l \mid p, q] \\ [N^\gamma \rightarrow P^\gamma \nu \bullet M^\gamma \omega, i, j \mid p', q'] \end{array}}{[N^\gamma \rightarrow P^\gamma \nu M^\gamma \bullet \omega, i, m \mid p \cup p', q \cup q']} \quad \beta \in \text{adj}(M^\gamma)$$

El conjunto de ítems finales del esquema es igual al del esquema **buE**:

$$\mathcal{F}_{\text{buLC}} = \mathcal{F}_{\text{buE}}$$

### 5.1.1 Resultados experimentales

El analizador **buLC** propuesto en esta sección presenta la misma complejidad teórica que el analizador **buE**, pero ahora vamos a ver como su comportamiento práctico es significativamente mejor. Aunque depende de la gramática y cadena de entrada consideradas, es habitual considerar el número de ítems deducidos en el proceso de análisis como medida fiable del comportamiento de un analizador deductivo.

Para llevar a cabo el estudio comparado del comportamiento de ambos analizadores para TAG vamos a usar: (1) una implementación de la máquina deductiva de análisis [122] usando el paradigma de programación lógica, la cual nos va a permitir animar sistemas de análisis; (2) un conjunto de siete gramáticas que recogen las características más relevantes de las gramáticas de adjunción de árboles.

El conjunto de gramáticas que se va a emplear como banco de pruebas lo construiremos a partir del conjunto de árboles elementales presentados en la figura 5.3, y lo podemos dividir en dos grupos atendiendo a su forma:

- Un conjunto de cuatro gramáticas que reconocen lenguajes regulares y que constan de árboles auxiliares recursivos por la izquierda y por la derecha. Este tipo de árboles elementales es muy frecuente en la definición de gramáticas para lenguajes naturales, como se puede comprobar en [144]. Los árboles elementales que constituyen estas cuatro gramáticas son los siguientes:

- Gramática G1 =  $\{\alpha_1, \beta_0\}$
- Gramática G2 =  $\{\alpha_1, \beta_1\}$
- Gramática G3 =  $\{\alpha_0, \beta_0, \beta_1\}$
- Gramática G4 =  $\{\alpha_1, \beta_0, \beta_1\}$

$n$	Gramática G1		Gramática G2		Gramática G3		Gramática G4	
	buLC	buE	buLC	buE	buLC	buE	buLC	buE
1	4	18	4	20	19	33	6	32
2	12	35	13	39	45	66	20	65
3	25	58	28	65	89	117	46	116
4	44	88	50	99	159	194	90	193
5	70	126	80	142	265	307	160	306
6	104	173	119	195	419	468	266	467
7	147	230	168	259	635	691	420	690
8	200	298	228	335	929	992	636	991
9	264	378	300	424	1319	1389	930	1388
10	340	471	384	527	1825	1902	1320	1901

Figura 5.1: Comparativa entre los esquemas tipo Earley y Left Corner ascendentes (1)

Hemos probado estas gramáticas con cadenas de entrada de longitudes entre 1 y 10 símbolos, obteniendo los resultados que se muestran en la figura 5.1. La media de reducción en el número de ítems deducidos es: 46% en las gramáticas G1 y G2, 16% en la gramática G3, y 49% en la gramática G4. Se observa una reducción muy significativa en el número de ítems deducidos, especialmente en la gramática que presenta recursión tanto por la izquierda como por la derecha. Por otra parte, también hemos observado una sensible disminución en el número de operaciones de compleción de adjunción.

- Un conjunto de tres gramáticas que reconocen lenguajes que muestran la capacidad generativa de las TAG, incluyendo lenguajes dependientes e independientes del contexto. Las tres incluyen árboles auxiliares con espigas de longitud tres y cuyos nodos son adjuntables. Los árboles elementales de estas gramáticas son los siguientes:

– Gramática  $G5 = \{\alpha_0, \beta_2, \beta_3\}$

– Gramática  $G6 = \{\alpha_0, \beta_4\}$

– Gramática  $G7 = \{\alpha_1, \beta_5\}$

Aquí las cadenas de entrada tendrán una longitud entre 2 y 41 símbolos, obteniendo los resultados mostrados en la figura 5.2. La media de reducción de ítems es muy elevada en los tres casos, siendo de un 74% en las gramáticas G5 y G6, y de un 91% en la gramática G7. En las gramáticas G5 y G6 también se produce una reducción significativa en el número de compleciones de adjunción, aunque este dato no se muestra en la tabla.

$n$	Gramática G5		Gramática G6		Gramática G7	
	$a^n b^n$		$a^n b^n c^n$		$a^n b^n e c^n d^n$	
	buLC	buE	buLC	buE	buLC	buE
1	20	49	20	46	12	70
2	38	102	37	99	20	146
3	57	173	54	168	28	245
4	77	264	71	253	36	367
5	98	377	88	354	44	512
6	120	514	105	471	52	680
7	143	677	122	604	60	871
8	167	868	139	753	68	1085
9	192	1089	156	918	75	1322
10	218	1342	173	1099	83	1582

Figura 5.2: Comparativa entre los esquemas tipo Earley y Left Corner ascendentes (2)

## 5.2 La relación Left Corner en TAG

Primero veamos que se entiende por relación de Left Corner en las CFG.

**Definición 5.2** *Relación de Left Corner (LC) en CFG*

El Left Corner es el símbolo de la izquierda del lado derecho de una producción.

La relación Left Corner  $>_\ell$  en  $V_N \times \{V_N \cup V_T \cup \{\varepsilon\}\}$  se define como

$A >_\ell B$  si hay una producción  $A \rightarrow Bv \in \mathcal{P}(G)$ .

La clausura reflexiva y transitiva de  $>_\ell$  la denotaremos como  $>_\ell^*$ .

Una relación LC en las CFG comienza en un nodo etiquetado con un símbolo no terminal y finaliza en un nodo etiquetado con un símbolo terminal o  $\varepsilon$ . Esta relación almacena la cadena de predicciones entre un símbolo, al que denominaremos *ancla* de la relación, y sus hijos izquierdos hasta llegar a un elemento terminal. Lo interesante es que estas predicciones se encuentran *precompiladas* y no son necesarias llevarlas a cabo durante el proceso de análisis de un algoritmo predictivo.

En las siguientes secciones vamos a definir esquemas que usan una extensión del concepto de relación de Left Corner en CFG para filtrar las predicciones en el analizador basado en el algoritmo de Earley para TAG. La complejidad temporal de todos estos algoritmos se mantienen en la cota de  $\mathcal{O}(n^6)$ , pero mejoran sus prestaciones mediante una reducción en el número de ítems deducidos.

Definamos el concepto de relación de Left Corner en los árboles elementales de las TAG.

**Definición 5.3** *Relación de Left Corner (LC) en los árboles elementales de las TAG*

El Left Corner de un nodo  $O^\gamma$  es su hijo izquierdo  $P^\gamma$  si y sólo si  $\text{adj}(P^\gamma) = \{\text{nil}\}$ .

La relación Left Corner  $>_\ell$  en  $V_N \times \{V_N \cup V_T \cup \{\varepsilon, \perp\}\}$  se define como

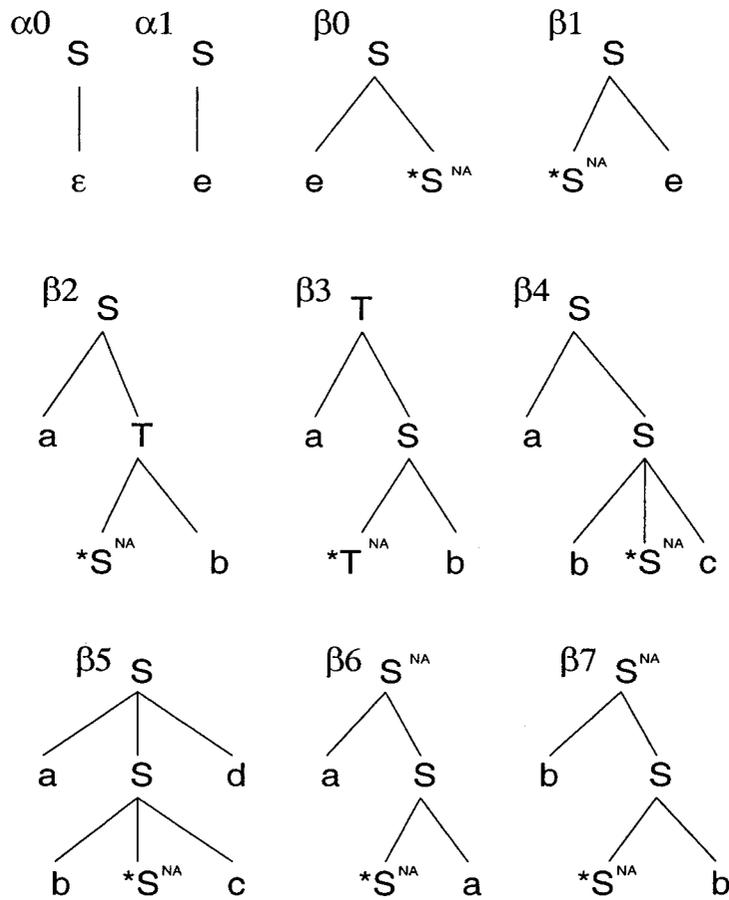


Figura 5.3: Conjunto de árboles elementales para experimentos

$O^\gamma >_\ell P^\gamma$  si hay una producción  $O^\gamma \rightarrow P^\gamma \nu \in \mathcal{P}(\gamma)$  y  $\text{adj}(P^\gamma) = \{\text{nil}\}$ .

La clausura reflexiva y transitiva de  $>_\ell$  la denotaremos como  $>_\ell^*$ .

Las diferencias fundamentales respecto a la definición de relación LC en las CFG son debidas a la operación de adjunción. Así, para evitar los problemas que pueden surgir con la adjunción cuando se lleva a cabo el reconocimiento ascendente a través de los nodos de una relación LC en las TAG, restringimos su ámbito a los límites de un árbol elemental. Obsérvese también que toda relación de LC en las TAG comienza en un nodo etiquetado con un símbolo no terminal y finaliza en un nodo adjuntable o un nodo etiquetado con un símbolo terminal,  $\epsilon$  o  $\perp$ . Ésto es debido a que la relación LC en las TAG no es posible cuando el hijo izquierdo de la producción es:

- Un nodo adjuntable, ya que la inclusión de un árbol auxiliar detiene la relación de LC con sus descendientes.

- Un nodo  $\perp$ , ya que la relación de esquina izquierda no va más allá de un árbol elemental.

Hay que tener en cuenta que el hecho de incluir, mediante una adjunción, un árbol auxiliar entre el nodo  $M^\gamma$  y el subárbol que domina, provoca que  $M^\gamma$  se duplique, apareciendo como raíz y pie del árbol auxiliar adjuntado. Es evidente que el nodo  $M^\gamma$  que funciona como raíz, al insertar un árbol auxiliar, ha perdido su relación de esquina izquierda respecto a su subárbol, pero el nodo  $M^\gamma$  que funciona como pie la sigue manteniendo.

Abusando de la notación, vamos a denotar como  $P^\gamma >_\ell \Delta$  si hay una producción  $O^\gamma \rightarrow P^\gamma \nu \in \mathcal{P}(\gamma)$  y existe un  $\beta$  tal que  $\beta \in \text{adj}(P^\gamma)$ .

### 5.3 Esquemas tipo Left Corner con ítems predictivos

En este esquema, al que denominaremos **pLC**, los pasos predictivos en el esquema **E** son reemplazados por objetivos que se intentan satisfacer de forma ascendente. La fase ascendente del proceso de reconocimiento es guiada hacia el correspondiente objetivo mediante la relación de esquina izquierda. En el dominio del esquema **pLC** vamos a distinguir dos tipos de ítems: predictivos y *left corner*. Por la propia naturaleza del analizador estos últimos se dividen en dos subtipos.

$$\mathcal{I}_{\text{pLC}} = \mathcal{I}_{\text{pLC}}^{\text{p}} \cup \mathcal{I}_{\text{pLC}}^{\text{lc}} \cup \mathcal{I}_{\text{pLC}}^{\text{lc}'}$$

Los ítems predictivos son aquellos que lanzan la predicción, de adjunción o subárbol, de un nodo  $M^\gamma$  desde una posición de la cadena de entrada. En este tipo de ítems sólo son necesarias la información del propio nodo y la posición de la cadena de entrada.

$$\mathcal{I}_{\text{pLC}}^{\text{p}} = \{[M^\gamma, j]\}$$

donde  $M^\gamma \in V_N$  y  $\gamma \in I \cup A$ . El índice  $0 \leq j$  establece la posición de la cadena de entrada donde se inicia el reconocimiento del nodo  $M^\gamma$ .

Los ítems *left corner* (ver figura 5.4) básicamente tienen la misma forma que los ítems del analizador tipo Earley para las TAG, pero se le añade delante la información del nodo  $C^\gamma$  que domina por una relación de esquina izquierda al nodo situado a la izquierda de la producción  $M^\gamma$ . Esta información adicional, aunque no es necesaria, va a permitir al analizador simplificar algunos de sus pasos deductivos y el reconocimiento ascendente a través de una secuencia de relaciones de esquinas izquierdas.

El esquema **pLC** filtra los ítems de **E** eliminando de su dominio aquellos ítems que inician el reconocimiento en la esquina izquierda de una producción.

$$\mathcal{I}_{\text{pLC}}^{\text{lc}} = \{[C^\gamma; M^\gamma \rightarrow \delta \bullet \nu, i, j \mid p, q]\}$$

donde  $M^\gamma \rightarrow \delta \nu \in \mathcal{P}(\gamma)$  y  $\gamma \in I \cup A$ . Los índices  $0 \leq i \leq j$  establecen las posiciones de la cadena de entrada que delimitan el fragmento reconocido por  $\delta$ . Si  $p$  y  $q$  presentan un valor definido, entonces  $M^\gamma$  pertenece a la espina de un árbol auxiliar,  $\delta$  domina el nodo pie

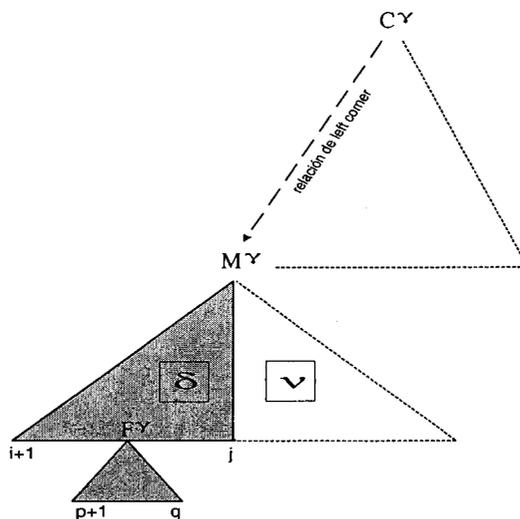


Figura 5.4: Ítems left corner del esquema pLC

de  $\gamma$  y se cumple  $(p, q) \leq (i, j)$ .  $C^\gamma$  es el ancla de una relación LC y domina al nodo  $M^\gamma$  ( $C^\gamma >_\ell^* M^\gamma$ ). Y para evitar que el punto aparezca a la izquierda de la producción, se exige que  $\delta \neq \varepsilon$ .

Sin embargo, este filtro no es posible cuando la esquina izquierda de la producción  $P^\gamma$  es un nodo adjuntable o un nodo  $\perp$ . Para recoger estos dos casos especiales, definimos la siguiente clase de ítems *left corner*:

$$I_{pLC}^{lc'} = \{[C^\gamma; M^\gamma \rightarrow \bullet P^\gamma \nu, j, j \mid -, -]\}$$

donde  $M^\gamma \rightarrow P^\gamma \nu \in \mathcal{P}(\gamma)$  y  $\gamma \in I \cup A$ . El índice  $0 \leq j$  establece la posición de la cadena de entrada donde se inicia el reconocimiento de la regla.  $C^\gamma$  es el ancla de una relación LC y domina al nodo  $M^\gamma$  ( $C^\gamma >_\ell^* M^\gamma$ ). Y  $P^\gamma >_\ell \Delta$  ó  $label(P^\gamma) = \perp$ .

Con respecto al conjunto de pasos deductivos, definimos subconjuntos para reconocimiento y compleción similares a los del esquema E para TAG. La relación de esquina izquierda se aplicará a los cuatro casos de predicción: inicial, subárbol, pie y adjunción. Los pasos de esquina izquierda vienen en tres variedades, según el tipo de nodo en que finalice la relación: terminal, cadena vacía y no terminal. El último caso es necesario cuando la esquina izquierda es un nodo de adjunción o el nodo *bottom* de un árbol auxiliar, ya que el árbol auxiliar o el subárbol escindido mediante una adjunción deben ser reconocidos. El conjunto de pasos deductivos del esquema es:

$$\begin{aligned} \mathcal{D}_{pLC} = & \mathcal{D}_{pLC}^{LI_t} \cup \mathcal{D}_{pLC}^{LI_\varepsilon} \cup \mathcal{D}_{pLC}^{LI_{pre}} \cup \mathcal{D}_{pLC}^{Scan} \cup \mathcal{D}_{pLC}^\varepsilon \cup \mathcal{D}_{pLC}^{LC_t} \cup \mathcal{D}_{pLC}^{LC_\varepsilon} \cup \\ & \mathcal{D}_{pLC}^{LC_{pre}} \cup \mathcal{D}_{pLC}^{LC_n} \cup \mathcal{D}_{pLC}^{Pre} \cup \mathcal{D}_{pLC}^{Comp} \cup \mathcal{D}_{pLC}^{LA_t} \cup \mathcal{D}_{pLC}^{LA_\varepsilon} \cup \end{aligned}$$



$$\mathcal{D}_{\text{pLC}}^{\text{LApre}} \cup \mathcal{D}_{\text{pLC}}^{\text{AdjComp}} \cup \mathcal{D}_{\text{pLC}}^{\text{LF}_t} \cup \mathcal{D}_{\text{pLC}}^{\text{LF}_\varepsilon} \cup \mathcal{D}_{\text{pLC}}^{\text{LFpre}} \cup \mathcal{D}_{\text{pLC}}^{\text{FootComp}}$$

### Filtrado de inicio

El reconocimiento comienza prediciendo todos los árboles iniciales. Dado que siempre se cumplirá que  $\top >_\ell^* O^\alpha$  y  $O^\alpha \rightarrow P^\alpha \nu \in \mathcal{P}(\alpha)$ , podemos aplicar un filtro LC y obtenemos los siguientes pasos:

$$\begin{aligned} \mathcal{D}_{\text{pLC}}^{\text{LI}_t} &= \frac{[a, 0, 1]}{[\top; O^\alpha \rightarrow P^\alpha \bullet \nu, 0, 1 \mid -, -]} & \text{label}(P^\gamma) &= a \\ & & \alpha &\in \mathbf{I} \\ \mathcal{D}_{\text{pLC}}^{\text{LI}_\varepsilon} &= \frac{[a, 0, 0]}{[\top; O^\alpha \rightarrow P^\alpha \bullet \nu, 0, 0 \mid -, -]} & \text{label}(P^\alpha) &= \varepsilon \\ & & \alpha &\in \mathbf{I} \\ \mathcal{D}_{\text{pLC}}^{\text{LIpre}} &= \frac{[a, 0, 0]}{[\top; O^\alpha \rightarrow \bullet P^\alpha \nu, 0, 0 \mid -, -]} & P^\alpha &>_\ell \Delta \\ & & \alpha &\in \mathbf{I} \end{aligned}$$

El paso  $\mathcal{D}_{\text{pLC}}^{\text{LI}_t}$  se aplicará cuando la esquina izquierda  $P^\alpha$  está etiquetada con un símbolo terminal. El paso  $\mathcal{D}_{\text{pLC}}^{\text{LI}_\varepsilon}$  se usa cuando  $P^\alpha$  es la cadena vacía. Y si  $P^\alpha$  es una nodo de adjunción entonces se aplica  $\mathcal{D}_{\text{pLC}}^{\text{LIpre}}$ . Los árboles elementales cuyas predicciones se filtran en este grupo son árboles iniciales, por lo que no es necesario considerar el caso en que la esquina izquierda sea un nodo  $\perp$ .

### Reconocimiento

El paso deductivo  $\mathcal{D}_{\text{pLC}}^{\text{Scan}}$  se aplica cuando el reconocimiento alcanza un nodo cuya etiqueta coincide con el símbolo terminal de la cadena de entrada. Es evidente que este terminal no puede ser la esquina izquierda de una producción, ya que ese tipo de ítems no pertenecen al dominio del esquema.

$$\mathcal{D}_{\text{pLC}}^{\text{Scan}} = \frac{[C^\gamma; N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \nu, i, j \mid p, q] \quad [a, j, j+1]}{[C^\gamma; N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, i, j+1 \mid p, q]} \quad \text{label}(M^\gamma) = a$$

Para efectuar el reconocimiento de símbolos  $\varepsilon$  se usa el paso deductivo  $\mathcal{D}_{\text{pLC}}^\varepsilon$ .

$$\mathcal{D}_{\text{pLC}}^\varepsilon = \frac{[C^\gamma; N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[C^\gamma; N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, i, j \mid p, q]} \quad \text{label}(M^\gamma) = \varepsilon$$

### Filtrado de predicción de subárbol

Si un nodo  $M^\gamma$ , cuyo reconocimiento ha sido predicho, no presenta una restricción de adjunción obligatoria y además domina por una relación de esquina izquierda a otro nodo  $O^\gamma$ , se podrían eliminar todas las predicciones entre ambos. En el caso de que el hijo izquierdo de  $O^\gamma$  esté etiquetado con un símbolo terminal se aplica el paso  $\mathcal{D}_{\text{pLC}}^{\text{LC}_t}$ .

$$\mathcal{D}_{\text{pLC}}^{\text{LC}_t} = \frac{[M^\gamma, j] \quad [a, j, j+1]}{[M^\gamma; O^\gamma \rightarrow P^\gamma \bullet \nu, j, j+1 \mid -, -]} \quad \begin{array}{l} \text{label}(P^\gamma) = a \\ \text{nil} \in \text{adj}(M^\gamma) \end{array}$$

El paso  $\mathcal{D}_{\text{pLC}}^{\text{LC}\varepsilon}$  tiene la misma función que  $\mathcal{D}_{\text{pLC}}^{\text{LC}i}$ , pero se aplica cuando el hijo izquierdo de  $O^\gamma$  es un símbolo etiquetado como  $\varepsilon$ .

$$\mathcal{D}_{\text{pLC}}^{\text{LC}\varepsilon} = \frac{[M^\gamma, j]}{[M^\gamma; O^\gamma \rightarrow P^\gamma \bullet \nu, j, j \mid -, -]} \quad \begin{array}{l} \text{label}(P^\gamma) = \varepsilon \\ \text{nil} \in \text{adj}(M^\gamma) \end{array}$$

Cuando el hijo izquierdo de  $O^\gamma$  sea un nodo adjuntable, se elimina la relación de esquina izquierda y se detiene el reconocimiento para lanzar los árboles auxiliares adjuntables en dicho nodo. Lo mismo ocurre cuando  $\gamma \in \mathbf{A}$  y el hijo izquierdo de  $O^\gamma$  está etiquetado con  $\perp$ , ya que hay que iniciar el reconocimiento del subárbol escindido en el nodo donde se ha llevado a cabo la adjunción de este árbol auxiliar.

$$\mathcal{D}_{\text{pLC}}^{\text{LCpre}} = \frac{[M^\gamma, j]}{[M^\gamma; O^\gamma \rightarrow \bullet P^\gamma \nu, j, j \mid -, -]} \quad \begin{array}{l} \text{nil} \in \text{adj}(M^\gamma) \\ P^\gamma >_\ell \Delta \text{ o } \text{label}(P^\gamma) = \perp \end{array}$$

### Compleción en la esquina izquierda

El recorrido ascendente a través de los nodos dominados por una relación de esquina izquierda se lleva a cabo mediante el paso  $\mathcal{D}_{\text{pLC}}^{\text{LCn}}$ . Esta operación es la equivalente a la completación de un subárbol para los nodos que se encuentran dentro de una relación de esquina izquierda. Es importante comprobar como el elemento auxiliar incluido en los ítems (nodo ancla de la relación esquina izquierda) permite establecer el fin de este recorrido ascendente, ya que la cadena de operaciones  $\mathcal{D}_{\text{pLC}}^{\text{LCn}}$  se detendrá cuando se alcance dicho nodo.

$$\mathcal{D}_{\text{pLC}}^{\text{LCn}} = \frac{[M^\gamma; O^\gamma \rightarrow \nu \bullet, j, k \mid p, q]}{[M^\gamma; Q^\gamma \rightarrow O^\gamma \bullet \omega, j, k \mid p, q]} \quad M^\gamma \neq O^\gamma$$

### Predicción

Mediante el paso  $\mathcal{D}_{\text{pLC}}^{\text{Pre}}$  se realiza la predicción de los nodos etiquetados con símbolos no terminales que no se encuentran dominados en una relación de esquina izquierda. Estos nodos se convertirán en anclas dentro de ítems de tipo *left corner*.

$$\mathcal{D}_{\text{pLC}}^{\text{Pre}} = \frac{[C^\gamma; N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[M^\gamma, j]} \quad \text{label}(M^\gamma) \in V_N$$

### Compleción de subárbol

El paso  $\mathcal{D}_{\text{pLC}}^{\text{Comp}}$  efectúa el reconocimiento ascendente de los subárboles dominados por nodos que no se encuentran en la esquina izquierda. También lleva a cabo la completación aunque el nodo sea esquina izquierda, siempre que éste sea un nodo adjuntable pero sin restricción de adjunción obligatoria.

$$\mathcal{D}_{\text{pLC}}^{\text{Comp}} = \frac{\begin{array}{l} [C^\gamma; N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q] \\ [M^\gamma; M^\gamma \rightarrow \nu \bullet, j, k \mid p', q'] \end{array}}{[C^\gamma; N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k \mid p \cup p', q \cup q']} \quad \text{nil} \in \text{adj}(M^\gamma)$$



### Filtrado de predicción de adjunción

Cuando se predice un nodo que es adjuntable, hay que lanzar el reconocimiento de todos los árboles auxiliares  $\beta$  que puedan ser adjuntados sobre él. No obstante, en lugar de comenzar el reconocimiento en la raíces de estos árboles auxiliares, se bajará hasta aquellos nodos  $O^\beta$  que estén dominados por una relación LC por los nodos  $\top$ . En función del tipo de nodo que sea hijo izquierdo de  $O^\beta$ , se distinguen tres pasos deductivos:  $\mathcal{D}_{\text{pLC}}^{\text{LA}_t}$ ,  $\mathcal{D}_{\text{pLC}}^{\text{LA}_\varepsilon}$  y  $\mathcal{D}_{\text{pLC}}^{\text{LA}_{\text{pre}}}$ .

Si el hijo izquierdo de  $O^\beta$  está etiquetado con un símbolo terminal se aplica  $\mathcal{D}_{\text{pLC}}^{\text{LA}_t}$ .

$$\mathcal{D}_{\text{pLC}}^{\text{LA}_t} = \frac{\begin{array}{c} [M^\gamma, j] \\ [a, j, j+1] \end{array}}{[\top; O^\beta \rightarrow P^\beta \bullet \nu, j, j+1 \mid -, -]} \quad \begin{array}{l} \text{label}(P^\beta) = a \\ \beta \in \text{adj}(M^\gamma) \end{array}$$

Si el hijo izquierdo de  $O^\beta$  está etiquetado con  $\varepsilon$  se aplica  $\mathcal{D}_{\text{pLC}}^{\text{LA}_\varepsilon}$ .

$$\mathcal{D}_{\text{pLC}}^{\text{LA}_\varepsilon} = \frac{[M^\gamma, j]}{[\top; O^\beta \rightarrow P^\beta \bullet \nu, j, j \mid -, -]} \quad \begin{array}{l} \text{label}(P^\beta) = \varepsilon \\ \beta \in \text{adj}(M^\gamma) \end{array}$$

Si el hijo izquierdo de  $O^\beta$  es un nodo adjuntable o el nodo etiquetado  $\perp$  se aplica  $\mathcal{D}_{\text{pLC}}^{\text{LA}_{\text{pre}}}$ , que posteriormente lanzará la predicción de los árboles auxiliares adjuntables o del subárbol escindido por el árbol auxiliar  $\beta$ .

$$\mathcal{D}_{\text{pLC}}^{\text{LA}_{\text{pre}}} = \frac{[M^\gamma, j]}{[\top; O^\beta \rightarrow \bullet P^\beta \nu, j, j \mid -, -]} \quad \begin{array}{l} \beta \in \text{adj}(M^\gamma) \\ P^\beta >_\ell \Delta \text{ o } \text{label}(P^\beta) = \perp \end{array}$$

### Compleción de adjunción

Una vez que se ha completado el reconocimiento de un árbol auxiliar hay que continuar el reconocimiento del árbol sobre el que se ha completado la adjunción.

$$\mathcal{D}_{\text{pLC}}^{\text{AdjComp}} = \frac{\begin{array}{c} [C^\gamma; N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p', q'] \\ [\top; \top \rightarrow \mathbf{R}^\beta \bullet, j, m \mid k, l] \\ [M^\gamma; M^\gamma \rightarrow \omega \bullet, k, l \mid p, q] \end{array}}{[C^\gamma; N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, m \mid p \cup p', q \cup q']} \quad \beta \in \text{adj}(M^\gamma)$$

### Filtrado de predicción de pie

Cuando un árbol auxiliar  $\beta$ , adjuntable en un nodo  $M^\gamma$ , se ha reconocido hasta su nodo  $\perp$ , se debe iniciar el reconocimiento del subárbol que domina  $M^\gamma$ . Al tratarse de un analizador que no posee la propiedad del prefijo válido, se debe lanzar el reconocimiento de todos los subárboles donde puede ser adjuntado el árbol auxiliar  $\beta$ .

Sobre el nodo  $M^\gamma$  que pende del pie se puede aplicar un filtro en las predicciones, bajando hasta el nodo que sea su esquina izquierda  $O^\gamma$ . Como en casos anteriores, en función del tipo de nodo que sea el hijo izquierdo de  $O^\gamma$  se aplica uno de los siguientes pasos deductivos:  $\mathcal{D}_{\text{pLC}}^{\text{LF}_t}$ ,  $\mathcal{D}_{\text{pLC}}^{\text{LF}_\varepsilon}$  y  $\mathcal{D}_{\text{pLC}}^{\text{LF}_{\text{pre}}}$ .

$$\mathcal{D}_{\text{pLC}}^{\text{LF}_t} = \frac{[E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -]}{[a, k, k+1]} \quad \begin{array}{l} \text{label}(P^\gamma) = a \\ \beta \in \text{adj}(M^\gamma) \end{array}$$

$$\mathcal{D}_{\text{pLC}}^{\text{LF}_\varepsilon} = \frac{[E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -]}{[M^\gamma; O^\gamma \rightarrow P^\gamma \bullet \nu, k, k \mid -, -]} \quad \begin{array}{l} \text{label}(P^\gamma) = \varepsilon \\ \beta \in \text{adj}(M^\gamma) \end{array}$$

$$\mathcal{D}_{\text{pLC}}^{\text{LF}_{\text{pre}}} = \frac{[E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -]}{[M^\gamma; O^\gamma \rightarrow \bullet P^\gamma \nu, k, k \mid -, -]} \quad \begin{array}{l} \beta \in \text{adj}(M^\gamma) \\ P^\gamma >_\ell \Delta \text{ ó } \text{label}(P^\gamma) = \perp \end{array}$$

### Completación de pie

Cuando un subárbol escindido por una adjunción es completamente reconocido, hay que pasar a reconocer el contexto derecho del árbol auxiliar adjuntado mediante el paso deductivo  $\mathcal{D}_{\text{pLC}}^{\text{FootComp}}$ .

$$\mathcal{D}_{\text{pLC}}^{\text{FootComp}} = \frac{[E^\beta; F^\beta \rightarrow \bullet \perp, k, k \mid -, -]}{[M^\gamma; M^\gamma \rightarrow \nu \bullet, k, l \mid p, q]} \quad \beta \in \text{adj}(M^\gamma)$$

El conjunto de ítems finales del esquema es:

$$\mathcal{F}_{\text{pLC}} = \{[T; T \rightarrow \mathbf{R}^\alpha \bullet, 0, n \mid -, -] \mid \alpha \in \mathbf{I}\}$$

### 5.3.1 Esquema tipo Left Corner con la propiedad del prefijo válido

En esta sección definimos una versión del esquema **pLC** que verifica la VPP. Este nuevo esquema lo presentamos en [26] y lo vamos a denominar **pLC<sup>vpp</sup>**. Su dominio es muy similar al del esquema sin VPP, por lo que también distinguiremos dos tipos de ítems: predictivos y *left corner*.

$$\mathcal{I}_{\text{pLC}^{\text{vpp}}} = \mathcal{I}_{\text{pLC}^{\text{vpp}}}^{\text{p}} \cup \mathcal{I}_{\text{pLC}^{\text{vpp}}}^{\text{lc}} \cup \mathcal{I}_{\text{pLC}^{\text{vpp}}}^{\text{lc}'}$$

La semántica y forma de ambos tipos de ítems del dominio son iguales a las del esquema **pLC**, pero para conseguir que el algoritmo verifique la VPP es necesario que incluyamos en ambos casos la información de la posición de comienzo, respecto a la cadena de entrada, del árbol al que pertenece el nodo (ver figura 5.5).

$$\mathcal{I}_{\text{pLC}^{\text{vpp}}}^{\text{p}} = \{[M^\gamma, h, j]\}$$

donde  $M^\gamma \in V_N$  y  $\gamma \in \mathbf{I} \cup \mathbf{A}$ . El índice  $0 \leq j$  establece la posición de la cadena de entrada donde se inicia el reconocimiento del nodo  $M^\gamma$ . El índice  $h$  ( $0 \leq h \leq j$ ) indica la posición en la cadena de entrada donde se inicia el reconocimiento de  $\gamma$ .

$$\mathcal{I}_{\text{pLC}^{\text{vpp}}}^{\text{lc}} = \{[C^\gamma; M^\gamma \rightarrow \delta \bullet \nu, h, i, j \mid p, q]\}$$



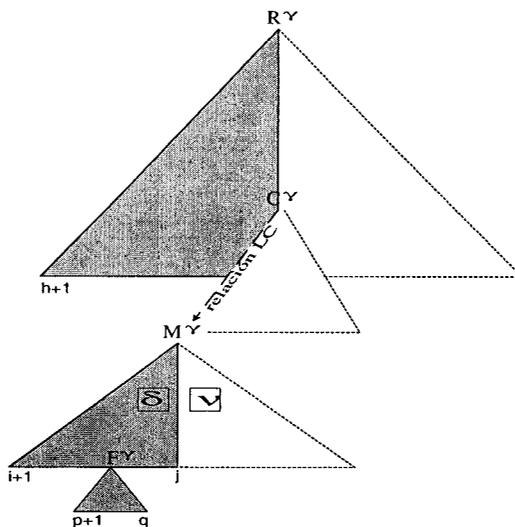


Figura 5.5: Ítems left corner del esquema  $pLC^{vpp}$

donde  $M^\gamma \rightarrow \delta\nu \in \mathcal{P}(\gamma)$  y  $\gamma \in I \cup A$ . Los índices  $0 \leq i \leq j$  establecen las posiciones de la cadena de entrada que delimitan el fragmento reconocido por  $\delta$ . El índice  $h$  ( $0 \leq h \leq i$ ) indica la posición en la cadena de entrada donde se inicia el reconocimiento de  $\gamma$ . Si  $p$  y  $q$  presentan un valor definido, entonces  $M^\gamma$  pertenece a la espina de un árbol auxiliar,  $\delta$  domina el nodo pie de  $\gamma$  y se cumple  $(p, q) \leq (i, j)$ .  $C^\gamma$  es el ancla de una relación LC y domina al nodo  $M^\gamma$  ( $C^\gamma >_\ell^* M^\gamma$ ). Para evitar que el punto aparezca a la izquierda de la producción, se exige que  $\delta \neq \varepsilon$ .

$$\mathcal{I}_{pLC^{vpp}}^{lc'} = \{[C^\gamma; M^\gamma \rightarrow \bullet P^\gamma \nu, h, j, j \mid -, -]\}$$

donde  $M^\gamma \rightarrow P^\gamma \nu \in \mathcal{P}(\gamma)$  y  $\gamma \in I \cup A$ . El índice  $0 \leq j$  establece la posición de la cadena de entrada donde se inicia el reconocimiento de la regla. El índice  $h$  ( $0 \leq h \leq j$ ) indica la posición en la cadena de entrada donde se inicia el reconocimiento de  $\gamma$ .  $C^\gamma$  es el ancla de una relación LC y domina al nodo  $M^\gamma$  ( $C^\gamma >_\ell^* M^\gamma$ ). Y  $P^\gamma >_\ell \Delta$  o  $label(P^\gamma) = \perp$ .

Los pasos deductivos del esquema son:

$$\begin{aligned} \mathcal{D}_{pLC^{vpp}} &= \mathcal{D}_{pLC^{vpp}}^{LI_t} \cup \mathcal{D}_{pLC^{vpp}}^{LI_\varepsilon} \cup \mathcal{D}_{pLC^{vpp}}^{LI_{pre}} \cup \mathcal{D}_{pLC^{vpp}}^{Scan} \cup \mathcal{D}_{pLC^{vpp}}^\varepsilon \cup \mathcal{D}_{pLC^{vpp}}^{LC_t} \cup \\ &\mathcal{D}_{pLC^{vpp}}^{LC_\varepsilon} \cup \mathcal{D}_{pLC^{vpp}}^{LC_{pre}} \cup \mathcal{D}_{pLC^{vpp}}^{LC_n} \cup \mathcal{D}_{pLC^{vpp}}^{Pre} \cup \mathcal{D}_{pLC^{vpp}}^{Comp} \cup \mathcal{D}_{pLC^{vpp}}^{LA_t} \cup \mathcal{D}_{pLC^{vpp}}^{LA_\varepsilon} \cup \\ &\mathcal{D}_{pLC^{vpp}}^{LA_{pre}} \cup \mathcal{D}_{pLC^{vpp}}^{AdjComp} \cup \mathcal{D}_{pLC^{vpp}}^{LF_t} \cup \mathcal{D}_{pLC^{vpp}}^{LF_\varepsilon} \cup \mathcal{D}_{pLC^{vpp}}^{LF_{pre}} \cup \mathcal{D}_{pLC^{vpp}}^{FootComp} \end{aligned}$$

$$\mathcal{D}_{pLC^{vpp}}^{LI_t} = \frac{[a, 0, 1]}{[T; O^\alpha \rightarrow P^\alpha \bullet \nu, 0, 0, 1 \mid -, -]} \quad \begin{array}{l} label(P^\gamma) = a \\ \alpha \in I \end{array}$$

$$\mathcal{D}_{\text{pLC}^{\text{vdp}}}^{\text{LI}_\varepsilon} = \frac{}{[\top; O^\alpha \rightarrow P^\alpha \bullet \nu, 0, 0, 0 \mid -, -]} \quad \begin{array}{l} \text{label}(P^\alpha) = \varepsilon \\ \alpha \in \mathbf{I} \end{array}$$

$$\mathcal{D}_{\text{pLC}^{\text{vdp}}}^{\text{LI}_{\text{pre}}} = \frac{}{[\top; O^\alpha \rightarrow \bullet P^\alpha \nu, 0, 0, 0 \mid -, -]} \quad \begin{array}{l} P^\alpha >_\ell \Delta \\ \alpha \in \mathbf{I} \end{array}$$

$$\mathcal{D}_{\text{pLC}^{\text{vdp}}}^{\text{Scan}} = \frac{\begin{array}{l} [C^\gamma; N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \nu, h, i, j \mid p, q] \\ [a, j, j+1] \end{array}}{[C^\gamma; N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, h, i, j+1 \mid p, q]} \quad \text{label}(M^\gamma) = a$$

$$\mathcal{D}_{\text{pLC}^{\text{vdp}}}^\varepsilon = \frac{[C^\gamma; N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \nu, h, i, j \mid p, q]}{[C^\gamma; N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, h, i, j \mid p, q]} \quad \text{label}(M^\gamma) = \varepsilon$$

$$\mathcal{D}_{\text{pLC}^{\text{vdp}}}^{\text{LC}_t} = \frac{\begin{array}{l} [M^\gamma, h, j] \\ [a, j, j+1] \end{array}}{[M^\gamma; O^\gamma \rightarrow P^\gamma \bullet \nu, h, j, j+1 \mid -, -]} \quad \begin{array}{l} \text{label}(P^\gamma) = a \\ \text{nil} \in \text{adj}(M^\gamma) \end{array}$$

$$\mathcal{D}_{\text{pLC}^{\text{vdp}}}^{\text{LC}_\varepsilon} = \frac{[M^\gamma, h, j]}{[M^\gamma; O^\gamma \rightarrow P^\gamma \bullet \nu, h, j, j \mid -, -]} \quad \begin{array}{l} \text{label}(P^\gamma) = \varepsilon \\ \text{nil} \in \text{adj}(M^\gamma) \end{array}$$

$$\mathcal{D}_{\text{pLC}^{\text{vdp}}}^{\text{LC}_{\text{pre}}} = \frac{[M^\gamma, h, j]}{[M^\gamma; O^\gamma \rightarrow \bullet P^\gamma \nu, h, j, j \mid -, -]} \quad \begin{array}{l} \text{label}(P^\gamma) = \perp \\ \text{nil} \in \text{adj}(M^\gamma) \end{array}$$

$$\mathcal{D}_{\text{pLC}^{\text{vdp}}}^{\text{LC}_n} = \frac{[M^\gamma; O^\gamma \rightarrow \nu \bullet, h, j, k \mid p, q]}{[M^\gamma; Q^\gamma \rightarrow O^\gamma \bullet \omega, h, j, k \mid p, q]} \quad M^\gamma \neq O^\gamma$$

$$\mathcal{D}_{\text{pLC}^{\text{vdp}}}^{\text{Pre}} = \frac{[C^\gamma; N^\gamma \rightarrow \delta \bullet M^\gamma \nu, h, i, j \mid p, q]}{[M^\gamma, h, j]} \quad \text{label}(M^\gamma) \in V_N$$

$$\mathcal{D}_{\text{pLC}^{\text{vdp}}}^{\text{Comp}} = \frac{\begin{array}{l} [C^\gamma; N^\gamma \rightarrow \delta \bullet M^\gamma \nu, h, i, j \mid p, q] \\ [M^\gamma; M^\gamma \rightarrow \omega \bullet, h, j, k \mid p', q'] \end{array}}{[C^\gamma; N^\gamma \rightarrow \delta M^\gamma \bullet \nu, h, i, k \mid p \cup p', q \cup q']} \quad \text{nil} \in \text{adj}(M^\gamma)$$

$$\mathcal{D}_{\text{pLC}^{\text{vdp}}}^{\text{LA}_t} = \frac{\begin{array}{l} [M^\gamma, h, j] \\ [a, j, j+1] \end{array}}{[\top; O^\beta \rightarrow P^\beta \bullet \nu, j, j, j+1 \mid -, -]} \quad \begin{array}{l} \text{label}(P^\gamma) = a \\ \beta \in \text{adj}(M^\gamma) \end{array}$$

$$\mathcal{D}_{\text{pLC}^{\text{vdp}}}^{\text{LA}_\varepsilon} = \frac{[M^\gamma, h, j]}{[\top; O^\beta \rightarrow P^\beta \bullet \nu, j, j, j \mid -, -]} \quad \begin{array}{l} \text{label}(P^\gamma) = \varepsilon \\ \beta \in \text{adj}(M^\gamma) \end{array}$$

$$\mathcal{D}_{\text{pLC}^{\text{VPP}}}^{\text{LA}_{\text{pre}}} = \frac{[M^\gamma, h, j]}{[\top; O^\beta \rightarrow \bullet P^\beta \nu, j, j, j \mid -, -]} \quad \begin{array}{l} P^\beta >_\ell \Delta \text{ o } \text{label}(P^\beta) = \perp \\ \beta \in \text{adj}(M^\gamma) \end{array}$$

$$\mathcal{D}_{\text{pLC}^{\text{VPP}}}^{\text{AdjComp}} = \frac{\begin{array}{l} [\top; \top \rightarrow \mathbf{R}^\beta \bullet, j, j, m \mid k, l] \\ [M^\gamma; M^\gamma \rightarrow \omega \bullet, h, k, l \mid p, q] \\ [C^\gamma; N^\gamma \rightarrow \delta \bullet M^\gamma \nu, h, i, j \mid p', q'] \end{array}}{[C^\gamma; N^\gamma \rightarrow \delta M^\gamma \bullet \nu, h, i, m \mid p \cup p', q \cup q']} \quad \beta \in \text{adj}(M^\gamma)$$

$$\mathcal{D}_{\text{pLC}^{\text{VPP}}}^{\text{LF}_t} = \frac{\begin{array}{l} [M^\gamma, h, j] \\ [E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, j, k, k \mid -, -] \\ [a, k, k + 1] \end{array}}{[M^\gamma; O^\gamma \rightarrow P^\gamma \bullet \nu, h, k, k + 1 \mid -, -]} \quad \begin{array}{l} \text{label}(P^\gamma) = a \\ \beta \in \text{adj}(M^\gamma) \end{array}$$

$$\mathcal{D}_{\text{pLC}^{\text{VPP}}}^{\text{LF}_\varepsilon} = \frac{\begin{array}{l} [M^\gamma, h, j] \\ [E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, j, k, k \mid -, -] \end{array}}{[M^\gamma; O^\gamma \rightarrow P^\gamma \bullet \nu, h, k, k \mid -, -]} \quad \begin{array}{l} \text{label}(P^\gamma) = \varepsilon \\ \beta \in \text{adj}(M^\gamma) \end{array}$$

$$\mathcal{D}_{\text{pLC}^{\text{VPP}}}^{\text{LF}_{\text{pre}}} = \frac{\begin{array}{l} [M^\gamma, h, j] \\ [E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, j, k, k \mid -, -] \end{array}}{[M^\gamma; O^\gamma \rightarrow \bullet P^\gamma \nu, h, k, k \mid -, -]} \quad \begin{array}{l} P^\gamma >_\ell \Delta \text{ o } \text{label}(P^\gamma) = \perp \\ \beta \in \text{adj}(M^\gamma) \end{array}$$

$$\mathcal{D}_{\text{pLC}^{\text{VPP}}}^{\text{FootComp}} = \frac{\begin{array}{l} [M^\gamma, h, j] \\ [E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, j, k, k \mid -, -] \\ [M^\gamma; M^\gamma \rightarrow \nu \bullet, h, k, l \mid p, q] \end{array}}{[E^\beta; \mathbf{F}^\beta \rightarrow \perp \bullet, j, k, l \mid k, l]} \quad \beta \in \text{adj}(M^\gamma)$$

El conjunto de ítems finales del esquema es:

$$\mathcal{F}_{\text{pLC}^{\text{VPP}}} = \{[\top; \top \rightarrow \mathbf{R}^\alpha \bullet, 0, 0, n \mid -, -] \mid \alpha \in \mathbf{I}\}$$

Todos los pasos funcionan de la misma forma que sus homólogos del esquema **pLC**. Al igual que ocurría en el esquema tipo Earley con VPP, el nuevo índice se limita a ser transmitido al consecuente, salvo en los pasos de predicción y completación de pie. En estos cuatro pasos se introduce un antecedente más, que se corresponde con el ítem predictivo que se lanzó para predecir la adjunción sobre un nodo, y que contiene el índice con la posición en la cadena de entrada donde comenzó el árbol al que pertenece dicho nodo. Este antecedente permite, en los tres pasos de predicción de pie, establecer sobre qué subárbol se debe continuar el reconocimiento y, en el paso de completación de pie, asociar el subárbol escindido a los nodos pie de los árboles auxiliares adecuados.

El algoritmo presenta una complejidad temporal de  $\mathcal{O}(n^8)$ , debida a que el paso  $\mathcal{D}_{\text{pLC}^{\text{VPP}}}^{\text{AdjComp}}$  involucra 8 índices. Sin embargo, mediante dos aplicaciones parciales sobre este paso, similares a las que describimos en el esquema **Nederhof**, la complejidad se reduce a  $\mathcal{O}(n^6)$ .

## 5.4 Esquema tipo Left Corner con ítems simplificados

Un ítem *left corner*  $[C^\gamma; N^\gamma \rightarrow \delta \bullet \nu, i, j \mid p, q]$  puede considerarse compuesto de dos partes claramente diferenciadas: la parte predicha  $[C^\gamma, i]$  y la parte reconocida  $[N^\gamma \rightarrow \delta \bullet \nu, i, j \mid p, q]$ . Si tenemos en cuenta que la información que conlleva la parte predicha ya está almacenada en el *chart* y que la parte reconocida son ítems Earley convencionales, podemos simplificar el esquema **pLC** para obtener un nuevo esquema **sLC**, cuya única función es la de servir de esquema intermedio para establecer la relación entre el esquema **pLC** de la sección anterior y el esquema **LC** que veremos en la siguiente sección.

Derivamos el esquema **sLC** a partir del esquema **pLC** de la siguiente manera:

- A los ítems *left corner* les eliminamos la parte predictiva (el ancla de la relación esquina izquierda) y lo convertimos en ítems Earley.
- Se modifican los pasos deductivos, añadiendo condiciones laterales en aquellos casos que sean necesarias.

En el dominio del esquema **sLC** vamos a distinguir dos tipos de ítems: predictivos y Earley. Por la propia naturaleza del analizador estos últimos se dividen en dos subtipos:

$$\mathcal{I}_{\text{sLC}} = \mathcal{I}_{\text{sLC}}^{\text{p}} \cup \mathcal{I}_{\text{sLC}}^{\text{e}} \cup \mathcal{I}_{\text{sLC}}^{\text{e}'}$$

Los ítems predictivos en este esquema son idénticos a los de la misma denominación en el esquema **pLC**:

$$\mathcal{I}_{\text{sLC}}^{\text{p}} = \mathcal{I}_{\text{pLC}}^{\text{p}}$$

Los ítems Earley tienen la misma forma que los ítems del analizador tipo Earley para las TAG pero con la particularidad de que se han filtrado los ítems con el punto al comienzo de la regla, salvo en los casos descritos en la definición de  $\mathcal{I}_{\text{pLC}}^{\text{lc}'}$  de la sección anterior. Por tanto, este tipo de ítems se define mediante:

$$\mathcal{I}_{\text{sLC}}^{\text{e}} = \{[M^\gamma \rightarrow \delta \bullet \nu, i, j \mid p, q]\}$$

donde  $M^\gamma \rightarrow \delta \nu \in \mathcal{P}(\gamma)$  y  $\gamma \in I \cup A$ . Los índices  $0 \leq i \leq j$  establecen las posiciones de la cadena de entrada que delimitan el fragmento reconocido por  $\delta$ . Si  $p$  y  $q$  presentan un valor definido, entonces  $M^\gamma$  pertenece a la espina de un árbol auxiliar,  $\delta$  domina el nodo pie de  $\gamma$  y se cumple  $(p, q) \leq (i, j)$ . Para evitar que el punto aparezca a la izquierda de la producción, se exige que  $\delta \neq \varepsilon$ .

$$\mathcal{I}_{\text{sLC}}^{\text{e}'} = \{[M^\gamma \rightarrow \bullet P^\gamma \nu, j, j \mid -, -]\}$$

donde  $M^\gamma \rightarrow P^\gamma \nu \in \mathcal{P}(\gamma)$  y  $\gamma \in I \cup A$ . El índice  $0 \leq j$  establece la posición de la cadena de entrada donde se inicia el reconocimiento de la regla. Y  $P^\gamma >_\ell \Delta$  o  $\text{label}(P^\gamma) = \perp$ .

El conjunto de pasos deductivos del esquema es:

$$\begin{aligned} \mathcal{D}_{\text{sLC}} = & \mathcal{D}_{\text{sLC}}^{\text{LI}_t} \cup \mathcal{D}_{\text{sLC}}^{\text{LI}_\varepsilon} \cup \mathcal{D}_{\text{sLC}}^{\text{LI}_{\text{pre}}} \cup \mathcal{D}_{\text{sLC}}^{\text{Scan}} \cup \mathcal{D}_{\text{sLC}}^\varepsilon \cup \mathcal{D}_{\text{sLC}}^{\text{LC}_t} \cup \mathcal{D}_{\text{sLC}}^{\text{LC}_\varepsilon} \cup \\ & \mathcal{D}_{\text{sLC}}^{\text{LC}_{\text{pre}}} \cup \mathcal{D}_{\text{sLC}}^{\text{LC}_n} \cup \mathcal{D}_{\text{sLC}}^{\text{Pre}} \cup \mathcal{D}_{\text{sLC}}^{\text{Comp}} \cup \mathcal{D}_{\text{sLC}}^{\text{LA}_t} \cup \mathcal{D}_{\text{sLC}}^{\text{LA}_\varepsilon} \cup \\ & \mathcal{D}_{\text{sLC}}^{\text{LA}_{\text{pre}}} \cup \mathcal{D}_{\text{sLC}}^{\text{AdjComp}} \cup \mathcal{D}_{\text{sLC}}^{\text{LF}_t} \cup \mathcal{D}_{\text{sLC}}^{\text{LF}_\varepsilon} \cup \mathcal{D}_{\text{sLC}}^{\text{LF}_{\text{pre}}} \cup \mathcal{D}_{\text{sLC}}^{\text{FootComp}} \end{aligned}$$

$$\mathcal{D}_{\text{sLC}}^{\text{LI}_t} = \frac{[a, 0, 1]}{[O^\alpha \rightarrow P^\alpha \bullet \nu, 0, 1 \mid -, -]} \quad \begin{array}{l} \alpha \in I \\ \top >_\ell^* O^\alpha \\ \text{label}(P^\alpha) = a \end{array}$$

$$\mathcal{D}_{\text{sLC}}^{\text{LI}_\varepsilon} = \frac{[a, 0, 0]}{[O^\alpha \rightarrow P^\alpha \bullet \nu, 0, 0 \mid -, -]} \quad \begin{array}{l} \alpha \in I \\ \top >_\ell^* O^\alpha \\ \text{label}(P^\alpha) = \varepsilon \end{array}$$

$$\mathcal{D}_{\text{sLC}}^{\text{LI}_{\text{pre}}} = \frac{[a, 0, 0]}{[O^\alpha \rightarrow \bullet P^\alpha \nu, 0, 0 \mid -, -]} \quad \begin{array}{l} \alpha \in I \\ \top >_\ell^* O^\alpha \\ P^\alpha >_\ell \Delta \end{array}$$

$$\mathcal{D}_{\text{sLC}}^{\text{Scan}} = \frac{\begin{array}{l} [N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \nu, i, j \mid p, q] \\ [a, j, j+1] \end{array}}{[N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, i, j+1 \mid p, q]} \quad \text{label}(M^\gamma) = a$$

$$\mathcal{D}_{\text{sLC}}^\varepsilon = \frac{[N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, i, j \mid p, q]} \quad \text{label}(M^\gamma) = \varepsilon$$

$$\mathcal{D}_{\text{sLC}}^{\text{LC}_t} = \frac{\begin{array}{l} [M^\gamma, j] \\ [a, j, j+1] \end{array}}{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, j+1 \mid -, -]} \quad \begin{array}{l} \text{nil} \in \text{adj}(M^\gamma) \\ M^\gamma >_\ell^* O^\alpha \\ \text{label}(P^\gamma) = a \end{array}$$

$$\mathcal{D}_{\text{sLC}}^{\text{LC}_\varepsilon} = \frac{[M^\gamma, j]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, j \mid -, -]} \quad \begin{array}{l} \text{nil} \in \text{adj}(M^\gamma) \\ M^\gamma >_\ell^* O^\alpha \\ \text{label}(P^\gamma) = \varepsilon \end{array}$$

$$\mathcal{D}_{\text{sLC}}^{\text{LC}_{\text{pre}}} = \frac{[M^\gamma, j]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j \mid -, -]} \quad \begin{array}{l} \text{nil} \in \text{adj}(M^\gamma) \\ M^\gamma >_\ell^* O^\alpha \\ P^\gamma >_\ell \Delta \circ \text{label}(P^\gamma) = \perp \end{array}$$

En este punto hay que hacer un comentario. En el paso homónimo del esquema **pLC** se disponía de información de cual era el ancla en una relación de esquina izquierda, pero esa información se suprime en este esquema. Este dato parece que se puede introducir en el paso como un antecedente adicional que informe del ancla de la relación LC, obteniendo el siguiente paso deductivo:

$$\mathcal{D}_{\text{sLC}}^{\text{LC}'_n} = \frac{\begin{array}{l} [N^\gamma \rightarrow \delta \bullet M^\gamma \alpha, i, j \mid p', q'] \\ [O^\gamma \rightarrow \nu \bullet, j, k \mid p, q] \end{array}}{[Q^\gamma \rightarrow O^\gamma \bullet \omega, j, k \mid p, q]} \quad \begin{array}{l} M^\gamma >_\ell^* O^\gamma \\ M^\gamma \neq O^\gamma \end{array}$$

Pero esta solución presenta problemas cuando el ancla de la relación LC no aparece en ningún ítem del *chart* con el punto delante, como ocurre cuando el ancla es la raíz de un árbol elemental o el nodo que domina el subárbol que pende del pie de un árbol auxiliar adjuntado. De cualquier manera, para ascender a través de los nodos de una relación de esquina izquierda no es necesario conocer el ancla, basta con saber el ascendente directo. Luego podemos corregir el paso deductivo  $D_{\text{sLC}}^{\text{LC}'_n}$  de la siguiente forma:

$$\begin{aligned}
 D_{\text{sLC}}^{\text{LC}'_n} &= \frac{[O^\gamma \rightarrow \nu \bullet, j, k \mid p, q]}{[Q^\gamma \rightarrow O^\gamma \bullet \omega, j, k \mid p, q]} \quad Q^\gamma >_\ell O^\gamma \\
 D_{\text{sLC}}^{\text{Pre}} &= \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[M^\gamma, j]} \quad \text{label}(M^\gamma) \in V_N \\
 D_{\text{sLC}}^{\text{Comp}} &= \frac{\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[M^\gamma \rightarrow \nu \bullet, j, k \mid p', q']}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k \mid p \cup p', q \cup q']} \quad \text{nil} \in \text{adj}(M^\gamma) \\
 D_{\text{sLC}}^{\text{LA}_t} &= \frac{\frac{[M^\gamma, j]}{[a, j, j+1]}}{[O^\beta \rightarrow P^\beta \bullet \nu, j, j+1 \mid -, -]} \quad \begin{array}{l} \top >_\ell^* O^\beta \\ \beta \in \text{adj}(M^\gamma) \\ \text{label}(P^\beta) = a \end{array} \\
 D_{\text{sLC}}^{\text{LA}_\varepsilon} &= \frac{[M^\gamma, j]}{[O^\beta \rightarrow P^\beta \bullet \nu, j, j \mid -, -]} \quad \begin{array}{l} \top >_\ell^* O^\beta \\ \beta \in \text{adj}(M^\gamma) \\ \text{label}(P^\beta) = \varepsilon \end{array} \\
 D_{\text{sLC}}^{\text{LA}_{\text{pre}}} &= \frac{[M^\gamma, j]}{[O^\beta \rightarrow \bullet P^\beta \nu, j, j \mid -, -]} \quad \begin{array}{l} \top >_\ell^* O^\beta \\ \beta \in \text{adj}(M^\gamma) \\ P^\beta >_\ell \Delta \circ \text{label}(P^\beta) = \perp \end{array} \\
 D_{\text{sLC}}^{\text{AdjComp}} &= \frac{\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p', q']}{\frac{[\top \rightarrow \mathbf{R}^\beta \bullet, j, m \mid k, l]}{[M^\gamma \rightarrow \omega \bullet, k, l \mid p, q]}}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, m \mid p \cup p', q \cup q']} \quad \beta \in \text{adj}(M^\gamma) \\
 D_{\text{sLC}}^{\text{LF}_t} &= \frac{\frac{[\mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -]}{[a, k, k+1]}}{[O^\gamma \rightarrow P^\gamma \bullet \nu, k, k+1 \mid -, -]} \quad \begin{array}{l} \beta \in \text{adj}(M^\gamma) \\ M^\gamma >_\ell^* O^\gamma \\ \text{label}(P^\gamma) = a \end{array} \\
 D_{\text{sLC}}^{\text{LF}_\varepsilon} &= \frac{[\mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, k, k \mid -, -]} \quad \begin{array}{l} \beta \in \text{adj}(M^\gamma) \\ M^\gamma >_\ell^* O^\gamma \\ \text{label}(P^\gamma) = \varepsilon \end{array} \\
 D_{\text{sLC}}^{\text{LF}_{\text{pre}}} &= \frac{[\mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, k, k \mid -, -]} \quad \begin{array}{l} \beta \in \text{adj}(M^\gamma) \\ M^\gamma >_\ell^* O^\gamma \\ P^\gamma >_\ell \Delta \circ \text{label}(P^\gamma) = \perp \end{array} \\
 D_{\text{sLC}}^{\text{FootComp}} &= \frac{\frac{[\mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -]}{[M^\gamma \rightarrow \nu \bullet, k, l \mid p, q]}}{[\mathbf{F}^\beta \rightarrow \perp \bullet, k, l \mid k, l]} \quad \beta \in \text{adj}(M^\gamma)
 \end{aligned}$$

El conjunto de ítems finales del esquema es igual al del esquema **E**:

$$\mathcal{F}_{\text{sLC}} = \mathcal{F}_{\text{E}}$$

## 5.5 Esquema tipo Left Corner

Como dijimos previamente, los ítems predictivos llevan información que está implícita en los ítems Earley que ya están almacenados en el *chart*. Por tanto, como evolución natural del esquema **sLC**, en esta sección vamos a definir el esquema **LC**, que elimina de su dominio este tipo de ítems, rehaciendo el conjunto de pasos deductivos para el nuevo dominio.

En el esquema **LC** vamos a distinguir dos subconjuntos de ítems:

$$\mathcal{I}_{\text{LC}} = \mathcal{I}_{\text{LC}}^e \cup \mathcal{I}_{\text{LC}}^{e'}$$

Los cuales son idénticos a los de la misma denominación del esquema **sLC**:

$$\begin{aligned} \mathcal{I}_{\text{LC}}^e &= \mathcal{I}_{\text{sLC}}^e \\ \mathcal{I}_{\text{LC}}^{e'} &= \mathcal{I}_{\text{sLC}}^{e'} \end{aligned}$$

El conjunto de pasos deductivos del esquema es:

$$\begin{aligned} \mathcal{D}_{\text{LC}} &= \mathcal{D}_{\text{LC}}^{\text{LI}_t} \cup \mathcal{D}_{\text{LC}}^{\text{LI}_\epsilon} \cup \mathcal{D}_{\text{LC}}^{\text{LI}_{\text{pre}}} \cup \mathcal{D}_{\text{LC}}^{\text{Scan}} \cup \mathcal{D}_{\text{LC}}^\epsilon \cup \mathcal{D}_{\text{LC}}^{\text{LC}_t} \cup \mathcal{D}_{\text{LC}}^{\text{LC}_\epsilon} \cup \\ &\quad \mathcal{D}_{\text{LC}}^{\text{LC}_{\text{pre}}} \cup \mathcal{D}_{\text{LC}}^{\text{LC}_n} \cup \mathcal{D}_{\text{LC}}^{\text{Comp}} \cup \mathcal{D}_{\text{LC}}^{\text{LA}_t} \cup \mathcal{D}_{\text{LC}}^{\text{LA}_\epsilon} \cup \\ &\quad \mathcal{D}_{\text{LC}}^{\text{LA}_{\text{pre}}} \cup \mathcal{D}_{\text{LC}}^{\text{AdjComp}} \cup \mathcal{D}_{\text{LC}}^{\text{LF}_t} \cup \mathcal{D}_{\text{LC}}^{\text{LF}_\epsilon} \cup \mathcal{D}_{\text{LC}}^{\text{LF}_{\text{pre}}} \cup \mathcal{D}_{\text{LC}}^{\text{FootComp}} \end{aligned}$$

Todos los pasos deductivos son iguales a los del esquema **sLC**, salvo los dedicados al filtrado de predicción de subárbol y adjunción:

$$\begin{aligned} \mathcal{D}_{\text{LC}}^{\text{LI}_t} &= \mathcal{D}_{\text{sLC}}^{\text{LI}_t} & \mathcal{D}_{\text{LC}}^{\text{LI}_\epsilon} &= \mathcal{D}_{\text{sLC}}^{\text{LI}_\epsilon} & \mathcal{D}_{\text{LC}}^{\text{LI}_{\text{pre}}} &= \mathcal{D}_{\text{sLC}}^{\text{LI}_{\text{pre}}} \\ \mathcal{D}_{\text{LC}}^{\text{Scan}} &= \mathcal{D}_{\text{sLC}}^{\text{Scan}} & \mathcal{D}_{\text{LC}}^\epsilon &= \mathcal{D}_{\text{sLC}}^\epsilon & \mathcal{D}_{\text{LC}}^{\text{LC}_n} &= \mathcal{D}_{\text{sLC}}^{\text{LC}_n} \\ \mathcal{D}_{\text{LC}}^{\text{Comp}} &= \mathcal{D}_{\text{sLC}}^{\text{Comp}} & \mathcal{D}_{\text{LC}}^{\text{AdjComp}} &= \mathcal{D}_{\text{sLC}}^{\text{AdjComp}} & \mathcal{D}_{\text{LC}}^{\text{LF}_t} &= \mathcal{D}_{\text{sLC}}^{\text{LF}_t} \\ \mathcal{D}_{\text{LC}}^{\text{LF}_\epsilon} &= \mathcal{D}_{\text{sLC}}^{\text{LF}_\epsilon} & \mathcal{D}_{\text{LC}}^{\text{LF}_{\text{pre}}} &= \mathcal{D}_{\text{sLC}}^{\text{LF}_{\text{pre}}} & \mathcal{D}_{\text{LC}}^{\text{FootComp}} &= \mathcal{D}_{\text{sLC}}^{\text{FootComp}} \end{aligned}$$

### Filtrado de predicción de subárbol

Al eliminar los ítems predictivos del dominio de esquema descrito en la sección anterior, nos vemos obligados a adaptar este grupo de pasos deductivos. Si tenemos en cuenta que en el esquema **sLC** los ítems predictivos sólo se generan mediante el paso deductivo  $\mathcal{D}_{\text{sLC}}^{\text{Pre}}$ ,

podemos sustituir este tipo de ítems por el antecedente de dicho paso y eliminar el mismo del conjunto de pasos deductivos.

$$\mathcal{D}_{\text{LC}}^{\text{LC}_t} = \frac{[N^\gamma \rightarrow \delta \bullet M^{\gamma\nu}, i, j \mid p, q]}{[a, j, j+1]} \quad \begin{array}{l} \text{nil} \in \text{adj}(M^\gamma) \\ M^\gamma >_\ell^* O^\alpha \\ \text{label}(P^\gamma) = a \end{array}$$

$$\mathcal{D}_{\text{LC}}^{\text{LC}_\varepsilon} = \frac{[N^\gamma \rightarrow \delta \bullet M^{\gamma\nu}, i, j \mid p, q]}{[O^\gamma \rightarrow P^\gamma \bullet \omega, j, j+1 \mid -, -]} \quad \begin{array}{l} \text{nil} \in \text{adj}(M^\gamma) \\ M^\gamma >_\ell^* O^\alpha \\ \text{label}(P^\gamma) = \varepsilon \end{array}$$

$$\mathcal{D}_{\text{LC}}^{\text{LC}_{\text{pre}}} = \frac{[N^\gamma \rightarrow \delta \bullet M^{\gamma\nu}, i, j \mid p, q]}{[O^\gamma \rightarrow \bullet P^\gamma \omega, j, j \mid -, -]} \quad \begin{array}{l} \text{nil} \in \text{adj}(M^\gamma) \\ M^\gamma >_\ell^* O^\alpha \\ P^\gamma >_\ell \Delta \vee \text{label}(P^\gamma) = \perp \end{array}$$

### Filtrado de predicción de adjunción

Para este grupo hacemos la misma adaptación que comentamos para los pasos deductivos de filtrado de predicción de subárbol.

$$\mathcal{D}_{\text{LC}}^{\text{LA}_t} = \frac{[N^\gamma \rightarrow \delta \bullet M^{\gamma\nu}, i, j \mid p, q]}{[a, j, j+1]} \quad \begin{array}{l} \top >_\ell^* O^\beta \\ \beta \in \text{adj}(M^\gamma) \\ \text{label}(P^\beta) = a \end{array}$$

$$\mathcal{D}_{\text{LC}}^{\text{LA}_\varepsilon} = \frac{[N^\gamma \rightarrow \delta \bullet M^{\gamma\nu}, i, j \mid p, q]}{[O^\beta \rightarrow P^\beta \bullet \omega, j, j \mid -, -]} \quad \begin{array}{l} \top >_\ell^* O^\beta \\ \beta \in \text{adj}(M^\gamma) \\ \text{label}(P^\beta) = \varepsilon \end{array}$$

$$\mathcal{D}_{\text{LC}}^{\text{LA}_{\text{pre}}} = \frac{[N^\gamma \rightarrow \delta \bullet M^{\gamma\nu}, i, j \mid p, q]}{[O^\beta \rightarrow \bullet P^\beta \omega, j, j \mid -, -]} \quad \begin{array}{l} \top >_\ell^* O^\beta \\ \beta \in \text{adj}(M^\gamma) \\ P^\beta >_\ell \Delta \vee \text{label}(P^\beta) = \perp \end{array}$$

El conjunto de ítems finales del esquema es igual al del esquema **sLC**:

$$\mathcal{F}_{\text{LC}} = \mathcal{F}_{\text{sLC}}$$

#### 5.5.1 Mejorando la eficiencia del esquema tipo Left Corner

En esta sección veremos una variante del esquema **LC** similar a la que presentamos en [43]. Este nuevo esquema, al que vamos a denominar **LC'**, busca filtrar aún más el número de predicciones del esquema anterior y mantiene sus cotas de complejidad espacial y temporal.

Tal como definimos la relación de esquina izquierda en las TAG, ésta se encuentra limitada a árboles elementales, de forma que en el esquema **LC** cuando el reconocimiento alcanza un nodo que domina por una relación **LC** al nodo pie (y, por consiguiente, al nodo *bottom*) de un árbol auxiliar, se baja el reconocimiento hasta el nodo *bottom*. Sin embargo, esta limitación en la definición de relación **LC** no es necesaria extenderla al analizador, es decir, el reconocimiento no tiene porqué bajar sólo hasta el nodo *bottom* de árbol auxiliar, sino

que puede seguir bajando a lo largo de la relación LC que mantiene el nodo sobre el que se está realizando la adjunción con sus descendientes. De esta forma se evita que el analizador almacene en el *chart* ítems que no aportan información adicional a la que ya existe de forma implícita en el mismo, y que son aquellos de la forma:

$$[F^\beta \rightarrow \bullet \perp, k, k \mid -, -]$$

No obstante, siempre no podremos obviar este tipo de ítems durante el proceso de reconocimiento, como veremos más adelante. Luego el dominio del esquema  $\mathbf{LC}'$  permanece igual al del esquema  $\mathbf{LC}$ :

$$\mathcal{I}_{\mathbf{LC}'} = \mathcal{I}_{\mathbf{LC}}^e \cup \mathcal{I}_{\mathbf{LC}'}^{e'}$$

donde  $\mathcal{I}_{\mathbf{LC}'}^e = \mathcal{I}_{\mathbf{LC}}^e$  y  $\mathcal{I}_{\mathbf{LC}'}^{e'} = \mathcal{I}_{\mathbf{LC}}^{e'}$ .

El conjunto de pasos deductivos del esquema es:

$$\begin{aligned} \mathcal{D}_{\mathbf{LC}'} = & \mathcal{D}_{\mathbf{LC}'}^{\text{LI}_t} \cup \mathcal{D}_{\mathbf{LC}'}^{\text{LI}_\varepsilon} \cup \mathcal{D}_{\mathbf{LC}'}^{\text{LI}_{\text{pre}}} \cup \mathcal{D}_{\mathbf{LC}'}^{\text{Scan}} \cup \mathcal{D}_{\mathbf{LC}'}^\varepsilon \cup \mathcal{D}_{\mathbf{LC}'}^{\text{LC}_t} \cup \mathcal{D}_{\mathbf{LC}'}^{\text{LC}_\varepsilon} \cup \\ & \mathcal{D}_{\mathbf{LC}'}^{\text{LC}_{\text{pre}}} \cup \mathcal{D}_{\mathbf{LC}'}^{\text{LC}_n} \cup \mathcal{D}_{\mathbf{LC}'}^{\text{Comp}} \cup \mathcal{D}_{\mathbf{LC}'}^{\text{LA}_t} \cup \mathcal{D}_{\mathbf{LC}'}^{\text{LA}_\varepsilon} \cup \\ & \mathcal{D}_{\mathbf{LC}'}^{\text{LA}_{\text{pre}}} \cup \mathcal{D}_{\mathbf{LC}'}^{\text{AdjComp}} \cup \mathcal{D}_{\mathbf{LC}'}^{\text{LF}_t} \cup \mathcal{D}_{\mathbf{LC}'}^{\text{LF}_\varepsilon} \cup \mathcal{D}_{\mathbf{LC}'}^{\text{LF}_{\text{pre}}} \cup \mathcal{D}_{\mathbf{LC}'}^{\text{FootComp}} \end{aligned}$$

Los pasos deductivos son básicamente iguales a los del esquema  $\mathbf{LC}$ , salvo en los pasos dedicados al filtrado de predicción de subárbol y pie, y el dedicado a la compleción del nodo pie:

$$\begin{aligned} \mathcal{D}_{\mathbf{LC}'}^{\text{LI}_t} = \mathcal{D}_{\mathbf{LC}}^{\text{LI}_t} \quad \mathcal{D}_{\mathbf{LC}'}^{\text{LI}_\varepsilon} = \mathcal{D}_{\mathbf{LC}}^{\text{LI}_\varepsilon} \quad \mathcal{D}_{\mathbf{LC}'}^{\text{LI}_{\text{pre}}} = \mathcal{D}_{\mathbf{LC}}^{\text{LI}_{\text{pre}}} \quad \mathcal{D}_{\mathbf{LC}'}^{\text{Scan}} = \mathcal{D}_{\mathbf{LC}}^{\text{Scan}} \\ \mathcal{D}_{\mathbf{LC}'}^\varepsilon = \mathcal{D}_{\mathbf{LC}}^\varepsilon \quad \mathcal{D}_{\mathbf{LC}'}^{\text{Comp}} = \mathcal{D}_{\mathbf{LC}}^{\text{Comp}} \quad \mathcal{D}_{\mathbf{LC}'}^{\text{AdjComp}} = \mathcal{D}_{\mathbf{LC}}^{\text{AdjComp}} \end{aligned}$$

### Filtrado de predicción de subárbol

Los pasos deductivos de predicción de subárbol donde la etiqueta del último nodo de la relación de LC es un símbolo terminal o  $\varepsilon$  son iguales a los del esquema  $\mathbf{LC}$ :

$$\mathcal{D}_{\mathbf{LC}'}^{\text{LC}_t} = \mathcal{D}_{\mathbf{LC}}^{\text{LC}_t}$$

$$\mathcal{D}_{\mathbf{LC}'}^{\text{LC}_\varepsilon} = \mathcal{D}_{\mathbf{LC}}^{\text{LC}_\varepsilon}$$

Pero cuando el último nodo de la relación de LC es un nodo de adjunción se aplica el paso siguiente:

$$\mathcal{D}_{\mathbf{LC}'}^{\text{LC}_{\text{pre}}} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q] \quad \begin{array}{l} \text{nil} \in \text{adj}(M^\gamma) \\ M^\gamma >_\ell^* O^\alpha \\ P^\gamma >_\ell \Delta \end{array}}{[O^\gamma \rightarrow \bullet P^\gamma \omega, j, j \mid -, -]}$$

Este paso difiere del  $\mathcal{D}_{\mathbf{LC}}^{\text{LC}_{\text{pre}}}$  en la condición lateral, en la cual hemos suprimido el caso en que el nodo  $P^\gamma$  estuviera etiquetado con  $\perp$ . De este caso ya se ocupan el conjunto de pasos de predicción de pie que veremos posteriormente.

### Filtrado de predicción de adjunción

Los pasos deductivos de predicción de adjunción son iguales a los del esquema LC:

$$\mathcal{D}_{LC'}^{LA_t} = \mathcal{D}_{LC}^{LA_t}$$

$$\mathcal{D}_{LC'}^{LA_\varepsilon} = \mathcal{D}_{LC}^{LA_\varepsilon}$$

$$\mathcal{D}_{LC'}^{LA_{pre}} = \mathcal{D}_{LC'}^{LA_{pre}}$$

Para el paso deductivo  $\mathcal{D}_{LC'}^{LA_{pre}}$  se mantiene la condición lateral en la que se detiene la relación LC cuando el último nodo sea un nodo adjuntable o esté etiquetado con  $\perp$ . Esto último es necesario porque en caso de no recoger esta situación, si en el árbol auxiliar que se va a adjuntar se cumple  $T >_\ell^* \perp$ , entonces se ignoraría en el proceso de reconocimiento.

### Filtrado de predicción de pie

Sea  $M^\gamma$  un nodo en un árbol elemental  $\gamma$  sobre el que se puede adjuntar un árbol auxiliar  $\beta$ . Supongamos que el reconocimiento ha alcanzado el nodo  $E^\beta$ , el cual que no presenta adjunción obligatoria y además domina el nodo pie por una relación LC. Estos nuevos pasos deductivos filtran tanto las predicciones sobre los nodos del árbol auxiliar  $\beta$  como sobre el árbol elemental  $\gamma$  donde se está llevando a cabo la adjunción:

$$\mathcal{D}_{LC'}^{LF_t} = \frac{[N^\beta \rightarrow \delta \bullet E^\beta \omega, j, k \mid -, -] \quad \beta \in \text{adj}(M^\gamma)}{[a, k, k + 1] \quad M^\gamma >_\ell^* O^\gamma} \quad \begin{array}{l} E^\beta >_\ell^* \perp \\ \text{label}(P^\gamma) = a \end{array}$$

$$\mathcal{D}_{LC'}^{LF_\varepsilon} = \frac{[N^\beta \rightarrow \delta \bullet E^\beta \omega, j, k \mid -, -] \quad \beta \in \text{adj}(M^\gamma)}{[O^\gamma \rightarrow P^\gamma \bullet \nu, k, k \mid -, -] \quad M^\gamma >_\ell^* O^\gamma} \quad \begin{array}{l} E^\beta >_\ell^* \perp \\ \text{label}(P^\gamma) = \varepsilon \end{array}$$

$$\mathcal{D}_{LC'}^{LF_{pre}} = \frac{[N^\beta \rightarrow \delta \bullet E^\beta \omega, j, k \mid -, -] \quad \beta \in \text{adj}(M^\gamma)}{[O^\gamma \rightarrow \bullet P^\gamma \nu, k, k \mid -, -] \quad M^\gamma >_\ell^* O^\gamma} \quad \begin{array}{l} E^\beta >_\ell^* \perp \\ P^\gamma >_\ell \Delta \text{ o } \text{label}(P^\gamma) = \perp \end{array}$$

Obsérvese que en el paso deductivo  $\mathcal{D}_{LC'}^{LF_{pre}}$ , al igual que ocurre en  $\mathcal{D}_{LC'}^{LA_{pre}}$ , también se mantiene la condición  $P^\gamma >_\ell \Delta$  o  $\text{label}(P^\gamma) = \perp$ , con objeto de poder realizar la completación del subárbol que domina el nodo pie si el árbol  $\gamma$  fuese un árbol auxiliar y además cumpliera que  $M^\gamma >_\ell^* \perp$ .

### Completación de pie



Frase	Reducción tiempo	Reducción ítems
Srini bought a book	-3%	-44%
Srini bought Beth a book	-5%	-47%
Srini bought a book at the bookstore	-6%	-46%
he put the book on the table	-8%	-44%
* he put the book	-13%	-42%
the sun melted the ice	-11%	-48%
the ice melted	-14%	-46%
Elmo borrowed a book	-7%	-45%
* a book borrowed	+5%	-41%
he hopes Muriel wins	-12%	-49%
he hopes that Muriel wins	-16%	-49%
the man who Muriel likes bought a book	+6%	-42%
the man that Muriel likes bought a book	-2%	-44%
the music should have been being played		
for the president	-28%	-56%
Clove caught a frisbee	-4%	-45%
who caught a frisbee	-7%	-45%
what did Clove catch	-13%	-49%
the aardvark smells terrible	-4%	-46%
the emu thinks that the aardvark smells terrible	-12%	-48%
who does the emu think smells terrible	-14%	-49%
who did the elephant think the panda		
heard the emu said smells terrible	-14%	-49%
Herbert is angry	-24%	-53%
Herbert is angry and furious	-21%	-54%
Herbert is more livid than angry	-25%	-51%
Herbert is more livid and furious than angry	-18%	-50%

Figura 5.6: Comparativa entre los esquemas tipo Earley y Left Corner

El paso deductivo de compleción de pie es similar al del esquema **LC**, pero hemos sustituido el antecedente  $[F^\beta \rightarrow \bullet \perp, k, l \mid k, l]$ , el cual ya no tenemos garantía que esté almacenado en el *chart*:

$$\mathcal{D}_{LC'}^{\text{FootComp}} = \frac{\begin{array}{l} [N^\beta \rightarrow \delta \bullet E^\beta \omega, j, k \mid -, -] \\ [M^\gamma \rightarrow \nu \bullet, k, l \mid p, q] \end{array}}{[F^\beta \rightarrow \perp \bullet, k, l \mid k, l]} \quad \begin{array}{l} \beta \in \text{adj}(M^\gamma) \\ E^\beta >_\ell^* \perp \end{array}$$

El conjunto de ítems finales del esquema es igual al del esquema **LC**:

$$\mathcal{F}_{LC'} = \mathcal{F}_{LC}$$

### 5.5.2 Resultados experimentales

Ninguno de estos analizadores basados en Left Corner reducen las cotas de complejidad teórica del analizador basado en Earley sin VPP que vimos en el capítulo anterior, per-

maneciendo en  $\mathcal{O}(n^6)$  la complejidad temporal en el caso peor en todos ellos. La mejora de las prestaciones de estos analizadores tipo Left Corner respecto al esquema **E** vienen, como en el caso del esquema **buLC**, por la reducción del número de ítems deducidos durante el proceso de análisis y la consiguiente disminución del espacio de búsqueda. Esta reducción siempre dependerá de las gramáticas y cadenas de entrada consideradas.

Hemos realizado estudios preliminares donde comparamos el comportamiento de un analizador con filtro left corner, concretamente el esquema **LC'**, frente al esquema **E** basado en Earley. Incorporamos ambos analizadores a una implementación en Prolog de la máquina deductiva de análisis presentada en [122]. Como gramática hemos tomado el subconjunto de la gramática del inglés de amplia cobertura [144] descrita en [35], que consta de 27 árboles elementales que cubren una amplia variedad de construcciones: cláusulas de relativo, verbos auxiliares, dependencias lejanas, extracción, etc. Con objeto de obviar el tiempo empleado en el proceso de unificación, se eliminaron las estructuras de rasgos asociadas a los nodos de los árboles elementales, simulándolas mediante restricciones locales. Cada una de las sentencias se analizó sin ningún tipo de filtro previo sobre el conjunto de árboles elementales. En la figura 5.6 se muestra la reducción porcentual, en tiempo y número de ítems deducidos, que presenta el esquema tipo Left Corner frente al esquema tipo Earley. Con el banco de pruebas expuesto, podemos decir que presenta una reducción media del 11% en tiempo y del 50% en el número de ítems deducidos.

## Capítulo 6

# Ampliando la red de analizadores para TAG

Este capítulo está dedicado a establecer las relaciones formales entre los esquemas basados en Left Corner para TAG propuestos en el capítulo anterior. También se incluyen las relaciones de éstos con los esquemas tipo Earley sin VPP, tanto en su versión ascendente (**buE**) como predictiva (**E**), con objeto de integrar estos nuevos analizadores en la red de esquemas de análisis para TAG basados en el algoritmo de Earley iniciada en [6] y [35].

### 6.1 Relaciones entre esquemas tipo Left Corner para TAG

Veamos a continuación las relaciones formales que existen entre los distintos esquemas que hemos propuesto en el capítulo anterior.

**Teorema 6.1** *Relación entre los esquemas pLC, sLC y LC*

*Se mantienen las siguientes relaciones de refinamiento de pasos e ítems:*

$$\text{LC} \xrightarrow{\text{sr}} \text{sLC} \xrightarrow{\text{ir}} \text{pLC}$$

*Prueba*

Primero vamos a probar la relación  $\text{sLC} \xrightarrow{\text{ir}} \text{pLC}$ . Debemos probar que existe una función regular  $f : \mathcal{I}_{\text{pLC}} \rightarrow \mathcal{I}_{\text{sLC}}$  tal que:

1.  $\mathcal{I}_{\text{sLC}} = f(\mathcal{I}_{\text{pLC}})$
2.  $\Delta_{\text{sLC}} = f(\Delta_{\text{pLC}})$

Una función regular de contracción de ítems que cumple estas condiciones es la siguiente:

$$f([C^\gamma; N^\gamma \rightarrow \delta \bullet \nu, i, j \mid p, q]) = [N^\gamma \rightarrow \delta \bullet \nu, i, j \mid p, q]$$

Puesto que los ítems predictivos son iguales en ambos esquemas:

$$f([C^\gamma, j]) = [C^\gamma, j]$$

De  $f$  se sigue inmediatamente que  $\mathcal{I}_{sLC} = f(\mathcal{I}_{pLC})$  y  $\Delta_{sLC} = f(\Delta_{pLC})$  por inducción en la longitud de las secuencias de derivación.

A continuación probaremos la relación  $LC \xrightarrow{sr} sLC$ . Para ello tenemos que demostrar que:

$$1. \mathcal{I}_{LC} \subseteq \mathcal{I}_{sLC}$$

$$2. \vdash_{LC}^* \subseteq \vdash_{sLC}^*$$

Lo primero es cierto por definición, ya que el dominio del esquema  $sLC$  está formado por los conjuntos de ítems del esquema  $LC$  y los predictivos:

$$\mathcal{I}_{sLC} = \mathcal{I}_{sLC}^P \cup \mathcal{I}_{LC}$$

Para demostrar que  $\vdash_{LC}^* \subseteq \vdash_{sLC}^*$  nos basta con probar que  $\mathcal{D}_{LC} \subseteq \vdash_{sLC}^*$ . Veamos cada paso:

- Un paso deductivo  $\mathcal{D}_{LC}^{LC_t}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{sLC}^{Pre}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[M^\gamma, j]}$$

seguido de otro  $\mathcal{D}_{sLC}^{LC_t}$

$$\frac{\frac{[M^\gamma, j]}{[a, j, j + 1]}}{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, j + 1 \mid -, -]}$$

- Un paso deductivo  $\mathcal{D}_{LC}^{LC_\epsilon}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{sLC}^{Pre}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[M^\gamma, j]}$$

seguido de otro  $\mathcal{D}_{sLC}^{LC_\epsilon}$

$$\frac{[M^\gamma, j]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, j \mid -, -]}$$

- Un paso deductivo  $\mathcal{D}_{LC}^{LC_{pre}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{sLC}^{Pre}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[M^\gamma, j]}$$

seguido de otro  $\mathcal{D}_{sLC}^{LC_{pre}}$

$$\frac{[M^\gamma, j]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j \mid -, -]}$$

- Un paso deductivo  $\mathcal{D}_{\text{LC}}^{\text{LA}_t}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{sLC}}^{\text{Pre}}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[M^\gamma, j]}$$

seguido de otro  $\mathcal{D}_{\text{sLC}}^{\text{LA}_t}$

$$\frac{\frac{[M^\gamma, j]}{[a, j, j+1]}}{[O^\beta \rightarrow P^\beta \bullet \nu, j, j+1 \mid -, -]}$$

- Un paso deductivo  $\mathcal{D}_{\text{LC}}^{\text{LA}_e}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{sLC}}^{\text{Pre}}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[M^\gamma, j]}$$

seguido de otro  $\mathcal{D}_{\text{sLC}}^{\text{LA}_e}$

$$\frac{[M^\gamma, j]}{[O^\beta \rightarrow P^\beta \bullet \nu, j, j \mid -, -]}$$

- Un paso deductivo  $\mathcal{D}_{\text{LC}}^{\text{LA}_{\text{pre}}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{sLC}}^{\text{Pre}}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[M^\gamma, j]}$$

seguido de otro  $\mathcal{D}_{\text{sLC}}^{\text{LA}_{\text{pre}}}$

$$\frac{[M^\gamma, j]}{[O^\beta \rightarrow \bullet P^\beta \nu, j, j \mid -, -]}$$

- El resto de pasos deductivos es igual en ambos esquemas

### Teorema 6.2 Relación entre los esquemas buLC y LC

Se mantiene la siguientes relaciones de filtros dinámicos y refinamiento de pasos:

$$\text{LC} \stackrel{\text{df}}{\Longrightarrow} \text{LC}_1 \stackrel{\text{df}}{\longleftarrow} \text{buLC}_1 \stackrel{\text{sr}}{\longleftarrow} \text{buLC}$$

*Prueba*

Antes de iniciar la demostración vamos a definir dos nuevos esquemas que nos servirán de enlaces para establecer la relación entre los dos esquemas de interés.

El primero, que denominaremos  $\text{LC}_1$ , se trata de una transformación trivial sobre el esquema  $\text{LC}$  que consiste en el desdoble de los pasos  $\mathcal{D}_{\text{LC}}^{\text{Comp}}$  y  $\mathcal{D}_{\text{LC}}^{\text{AdjComp}}$  en cuatro:  $\mathcal{D}_{\text{LC}_1}^{\text{Comp}_1}$ ,  $\mathcal{D}_{\text{LC}_1}^{\text{Comp}_2}$ ,  $\mathcal{D}_{\text{LC}_1}^{\text{AdjComp}_1}$  y  $\mathcal{D}_{\text{LC}_1}^{\text{AdjComp}_2}$ . De forma que los que presentan subíndice 1 se encargan de completar los subárboles y las adjunciones en nodos que no son hijos izquierdos y los que tienen subíndice 2 las completan en los hijos izquierdos.

El dominio de  $\text{LC}_1$  viene dado por:

$$\mathcal{I}_{\text{LC}_1} = \mathcal{I}_{\text{LC}}$$

El conjunto de pasos deductivos del nuevo esquema es:

$$\begin{aligned} \mathcal{D}_{LC_1} &= \mathcal{D}_{LC_1}^{LI_t} \cup \mathcal{D}_{LC_1}^{LI_\epsilon} \cup \mathcal{D}_{LC_1}^{LI_{pre}} \cup \mathcal{D}_{LC_1}^{Scan} \cup \mathcal{D}_{LC_1}^\epsilon \cup \mathcal{D}_{LC_1}^{LC_t} \cup \mathcal{D}_{LC_1}^{LC_\epsilon} \cup \\ &\mathcal{D}_{LC_1}^{LC_{pre}} \cup \mathcal{D}_{LC_1}^{LC_n} \cup \mathcal{D}_{LC_1}^{Comp_1} \cup \mathcal{D}_{LC_1}^{Comp_2} \cup \mathcal{D}_{LC_1}^{LA_t} \cup \mathcal{D}_{LC_1}^{LA_\epsilon} \cup \mathcal{D}_{LC_1}^{LA_{pre}} \cup \\ &\mathcal{D}_{LC_1}^{AdjComp_1} \cup \mathcal{D}_{LC_1}^{AdjComp_2} \cup \mathcal{D}_{LC_1}^{LF_t} \cup \mathcal{D}_{LC_1}^{LF_\epsilon} \cup \mathcal{D}_{LC_1}^{LF_{pre}} \cup \mathcal{D}_{LC_1}^{FootComp} \end{aligned}$$

donde todos los pasos deductivos son iguales a sus homónimos del esquema **LC**, salvo los cuatro nombrados anteriormente.

$$\begin{aligned} \mathcal{D}_{LC_1}^{LI_t} &= \mathcal{D}_{LC}^{LI_t} & \mathcal{D}_{LC_1}^{LI_\epsilon} &= \mathcal{D}_{LC}^{LI_\epsilon} & \mathcal{D}_{LC_1}^{LI_{pre}} &= \mathcal{D}_{LC}^{LI_{pre}} & \mathcal{D}_{LC_1}^{Scan} &= \mathcal{D}_{LC}^{Scan} \\ \mathcal{D}_{LC_1}^\epsilon &= \mathcal{D}_{LC}^\epsilon & \mathcal{D}_{LC_1}^{LC_t} &= \mathcal{D}_{LC}^{LC_t} & \mathcal{D}_{LC_1}^{LC_\epsilon} &= \mathcal{D}_{LC}^{LC_\epsilon} \\ \mathcal{D}_{LC_1}^{LC_{pre}} &= \mathcal{D}_{LC}^{LC_{pre}} & \mathcal{D}_{LC_1}^{LC_n} &= \mathcal{D}_{LC}^{LC_n} & \mathcal{D}_{LC_1}^{LA_t} &= \mathcal{D}_{LC}^{LA_t} \\ \mathcal{D}_{LC_1}^{LA_\epsilon} &= \mathcal{D}_{LC}^{LA_\epsilon} & \mathcal{D}_{LC_1}^{LA_{pre}} &= \mathcal{D}_{LC}^{LA_{pre}} & \mathcal{D}_{LC_1}^{LF_t} &= \mathcal{D}_{LC}^{LF_t} \\ \mathcal{D}_{LC_1}^{LF_\epsilon} &= \mathcal{D}_{LC}^{LF_\epsilon} & \mathcal{D}_{LC_1}^{LF_{pre}} &= \mathcal{D}_{LC}^{LF_{pre}} & \mathcal{D}_{LC_1}^{FootComp} &= \mathcal{D}_{LC}^{FootComp} \\ \mathcal{D}_{LC_1}^{Comp_1} &= \frac{\begin{array}{l} [N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \nu, i, j \mid p, q] \\ [M^\gamma \rightarrow \nu \bullet, j, k \mid p', q'] \end{array}}{[N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, i, k \mid p \cup p', q \cup q']} \quad \text{nil} \in \text{adj}(M^\gamma) \\ \mathcal{D}_{LC_1}^{Comp_2} &= \frac{\begin{array}{l} [N^\gamma \rightarrow \bullet M^\gamma \nu, j, j \mid -, -] \\ [M^\gamma \rightarrow \nu \bullet, j, k \mid p, q] \end{array}}{[N^\gamma \rightarrow M^\gamma \bullet \nu, j, k \mid p, q]} \quad \text{nil} \in \text{adj}(M^\gamma) \\ \mathcal{D}_{LC_1}^{AdjComp_1} &= \frac{\begin{array}{l} [N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \nu, i, j \mid p', q'] \\ [\top \rightarrow \mathbf{R}^\beta \bullet, j, m \mid k, l] \\ [M^\gamma \rightarrow \omega \bullet, k, l \mid p, q] \end{array}}{[N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, i, m \mid p \cup p', q \cup q']} \quad \beta \in \text{adj}(M^\gamma) \\ \mathcal{D}_{LC_1}^{AdjComp_2} &= \frac{\begin{array}{l} [N^\gamma \rightarrow \bullet M^\gamma \nu, j, j \mid -, -] \\ [\top \rightarrow \mathbf{R}^\beta \bullet, j, m \mid k, l] \\ [M^\gamma \rightarrow \omega \bullet, k, l \mid p, q] \end{array}}{[N^\gamma \rightarrow M^\gamma \bullet \nu, j, m \mid p, q]} \quad \beta \in \text{adj}(M^\gamma) \end{aligned}$$

El conjunto de ítems finales también es igual al del esquema **LC**:

$$\mathcal{I}_{LC_1} = \mathcal{I}_{LC}$$

El segundo esquema intermedio, al que vamos a denominar **buLC<sub>1</sub>**, consiste en la ampliación del dominio de **buLC** con ítems con el punto al comienzo de la regla cuando el hijo izquierdo sea un nodo adjutable. El dominio viene dado por:

$$\mathcal{I}_{buLC_1} = \mathcal{I}_{buLC_1}^e \cup \mathcal{I}_{buLC_1}^{e'}$$

$$\mathcal{I}_{buLC_1}^e = \{[M^\gamma \rightarrow \delta \bullet \nu, i, j \mid p, q]\}$$

tal que  $M^\gamma \rightarrow \delta\nu \in \mathcal{P}(\gamma)$ ,  $\gamma \in \mathbf{I} \cup \mathbf{A}$ ,  $0 \leq i \leq j$ ,  $\delta \neq \varepsilon$ ,  $((p, q) \leq (i, j)$  o  $(p, q) = (-, -)$ ).

$$\mathcal{I}_{\text{buLC}_1}^{\varepsilon'} = \{[M^\gamma \rightarrow \bullet P^\gamma \nu, j, j \mid -, -]\}$$

tal que  $M^\gamma \rightarrow P^\gamma \nu \in \mathcal{P}(\gamma)$ ,  $\gamma \in \mathbf{I} \cup \mathbf{A}$ ,  $0 \leq j$ ,  $(P^\gamma >_\ell \Delta$  o  $\text{label}(P^\gamma) = \perp$ ).

Por tanto, se cumple que:

$$\mathcal{I}_{\text{buLC}_1} = \mathcal{I}_{\text{LC}} = \mathcal{I}_{\text{LC}_1}$$

Hay que adaptar el conjunto de pasos deductivos al nuevo dominio:

$$\begin{aligned} \mathcal{D}_{\text{buLC}_1} &= \mathcal{D}_{\text{buLC}_1}^{\text{LC}_t} \cup \mathcal{D}_{\text{buLC}_1}^{\text{LC}_\varepsilon} \cup \mathcal{D}_{\text{buLC}_1}^{\text{LC}_n} \cup \mathcal{D}_{\text{buLC}_1}^{\text{LC}_{\text{pro}}} \cup \mathcal{D}_{\text{buLC}_1}^{\text{LC}_{\text{cad}}} \\ &\cup \mathcal{D}_{\text{buLC}_1}^{\text{Foot}} \cup \mathcal{D}_{\text{buLC}_1}^{\text{Scan}} \cup \mathcal{D}_{\text{buLC}_1}^\varepsilon \cup \mathcal{D}_{\text{buLC}_1}^{\text{Comp}} \cup \mathcal{D}_{\text{buLC}_1}^{\text{AdjComp}} \end{aligned}$$

donde todos los pasos son idénticos al los homónimos de **buLC**, excepto el paso  $\mathcal{D}_{\text{buLC}}^{\text{LC}_{\text{cad}}}$  y el nuevo paso  $\mathcal{D}_{\text{buLC}_1}^{\text{LC}_{\text{pre}}}$ .

$$\begin{aligned} \mathcal{D}_{\text{buLC}_1}^{\text{LC}_t} &= \mathcal{D}_{\text{buLC}}^{\text{LC}_t} & \mathcal{D}_{\text{buLC}_1}^{\text{LC}_\varepsilon} &= \mathcal{D}_{\text{buLC}}^{\text{LC}_\varepsilon} & \mathcal{D}_{\text{buLC}_1}^{\text{LC}_n} &= \mathcal{D}_{\text{buLC}}^{\text{LC}_n} \\ \mathcal{D}_{\text{buLC}_1}^{\text{Foot}} &= \mathcal{D}_{\text{buLC}}^{\text{Foot}} & \mathcal{D}_{\text{buLC}_1}^{\text{Scan}} &= \mathcal{D}_{\text{buLC}}^{\text{Scan}} & \mathcal{D}_{\text{buLC}_1}^\varepsilon &= \mathcal{D}_{\text{buLC}}^\varepsilon \\ \mathcal{D}_{\text{buLC}_1}^{\text{Comp}} &= \mathcal{D}_{\text{buLC}}^{\text{Comp}} & \mathcal{D}_{\text{buLC}_1}^{\text{AdjComp}} &= \mathcal{D}_{\text{buLC}}^{\text{AdjComp}} \\ \mathcal{D}_{\text{buLC}_1}^{\text{LC}_{\text{pre}}} &= \frac{[Q^\gamma \rightarrow \bullet O^\gamma \omega, j, j, -, -]}{[Q^\gamma \rightarrow \bullet O^\gamma \omega, j, j, -, -]} & O^\gamma >_\ell \Delta \\ & & [Q^\gamma \rightarrow \bullet O^\gamma \omega, j, j, -, -] \\ & & [\top \rightarrow \mathbf{R}^\beta \bullet, j, m, k, l] \\ \mathcal{D}_{\text{buLC}_1}^{\text{LC}_{\text{cad}}} &= \frac{[O^\gamma \rightarrow \nu \bullet, k, l, p, q]}{[Q^\gamma \rightarrow O^\gamma \bullet \omega, j, m, p, q]} & \beta \in \text{adj}(O^\gamma) \end{aligned}$$

El conjunto de ítems finales es igual al del esquema **buLC**:

$$\mathcal{F}_{\text{buLC}_1} = \mathcal{F}_{\text{buLC}}$$

Para probar que  $\text{LC} \xrightarrow{\text{df}} \text{LC}_1$  tenemos que demostrar que:

1.  $\mathcal{I}_{\text{LC}_1} \subseteq \mathcal{I}_{\text{LC}}$
2.  $\vdash_{\text{LC}_1} \subseteq \vdash_{\text{LC}}$

Lo primero es cierto por definición, ya que los dominios de ambos esquemas son iguales:

$$\mathcal{I}_{\text{LC}_1} = \mathcal{I}_{\text{LC}}$$

Para probar que  $\vdash_{\text{LC}_1} \subseteq \vdash_{\text{LC}}$  debemos demostrar que  $\mathcal{D}_{\text{LC}_1} \subseteq \vdash_{\text{LC}}$ . Sólo vamos a considerar los cuatro pasos sobre los que se aplica el filtro dinámico, el resto son idénticos:

- Dado un paso

$$\frac{\begin{array}{l} [N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \nu, i, j \mid p', q'] \\ [\top \rightarrow \mathbf{R}^\beta \bullet, j, m \mid k, l] \\ [M^\gamma \rightarrow \omega \bullet, k, l \mid p, q] \end{array}}{[N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, i, m \mid p \cup p', q \cup q']} \in \mathcal{D}_{LC_1}^{\text{AdjComp}_1}$$

existe un paso

$$\frac{\begin{array}{l} [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p', q'] \\ [\top \rightarrow \mathbf{R}^\beta \bullet, j, m \mid k, l] \\ [M^\gamma \rightarrow \omega \bullet, k, l \mid p, q] \end{array}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, m \mid p \cup p', q \cup q']} \in \mathcal{D}_{LC}^{\text{AdjComp}}$$

y, por tanto, existe la inferencia

$$\frac{\begin{array}{l} [N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \nu, i, j \mid p', q'] \\ [\top \rightarrow \mathbf{R}^\beta \bullet, j, m \mid k, l] \\ [M^\gamma \rightarrow \omega \bullet, k, l \mid p, q] \end{array}}{\vdash_{LC} [N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, i, m \mid p \cup p', q \cup q']}$$

- Dado un paso

$$\frac{\begin{array}{l} [N^\gamma \rightarrow \bullet M^\gamma \nu, j, j \mid -, -] \\ [\top \rightarrow \mathbf{R}^\beta \bullet, j, m \mid k, l] \\ [M^\gamma \rightarrow \omega \bullet, k, l \mid p, q] \end{array}}{[N^\gamma \rightarrow M^\gamma \bullet \nu, j, m \mid p, q]} \in \mathcal{D}_{LC_1}^{\text{AdjComp}_2}$$

existe un paso

$$\frac{\begin{array}{l} [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p', q'] \\ [\top \rightarrow \mathbf{R}^\beta \bullet, j, m \mid k, l] \\ [M^\gamma \rightarrow \omega \bullet, k, l \mid p, q] \end{array}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, m \mid p \cup p', q \cup q']} \in \mathcal{D}_{LC}^{\text{AdjComp}}$$

y, por tanto, existe la inferencia

$$\frac{\begin{array}{l} [N^\gamma \rightarrow \bullet M^\gamma \nu, j, j \mid -, -] \\ [\top \rightarrow \mathbf{R}^\beta \bullet, j, m \mid k, l] \\ [M^\gamma \rightarrow \omega \bullet, k, l \mid p, q] \end{array}}{\vdash_{LC} [N^\gamma \rightarrow M^\gamma \bullet \nu, j, m \mid p, q]}$$

- Dado un paso

$$\frac{\begin{array}{l} [N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \nu, i, j \mid p, q] \\ [M^\gamma \rightarrow \nu \bullet, j, k \mid p', q'] \end{array}}{[N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, i, k \mid p \cup p', q \cup q']} \in \mathcal{D}_{LC_1}^{\text{Comp}_1}$$

existe un paso

$$\frac{\begin{array}{l} [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q] \\ [M^\gamma \rightarrow \nu \bullet, j, k \mid p', q'] \end{array}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k \mid p \cup p', q \cup q']} \in \mathcal{D}_{LC}^{\text{Comp}}$$

y, por tanto, existe la inferencia

$$\frac{\begin{array}{l} [N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \nu, i, j \mid p, q] \\ [M^\gamma \rightarrow \nu \bullet, j, k \mid p', q'] \end{array}}{\vdash_{LC} [N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, i, k \mid p \cup p', q \cup q']}$$

- Dado un paso

$$\frac{\begin{array}{l} [N^\gamma \rightarrow \bullet M^\gamma \nu, j, j \mid -, -] \\ [M^\gamma \rightarrow \nu \bullet, j, k \mid p, q] \end{array}}{[N^\gamma \rightarrow M^\gamma \bullet \nu, j, k \mid p, q]} \in \mathcal{D}_{LC_1}^{\text{Comp}_2}$$

existe un paso

$$\frac{\begin{array}{l} [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q] \\ [M^\gamma \rightarrow \nu \bullet, j, k \mid p', q'] \end{array}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k \mid p \cup p', q \cup q']} \in \mathcal{D}_{LC}^{\text{Comp}}$$

y, por tanto, existe la inferencia

$$\frac{[N^\gamma \rightarrow \bullet M^\gamma \nu, j, j \mid -, -]}{[M^\gamma \rightarrow \nu \bullet, j, k \mid p, q]} \vdash_{LC} [N^\gamma \rightarrow M^\gamma \bullet \nu, j, k \mid p, q]$$

Probemos ahora que  $\text{buLC} \xrightarrow{\text{sr}} \text{buLC}_1$ , para lo cual se tiene que cumplir que:

1.  $\mathcal{I}_{\text{buLC}} \subseteq \mathcal{I}_{\text{buLC}_1}$
2.  $\vdash_{\text{buLC}}^* \subseteq \vdash_{\text{buLC}_1}^*$

Lo primero es cierto, ya que el dominio del esquema  $\text{buLC}_1$  está formado por los conjuntos de ítems del esquema  $\text{buLC}$  más los ítems con el punto a comienzo de la regla. Por tanto:

$$\mathcal{I}_{\text{buLC}} \subset \mathcal{I}_{\text{buLC}_1}$$

Para demostrar que  $\vdash_{\text{buLC}}^* \subseteq \vdash_{\text{buLC}_1}^*$  nos basta con probar que  $\mathcal{D}_{\text{buLC}} \subseteq \mathcal{D}_{\text{buLC}_1}^*$ . Veamos únicamente el paso  $\mathcal{D}_{\text{buLC}}^{\text{LCcad}}$ , puesto que el resto son iguales a sus homónimos en el esquema  $\text{buLC}_1$ :

- Un paso deductivo  $\mathcal{D}_{\text{buLC}}^{\text{LCcad}}$  es equivalente a una secuencia formada por una paso  $\mathcal{D}_{\text{buLC}_1}^{\text{LCpre}}$

$$\overline{[Q^\gamma \rightarrow \bullet O^\gamma \omega, j, j, -, -]}$$

seguido de otro  $\mathcal{D}_{\text{buLC}_1}^{\text{LCcad}}$

$$\frac{\begin{array}{l} [Q^\gamma \rightarrow \bullet O^\gamma \omega, j, j, -, -] \\ [\top \rightarrow \mathbf{R}^\beta \bullet, j, m, k, l] \\ [O^\gamma \rightarrow \nu \bullet, k, l, p, q] \end{array}}{[Q^\gamma \rightarrow O^\gamma \bullet \omega, j, m, p, q]}$$

Nos queda probar que  $\text{buLC}_1 \xrightarrow{\text{df}} \text{LC}_1$ , para lo cual tenemos que demostrar que:

1.  $\mathcal{I}_{\text{LC}_1} \subseteq \mathcal{I}_{\text{buLC}_1}$
2.  $\vdash_{\text{LC}_1} \subseteq \vdash_{\text{buLC}_1}$



Lo primero es cierto por definición, ya que los dominios de ambos esquemas son iguales:

$$\mathcal{I}_{LC_1} = \mathcal{I}_{buLC_1}$$

Para probar que  $\vdash_{LC_1} \subseteq \vdash_{buLC_1}$  debemos demostrar que  $\mathcal{D}_{LC_1} \subseteq \mathcal{D}_{buLC_1}$ . Vamos a ver cada paso:

- Dado un paso

$$\frac{[a, 0, 1]}{[O^\alpha \rightarrow P^\alpha \bullet \nu, 0, 1 \mid -, -]} \in \mathcal{D}_{LC_1}^{LI_t}$$

existe un paso

$$\frac{[a, j, j+1]}{[O^\alpha \rightarrow P^\alpha \bullet \nu, j, j+1 \mid -, -]} \in \mathcal{D}_{buLC_1}^{LC_t}$$

y, por tanto, existe la inferencia

$$[a, 0, 1] \vdash_{buLC_1} [O^\alpha \rightarrow P^\alpha \bullet \nu, 0, 1 \mid -, -]$$

- Dado un paso

$$\overline{[O^\alpha \rightarrow P^\alpha \bullet \nu, 0, 0 \mid -, -]} \in \mathcal{D}_{LC_1}^{LI_e}$$

existe un paso

$$\overline{[O^\alpha \rightarrow P^\alpha \bullet \nu, j, j \mid -, -]} \in \mathcal{D}_{buLC_1}^{LC_e}$$

y, por tanto, existe la inferencia

$$\vdash_{buLC_1} [O^\alpha \rightarrow P^\alpha \bullet \nu, 0, 0 \mid -, -]$$

- Dado un paso

$$\overline{[O^\alpha \rightarrow \bullet P^\alpha \nu, 0, 0 \mid -, -]} \in \mathcal{D}_{LC_1}^{LI_{pre}}$$

existe un paso

$$\overline{[O^\alpha \rightarrow \bullet P^\alpha \nu, j, j \mid -, -]} \in \mathcal{D}_{buLC_1}^{LC_{pre}}$$

y, por tanto, existe la inferencia

$$\vdash_{buLC_1} [O^\alpha \rightarrow \bullet P^\alpha \nu, 0, 0 \mid -, -]$$

- Dado un paso

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q] \quad [a, j, j+1]}{[O^\gamma \rightarrow P^\gamma \bullet \omega, j, j+1 \mid -, -]} \in \mathcal{D}_{LC_1}^{LC_t}$$

existe un paso

$$\frac{[a, j, j+1]}{[O^\gamma \rightarrow P^\gamma \bullet \omega, j, j+1 \mid -, -]} \in \mathcal{D}_{buLC_1}^{LC_t}$$

y, por tanto, existe la inferencia

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q] \quad [a, j, j+1]}{[O^\gamma \rightarrow P^\gamma \bullet \omega, j, j+1 \mid -, -]} \vdash_{buLC_1} [O^\gamma \rightarrow P^\gamma \bullet \omega, j, j+1 \mid -, -]$$

- Dado un paso

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[O^\gamma \rightarrow P^\gamma \bullet \omega, j, j \mid -, -]} \in \mathcal{D}_{LC_1}^{LC_\varepsilon}$$

existe un paso

$$\overline{[O^\gamma \rightarrow P^\gamma \bullet \omega, j, j \mid -, -]} \in \mathcal{D}_{buLC_1}^{LC_\varepsilon}$$

y, por tanto, existe la inferencia

$$[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q] \vdash_{buLC_1} [O^\gamma \rightarrow P^\gamma \bullet \omega, j, j \mid -, -]$$

- Dado un paso

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[O^\gamma \rightarrow \bullet P^\gamma \omega, j, j \mid -, -]} \in \mathcal{D}_{LC_1}^{LC_{pre}}$$

existe un paso

$$\overline{[O^\gamma \rightarrow \bullet P^\gamma \omega, j, j \mid -, -]} \in \mathcal{D}_{buLC_1}^{LC_{pre}}$$

y, por tanto, existe la inferencia

$$[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q] \vdash_{buLC_1} [O^\gamma \rightarrow \bullet P^\gamma \omega, j, j \mid -, -]$$

- Dado un paso

$$\frac{\frac{[N^\gamma \rightarrow \bullet M^\gamma \nu, j, j \mid -, -]}{[M^\gamma \rightarrow \nu \bullet, j, k \mid p, q]} \in \mathcal{D}_{LC_1}^{Comp_2}}{[N^\gamma \rightarrow M^\gamma \bullet \nu, j, k \mid p, q]} \in \mathcal{D}_{LC_1}^{Comp_2}$$

existe un paso

$$\frac{[M^\gamma \rightarrow \nu \bullet, j, k \mid p, q]}{[N^\gamma \rightarrow M^\gamma \bullet \omega, j, k \mid p, q]} \in \mathcal{D}_{buLC_1}^{LC_n}$$

y, por tanto, existe la inferencia

$$\frac{[N^\gamma \rightarrow \bullet M^\gamma \nu, j, j \mid -, -]}{[M^\gamma \rightarrow \nu \bullet, j, k \mid p, q]} \vdash_{buLC_1} [N^\gamma \rightarrow M^\gamma \bullet \nu, j, k \mid p, q]$$

- Dado un paso

$$\frac{\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[a, j, j+1]} \in \mathcal{D}_{LC_1}^{LA_t}}{[O^\beta \rightarrow P^\beta \bullet \omega, j, j+1 \mid -, -]} \in \mathcal{D}_{LC_1}^{LA_t}$$

existe un paso

$$\overline{[a, j, j+1]} \in \mathcal{D}_{buLC_1}^{LC_t}$$

y, por tanto, existe la inferencia

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[a, j, j+1]} \vdash_{buLC_1} [O^\beta \rightarrow P^\beta \bullet \omega, j, j+1 \mid -, -]$$

- Dado un paso

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[O^\beta \rightarrow P^\beta \bullet \omega, j, j \mid -, -]} \in \mathcal{D}_{LC_1}^{LA_\varepsilon}$$

existe un paso

$$\overline{[O^\beta \rightarrow P^\beta \bullet \omega, j, j \mid -, -]} \in \mathcal{D}_{buLC_1}^{LC_\varepsilon}$$

y, por tanto, existe la inferencia

$$[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q] \vdash_{buLC_1} [O^\beta \rightarrow P^\beta \bullet \omega, j, j \mid -, -]$$

- Dado un paso

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[O^\beta \rightarrow \bullet P^\beta \omega, j, j \mid -, -]} \in \mathcal{D}_{LC_1}^{LA_{pre}}$$

existe un paso

$$\overline{[O^\beta \rightarrow \bullet P^\beta \omega, j, j \mid -, -]} \in \mathcal{D}_{buLC_1}^{LC_{pre}}$$

y, por tanto, existe la inferencia

$$[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q] \vdash_{buLC_1} [O^\beta \rightarrow \bullet P^\beta \omega, j, j \mid -, -]$$

- Dado un paso

$$\frac{\frac{[F^\beta \rightarrow \bullet \perp, k, k \mid -, -]}{[a, k, k + 1]}}{[O^\gamma \rightarrow P^\gamma \bullet \nu, k, k + 1 \mid -, -]} \in \mathcal{D}_{LC_1}^{LF_t}$$

existe un paso

$$\overline{[a, j, j + 1]} \in \mathcal{D}_{buLC_1}^{LC_t}$$

y, por tanto, existe la inferencia

$$\frac{[F^\beta \rightarrow \bullet \perp, k, k \mid -, -]}{[a, k, k + 1]} \vdash_{buLC_1} [O^\gamma \rightarrow P^\gamma \bullet \nu, k, k + 1 \mid -, -]$$

- Dado un paso

$$\frac{[F^\beta \rightarrow \bullet \perp, k, k \mid -, -]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, k, k \mid -, -]} \in \mathcal{D}_{LC_1}^{LF_\varepsilon}$$

existe un paso

$$\overline{[O^\gamma \rightarrow P^\gamma \bullet \nu, k, k \mid -, -]} \in \mathcal{D}_{buLC_1}^{LC_\varepsilon}$$

y, por tanto, existe la inferencia

$$[F^\beta \rightarrow \bullet \perp, k, k \mid -, -] \vdash_{buLC_1} [O^\gamma \rightarrow P^\gamma \bullet \nu, k, k \mid -, -]$$

- Dado un paso

$$\frac{[\mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, k, k \mid -, -]} \in \mathcal{D}_{LC_1}^{LF_{pre}}$$

existe un paso

$$\frac{}{[O^\gamma \rightarrow \bullet P^\gamma \nu, k, k \mid -, -]} \in \mathcal{D}_{buLC_1}^{LC_{pre}}$$

y, por tanto, existe la inferencia

$$[\mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -] \vdash_{buLC_1} [O^\gamma \rightarrow \bullet P^\gamma \nu, k, k \mid -, -]$$

- Dado un paso

$$\frac{\frac{[\mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -]}{[M^\gamma \rightarrow \nu \bullet, k, l \mid p, q]}}{[\mathbf{F}^\beta \rightarrow \perp \bullet, k, l \mid k, l]} \in \mathcal{D}_{LC_1}^{FootComp}$$

existe un paso

$$\frac{[M^\gamma \rightarrow \nu \bullet, k, l \mid p, q]}{[\mathbf{F}^\beta \rightarrow \perp \bullet, k, l \mid k, l]} \mathcal{D}_{buLC_1}^{Foot}$$

y, por tanto, existe la inferencia

$$[\mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -] \vdash_{buLC_1} [M^\gamma \rightarrow \nu \bullet, k, l \mid p, q] \vdash_{buLC_1} [\mathbf{F}^\beta \rightarrow \perp \bullet, k, l \mid k, l]$$

- El resto de pasos deductivos cumplen las siguientes igualdades:

$$\mathcal{D}_{LC_1}^{Scan} = \mathcal{D}_{buLC_1}^{Scan}$$

$$\mathcal{D}_{LC_1}^e = \mathcal{D}_{buLC_1}^e$$

$$\mathcal{D}_{LC_1}^{LC_n} = \mathcal{D}_{buLC_1}^{LC_n}$$

$$\mathcal{D}_{LC_1}^{Comp_1} = \mathcal{D}_{buLC_1}^{Comp}$$

$$\mathcal{D}_{LC_1}^{AdjComp_1} = \mathcal{D}_{buLC_1}^{AdjComp}$$

$$\mathcal{D}_{LC_1}^{AdjComp_2} = \mathcal{D}_{buLC_1}^{LC_{cad}}$$

### Teorema 6.3 Relación entre los esquemas LC y LC'

Se mantiene la siguiente relación de contracción de secuencias deductivas:

$$LC \xrightarrow{sc} LC'$$

Prueba

Tenemos que probar que:

1.  $\mathcal{I}_{LC'} \subseteq \mathcal{I}_{LC}$
2.  $\vdash_{LC'}^* \subseteq \vdash_{LC}^*$

Lo primero es cierto por definición, ya que los dominios de ambos esquemas son iguales:

$$\mathcal{I}_{LC} = \mathcal{I}_{LC'}$$

Para probar que  $\vdash_{LC'}^* \sqsubseteq \vdash_{LC}^*$  debemos demostrar que  $\mathcal{D}_{LC'} \sqsubseteq \mathcal{D}_{LC}^*$ . Veamos cada caso particular:

- Un paso deductivo  $\mathcal{D}_{LC'}^{\text{LF}_i}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{LC}^{\text{LC}_{\text{pre}}}$

$$\frac{[N^\beta \rightarrow \delta \bullet E^\beta \nu, j, k \mid -, -]}{[\mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -]}$$

seguido de otro  $\mathcal{D}_{LC}^{\text{LF}_t}$

$$\frac{[\mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -] \quad [a, k, k + 1]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, k, k + 1 \mid -, -]}$$

- Un paso deductivo  $\mathcal{D}_{LC'}^{\text{LF}_e}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{LC}^{\text{LC}_{\text{pre}}}$

$$\frac{[N^\beta \rightarrow \delta \bullet E^\beta \nu, j, k \mid -, -]}{[\mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -]}$$

seguido de otro  $\mathcal{D}_{LC}^{\text{LF}_e}$

$$\frac{[\mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, k, k \mid -, -]}$$

- Un paso deductivo  $\mathcal{D}_{LC'}^{\text{LF}_{\text{pre}}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{LC}^{\text{LC}_{\text{pre}}}$

$$\frac{[N^\beta \rightarrow \delta \bullet E^\beta \nu, j, k \mid -, -]}{[\mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -]}$$

seguido de otro  $\mathcal{D}_{LC}^{\text{LF}_{\text{pre}}}$

$$\frac{[\mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, k, k \mid -, -]}$$

- Un paso deductivo  $\mathcal{D}_{LC'}^{\text{FootComp}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{LC}^{\text{LC}_{\text{pre}}}$

$$\frac{[N^\beta \rightarrow \delta \bullet E^\beta \nu, j, k \mid -, -]}{[\mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -]}$$

seguido de otro  $\mathcal{D}_{LC}^{\text{FootComp}}$

$$\frac{[\mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -] \quad [M^\gamma \rightarrow \nu \bullet, k, l \mid p, q]}{[\mathbf{F}^\beta \rightarrow \perp \bullet, k, l \mid k, l]}$$

- El resto de pasos deductivos son iguales en ambos esquemas

**Teorema 6.4** *Relación entre los esquemas pLC y pLC<sup>vpp</sup>*

Se mantiene la siguientes relaciones de filtro dinámico y refinamiento de ítems:

$$\text{pLC} \xrightarrow{\text{df}} \text{pLC}_1 \xrightarrow{\text{ir}} \text{pLC}^{\text{vpp}}$$

*Prueba*

Primero vamos a definir el esquema intermedio  $\text{pLC}_1$  añadiendo el antecedente  $[M^\gamma, j]$  en los pasos de filtrado de predicción y completión de pie del esquema  $\text{pLC}$ .

El dominio de  $\text{pLC}_1$  es:

$$\mathcal{I}_{\text{pLC}_1} = \mathcal{I}_{\text{pLC}}$$

El conjunto de pasos deductivos del nuevo esquema es:

$$\begin{aligned} \mathcal{D}_{\text{LC}_1} = & \mathcal{D}_{\text{pLC}_1}^{\text{LI}_t} \cup \mathcal{D}_{\text{pLC}_1}^{\text{LI}_\varepsilon} \cup \mathcal{D}_{\text{pLC}_1}^{\text{LI}_{\text{pre}}} \cup \mathcal{D}_{\text{pLC}_1}^{\text{Scan}} \cup \mathcal{D}_{\text{pLC}_1}^\varepsilon \cup \mathcal{D}_{\text{pLC}_1}^{\text{LC}_t} \cup \mathcal{D}_{\text{pLC}_1}^{\text{LC}_\varepsilon} \cup \\ & \mathcal{D}_{\text{pLC}_1}^{\text{LC}_{\text{pre}}} \cup \mathcal{D}_{\text{pLC}_1}^{\text{LC}_n} \cup \mathcal{D}_{\text{pLC}_1}^{\text{Pre}} \cup \mathcal{D}_{\text{pLC}_1}^{\text{Comp}} \cup \mathcal{D}_{\text{pLC}_1}^{\text{LA}_t} \cup \mathcal{D}_{\text{pLC}_1}^{\text{LA}_\varepsilon} \cup \\ & \mathcal{D}_{\text{pLC}_1}^{\text{LA}_{\text{pre}}} \cup \mathcal{D}_{\text{LC}_1}^{\text{AdjComp}} \cup \mathcal{D}_{\text{LC}_1}^{\text{LF}_t} \cup \mathcal{D}_{\text{LC}_1}^{\text{LF}_\varepsilon} \cup \mathcal{D}_{\text{LC}_1}^{\text{LF}_{\text{pre}}} \cup \mathcal{D}_{\text{LC}_1}^{\text{FootComp}} \end{aligned}$$

donde todos los pasos deductivos son iguales a sus homónimos del esquema  $\text{pLC}$ , salvo los cuatro nombrados anteriormente.

$$\begin{aligned} \mathcal{D}_{\text{pLC}_1}^{\text{LI}_t} &= \mathcal{D}_{\text{LC}}^{\text{pLI}_t} & \mathcal{D}_{\text{pLC}_1}^{\text{LI}_\varepsilon} &= \mathcal{D}_{\text{LC}}^{\text{pLI}_\varepsilon} & \mathcal{D}_{\text{pLC}_1}^{\text{LI}_{\text{pre}}} &= \mathcal{D}_{\text{LC}}^{\text{pLI}_{\text{pre}}} \\ \mathcal{D}_{\text{pLC}_1}^{\text{Scan}} &= \mathcal{D}_{\text{pLC}}^{\text{Scan}} & \mathcal{D}_{\text{pLC}_1}^\varepsilon &= \mathcal{D}_{\text{pLC}}^\varepsilon & \mathcal{D}_{\text{pLC}_1}^{\text{LC}_t} &= \mathcal{D}_{\text{pLC}}^{\text{LC}_t} \\ \mathcal{D}_{\text{pLC}_1}^{\text{LC}_\varepsilon} &= \mathcal{D}_{\text{pLC}}^{\text{LC}_\varepsilon} & \mathcal{D}_{\text{pLC}_1}^{\text{LC}_{\text{pre}}} &= \mathcal{D}_{\text{pLC}}^{\text{LC}_{\text{pre}}} & \mathcal{D}_{\text{pLC}_1}^{\text{LC}_n} &= \mathcal{D}_{\text{pLC}}^{\text{LC}_n} \\ \mathcal{D}_{\text{pLC}_1}^{\text{LA}_t} &= \mathcal{D}_{\text{pLC}}^{\text{LA}_t} & \mathcal{D}_{\text{pLC}_1}^{\text{LA}_\varepsilon} &= \mathcal{D}_{\text{pLC}}^{\text{LA}_\varepsilon} & \mathcal{D}_{\text{pLC}_1}^{\text{LA}_{\text{pre}}} &= \mathcal{D}_{\text{pLC}}^{\text{LA}_{\text{pre}}} \\ \mathcal{D}_{\text{pLC}_1}^{\text{Pre}} &= \mathcal{D}_{\text{pLC}}^{\text{Pre}} & \mathcal{D}_{\text{pLC}_1}^{\text{Comp}} &= \mathcal{D}_{\text{pLC}}^{\text{Comp}} & \mathcal{D}_{\text{pLC}_1}^{\text{AdjComp}} &= \mathcal{D}_{\text{pLC}}^{\text{AdjComp}} \\ \mathcal{D}_{\text{pLC}_1}^{\text{LF}_t} &= \frac{\begin{array}{l} [M^\gamma, j] \\ [E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -] \\ [a, k, k+1] \end{array}}{[M^\gamma; O\gamma \rightarrow P^\gamma \bullet \nu, k, k+1 \mid -, -]} & & & & \text{label}(P^\gamma) = a \\ & & & & & \beta \in \text{adj}(M^\gamma) \\ \mathcal{D}_{\text{pLC}_1}^{\text{LF}_\varepsilon} &= \frac{\begin{array}{l} [M^\gamma, j] \\ [E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -] \end{array}}{[M^\gamma; O\gamma \rightarrow P^\gamma \bullet \nu, k, k \mid -, -]} & & & & \text{label}(P^\gamma) = \varepsilon \\ & & & & & \beta \in \text{adj}(M^\gamma) \\ \mathcal{D}_{\text{pLC}_1}^{\text{LF}_{\text{pre}}} &= \frac{\begin{array}{l} [M^\gamma, j] \\ [E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -] \end{array}}{[M^\gamma; O\gamma \rightarrow \bullet P^\gamma \nu, k, k \mid -, -]} & & & & \beta \in \text{adj}(M^\gamma) \\ & & & & & P^\gamma >_\ell \Delta \text{ ó label}(P^\gamma) = \perp \\ \mathcal{D}_{\text{pLC}_1}^{\text{FootComp}} &= \frac{\begin{array}{l} [M^\gamma, j] \\ [E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -] \\ [M^\gamma; M^\gamma \rightarrow \nu \bullet, k, l \mid p, q] \end{array}}{[E^\beta; \mathbf{F}^\beta \rightarrow \perp \bullet, k, l \mid k, l]} & & & & \beta \in \text{adj}(M^\gamma) \end{aligned}$$

El conjunto de ítems finales es igual al del esquema **pLC**:

$$\mathcal{F}_{\text{pLC}_1} = \mathcal{F}_{\text{pLC}}$$

Para probar que  $\text{pLC} \xrightarrow{\text{df}} \text{pLC}_1$  tenemos que demostrar que:

$$1. \mathcal{I}_{\text{pLC}_1} \subseteq \mathcal{I}_{\text{pLC}}$$

$$2. \vdash_{\text{pLC}_1} \subseteq \vdash_{\text{pLC}}$$

Lo primero es cierto por definición, ya que los dominios de ambos esquemas son iguales:

$$\mathcal{I}_{\text{pLC}_1} = \mathcal{I}_{\text{pLC}}$$

Para probar que  $\vdash_{\text{pLC}_1} \subseteq \vdash_{\text{pLC}}$  debemos demostrar que  $\mathcal{D}_{\text{pLC}_1} \subseteq \mathcal{D}_{\text{pLC}}$ . Sólo vamos a considerar los cuatro pasos en los que difieren ambos esquemas:

- Dado un paso

$$\frac{\begin{array}{l} [M^\gamma, j] \\ [E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -] \\ [a, k, k + 1] \end{array}}{[M^\gamma; O^\gamma \rightarrow P^\gamma \bullet \nu, k, k + 1 \mid -, -]} \in \mathcal{D}_{\text{pLC}_1}^{\text{LF}_t}$$

existe un paso

$$\frac{\begin{array}{l} [E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -] \\ [a, k, k + 1] \end{array}}{[M^\gamma; O^\gamma \rightarrow P^\gamma \bullet \nu, k, k + 1 \mid -, -]} \in \mathcal{D}_{\text{pLC}}^{\text{LF}_t}$$

y, por tanto, existe la inferencia

$$\frac{\begin{array}{l} [M^\gamma, j] \\ [E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -] \\ [a, k, k + 1] \end{array}}{\vdash_{\text{pLC}} [M^\gamma; O^\gamma \rightarrow P^\gamma \bullet \nu, k, k + 1 \mid -, -]}$$

- Dado un paso

$$\frac{\begin{array}{l} [M^\gamma, j] \\ [E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -] \end{array}}{[M^\gamma; O^\gamma \rightarrow P^\gamma \bullet \nu, k, k \mid -, -]} \in \mathcal{D}_{\text{pLC}_1}^{\text{LF}_\epsilon}$$

existe un paso

$$\frac{\begin{array}{l} [E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -] \end{array}}{[M^\gamma; O^\gamma \rightarrow P^\gamma \bullet \nu, k, k \mid -, -]} \in \mathcal{D}_{\text{pLC}}^{\text{LF}_\epsilon}$$

y, por tanto, existe la inferencia

$$\frac{\begin{array}{l} [M^\gamma, j] \\ [E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -] \end{array}}{\vdash_{\text{pLC}} [M^\gamma; O^\gamma \rightarrow P^\gamma \bullet \nu, k, k \mid -, -]}$$

- Dado un paso

$$\frac{\begin{array}{c} [M^\gamma, j] \\ [E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -] \end{array}}{[M^\gamma; O^\gamma \rightarrow \bullet P^\gamma \nu, k, k \mid -, -]} \in \mathcal{D}_{\text{pLC}_1}^{\text{LF}_{\text{pre}}}$$

existe un paso

$$\frac{[E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -]}{[M^\gamma; O^\gamma \rightarrow \bullet P^\gamma \nu, k, k \mid -, -]} \in \mathcal{D}_{\text{pLC}}^{\text{LF}_{\text{pre}}}$$

y, por tanto, existe la inferencia

$$\frac{[M^\gamma, j]}{[E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -]} \vdash_{\text{pLC}} [M^\gamma; O^\gamma \rightarrow \bullet P^\gamma \nu, k, k \mid -, -]$$

- Dado un paso

$$\frac{\begin{array}{c} [M^\gamma, j] \\ [E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -] \\ [M^\gamma; M^\gamma \rightarrow \nu \bullet, k, l \mid p, q] \end{array}}{[E^\beta; \mathbf{F}^\beta \rightarrow \perp \bullet, k, l \mid k, l]} \in \mathcal{D}_{\text{pLC}_1}^{\text{FootComp}}$$

existe un paso

$$\frac{\begin{array}{c} [E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -] \\ [M^\gamma; M^\gamma \rightarrow \nu \bullet, k, l \mid p, q] \end{array}}{[E^\beta; \mathbf{F}^\beta \rightarrow \perp \bullet, k, l \mid k, l]} \in \mathcal{D}_{\text{pLC}}^{\text{FootComp}}$$

y, por tanto, existe la inferencia

$$\frac{\begin{array}{c} [M^\gamma, j] \\ [E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -] \\ [M^\gamma; M^\gamma \rightarrow \nu \bullet, k, l \mid p, q] \end{array}}{[E^\beta; \mathbf{F}^\beta \rightarrow \perp \bullet, k, l \mid k, l]} \vdash_{\text{pLC}} [E^\beta; \mathbf{F}^\beta \rightarrow \perp \bullet, k, l \mid k, l]$$

Queda demostrar la relación  $\text{pLC}_1 \xrightarrow{\text{ir}} \text{pLC}^{\text{vpp}}$ , para lo que debemos probar que existe una función regular  $f: \mathcal{I}_{\text{pLC}^{\text{vpp}}} \rightarrow \mathcal{I}_{\text{pLC}}$  tal que:

1.  $\mathcal{I}_{\text{pLC}} = f(\mathcal{I}_{\text{pLC}^{\text{vpp}}})$
2.  $\Delta_{\text{pLC}} = f(\Delta_{\text{pLC}^{\text{vpp}}})$

Una función regular de contracción de ítems que cumple estas condiciones es la siguiente:

$$f([C^\gamma; N^\gamma \rightarrow \delta \bullet \nu, h, i, j \mid p, q]) = [C^\gamma; N^\gamma \rightarrow \delta \bullet \nu, i, j \mid p, q]$$

$$f([C^\gamma, h, j]) = [C^\gamma, j]$$

De  $f$  se sigue inmediatamente que  $\mathcal{I}_{\text{pLC}} = f(\mathcal{I}_{\text{pLC}^{\text{vpp}}})$  y  $\Delta_{\text{pLC}} = f(\Delta_{\text{pLC}^{\text{vpp}}})$  por inducción en la longitud de las secuencias de derivación.

## 6.2 Integrando los esquemas tipo Left Corner en la red de analizadores para TAG

Ahora veremos las relaciones que existen entre los esquemas tipo Left Corner y los esquemas tipo Earley, tanto en sus versiones ascendentes como predictivas, de manera que podamos fusionar la redes de analizadores obtenida en la sección anterior y la basada en el algoritmo de Earley.

**Teorema 6.5** *Relación entre los esquemas buE y buLC*

*Se mantiene las siguientes relaciones de contracción de secuencias deductivas y filtros dinámicos:*

$$\text{buE} \xrightarrow{\text{sc}} \text{buLC}_2 \xrightarrow{\text{df}} \text{buLC}$$

*Prueba*

*Antes de iniciar la demostración vamos a definir el nuevo esquema que nos servirá de enlace para establecer la relación entre los dos esquemas de interés.*

*El dominio del esquema buLC<sub>2</sub> es igual al del esquema buLC:*

$$\mathcal{I}_{\text{buLC}_2} = \mathcal{I}_{\text{buLC}}$$

*El conjunto de pasos deductivos es:*

$$\begin{aligned} \mathcal{D}_{\text{buLC}_2} = & \mathcal{D}_{\text{buLC}_2}^{\text{LC}_t} \cup \mathcal{D}_{\text{buLC}_2}^{\text{LC}_e} \cup \mathcal{D}_{\text{buLC}_2}^{\text{LC}_n} \cup \mathcal{D}_{\text{buLC}_2}^{\text{LC}_{\text{cad}}} \cup \mathcal{D}_{\text{buLC}_2}^{\text{Foot}} \cup \\ & \mathcal{D}_{\text{buLC}_2}^{\text{Scan}} \cup \mathcal{D}_{\text{buLC}_2}^{\epsilon} \cup \mathcal{D}_{\text{buLC}_2}^{\text{Comp}} \cup \mathcal{D}_{\text{buLC}_2}^{\text{AdjComp}} \end{aligned}$$

*donde todos los pasos son idénticos al los homónimos de buLC, excepto el paso  $\mathcal{D}_{\text{buLC}_2}^{\text{Foot}}$ .*

$$\begin{aligned} \mathcal{D}_{\text{buLC}_2}^{\text{LC}_t} &= \mathcal{D}_{\text{buLC}}^{\text{LC}_t} & \mathcal{D}_{\text{buLC}_2}^{\text{LC}_e} &= \mathcal{D}_{\text{buLC}}^{\text{LC}_e} & \mathcal{D}_{\text{buLC}_2}^{\text{LC}_n} &= \mathcal{D}_{\text{buLC}}^{\text{LC}_n} & \mathcal{D}_{\text{buLC}_2}^{\text{LC}_{\text{cad}}} &= \mathcal{D}_{\text{buLC}}^{\text{LC}_{\text{cad}}} \\ \mathcal{D}_{\text{buLC}_2}^{\text{Scan}} &= \mathcal{D}_{\text{buLC}}^{\text{Scan}} & \mathcal{D}_{\text{buLC}_2}^{\epsilon} &= \mathcal{D}_{\text{buLC}}^{\epsilon} & \mathcal{D}_{\text{buLC}_2}^{\text{Comp}} &= \mathcal{D}_{\text{buLC}}^{\text{Comp}} & \mathcal{D}_{\text{buLC}_2}^{\text{AdjComp}} &= \mathcal{D}_{\text{buLC}}^{\text{AdjComp}} \\ \mathcal{D}_{\text{buLC}_2}^{\text{Foot}} &= \overline{[\mathbf{F}^\beta \rightarrow \perp \bullet, k, l, k, l]} \end{aligned}$$

*El conjunto de ítems finales también es igual al del esquema buLC:*

$$\mathcal{F}_{\text{buLC}_2} = \mathcal{F}_{\text{buLC}}$$

*Ahora probaremos que  $\text{buE} \xrightarrow{\text{sc}} \text{buLC}_2$ , para lo cual tenemos que demostrar que:*

1.  $\mathcal{I}_{\text{buLC}_2} \subseteq \mathcal{I}_{\text{buE}}$
2.  $\vdash_{\text{buLC}_2}^* \subseteq \vdash_{\text{buE}}^*$

*Puesto que el dominio del esquema buLC<sub>2</sub> es igual al del esquema buE, excepto los ítems con el punto al comienzo de la regla, podemos concluir que:*

$$\mathcal{I}_{\text{buLC}_2} \subset \mathcal{I}_{\text{buE}}$$

*Para probar que  $\vdash_{\text{buLC}_2}^* \subseteq \vdash_{\text{buE}}^*$  debemos demostrar que  $\mathcal{D}_{\text{buLC}_2} \subseteq \vdash_{\text{buE}}^*$ . Veamos cada uno de los pasos:*

- Un paso deductivo  $\mathcal{D}_{\text{buLC}_2}^{\text{LC}_t}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{buE}}^{\text{Ini}}$

$$\overline{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j \mid -, -]}$$

seguido de otro  $\mathcal{D}_{\text{buE}}^{\text{Scan}}$

$$\frac{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j \mid -, -], [a, j, j+1]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, j+1 \mid -, -]}$$

- Un paso deductivo  $\mathcal{D}_{\text{buLC}_2}^{\text{LC}_\varepsilon}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{buE}}^{\text{Ini}}$

$$\overline{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j \mid -, -]}$$

seguido de otro  $\mathcal{D}_{\text{buE}}^\varepsilon$

$$\frac{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j \mid -, -]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, j \mid -, -]}$$

- Un paso deductivo  $\mathcal{D}_{\text{buLC}_2}^{\text{LC}_n}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{buE}}^{\text{Ini}}$

$$\overline{[Q^\gamma \rightarrow \bullet O^\gamma \omega, j, j \mid -, -]}$$

seguido de otro  $\mathcal{D}_{\text{buE}}^{\text{Comp}}$

$$\frac{[O^\gamma \rightarrow \nu \bullet, j, k \mid p, q], [Q^\gamma \rightarrow \bullet O^\gamma \omega, j, j \mid -, -]}{[Q^\gamma \rightarrow O^\gamma \bullet \omega, j, k \mid p, q]}$$

- Un paso deductivo  $\mathcal{D}_{\text{buLC}_2}^{\text{LC}_{\text{cad}}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{buE}}^{\text{Ini}}$

$$\overline{[Q^\gamma \rightarrow \bullet O^\gamma \omega, j, j \mid -, -]}$$

seguido de otro  $\mathcal{D}_{\text{buE}}^{\text{AdjComp}}$

$$\frac{[\Gamma \rightarrow \mathbf{R}^\beta \bullet, j, m \mid k, l], [O^\gamma \rightarrow \delta \bullet, k, l \mid p, q], [Q^\gamma \rightarrow \bullet O^\gamma \nu, j, j \mid -, -]}{[Q^\gamma \rightarrow O^\gamma \bullet \nu, j, m \mid p, q]}$$

- El paso de completación de pie es idéntico en ambos esquemas:

$$\mathcal{D}_{\text{buLC}_2}^{\text{Foot}} = \mathcal{D}_{\text{buE}}^{\text{Foot}}$$

- El resto de pasos deductivos del esquema  $\text{buLC}_2$  están incluidos en las inferencias de sus homónimos del esquema  $\text{buE}$ , ya que estos últimos contemplan las operaciones en cualquier posición, mientras que los pertenecientes a  $\text{buLC}_2$  sólo son aplicables en nodos que no sean hijos izquierdos. Por tanto, se cumple:

$$\mathcal{D}_{\text{buLC}_2}^{\text{Scan}} \subset \mathcal{D}_{\text{buE}}^{\text{Scan}}$$

$$\begin{aligned} \mathcal{D}_{\text{buLC}_2}^\varepsilon &\subset \mathcal{D}_{\text{buE}}^\varepsilon \\ \mathcal{D}_{\text{buLC}_2}^{\text{Comp}} &\subset \mathcal{D}_{\text{buE}}^{\text{Comp}} \\ \mathcal{D}_{\text{buLC}_2}^{\text{AdjComp}} &\subset \mathcal{D}_{\text{buE}}^{\text{AdjComp}} \end{aligned}$$

Ahora probaremos que  $\text{buLC}_2 \xrightarrow{\text{df}} \text{buLC}$ , para lo cual tenemos que demostrar que:

1.  $\mathcal{I}_{\text{buLC}} \subseteq \mathcal{I}_{\text{buLC}_2}$
2.  $\vdash_{\text{buLC}} \subseteq \vdash_{\text{buLC}_2}$

Lo primero es cierto por definición, ya que los dominios de ambos esquemas son iguales:

$$\mathcal{I}_{\text{buLC}} = \mathcal{I}_{\text{buLC}_2}$$

Para probar que  $\vdash_{\text{buLC}} \subseteq \vdash_{\text{buLC}_2}$  debemos demostrar que  $\mathcal{D}_{\text{buLC}} \subseteq \mathcal{D}_{\text{buLC}_2}$ . Todos los pasos deductivos con la misma denominación son idénticos en ambos esquemas, con la excepción del paso  $\mathcal{D}_{\text{buLC}}^{\text{Foot}}$  que incorpora un antecedente. Para este caso, se cumple que dado un paso

$$\frac{[O^\gamma \rightarrow \nu \bullet, k, l, p, q]}{[\mathbf{F}^\beta \rightarrow \perp \bullet, k, l, k, l]} \in \mathcal{D}_{\text{buLC}}^{\text{Foot}}$$

existe un paso

$$\frac{}{[\mathbf{F}^\beta \rightarrow \perp \bullet, k, l, k, l]} \in \mathcal{D}_{\text{buLC}_2}^{\text{Foot}}$$

y, por tanto, existe la inferencia

$$[O^\gamma \rightarrow \nu \bullet, k, l, p, q] \vdash_{\text{buLC}_2} [\mathbf{F}^\beta \rightarrow \perp \bullet, k, l, k, l]$$

### **Teorema 6.6** Relación entre los esquemas **E** y **LC**

Se mantiene la siguiente relación de contracción de secuencias deductivas:

$$\mathbf{E} \xrightarrow{\text{sc}} \mathbf{LC}$$

*Prueba*

Tenemos que demostrar que:

1.  $\mathcal{I}_{\text{LC}} \subseteq \mathcal{I}_{\mathbf{E}}$
2.  $\vdash_{\text{LC}}^* \subseteq \vdash_{\mathbf{E}}^*$

El dominio del esquema **LC** es igual al del esquema **E**, pero suprimiendo los ítems con el punto al comienzo de la regla en ciertos casos:

$$\mathcal{I}_{\text{LC}} \subset \mathcal{I}_{\mathbf{E}}$$

Para probar que  $\vdash_{\text{LC}}^* \subseteq \vdash_{\mathbf{E}}^*$  debemos demostrar que  $\mathcal{D}_{\text{LC}} \subseteq \mathcal{D}_{\mathbf{E}}^*$ . Veamos cada uno de los pasos:

- Un paso  $\mathcal{D}_{LC}^{LI_t}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_E^{Ini}$

$$\overline{[\top \rightarrow \bullet \mathbf{R}^\alpha, 0, 0 \mid -, -]}$$

seguido de una secuencia de pasos  $\mathcal{D}_E^{Pred}$

$$\frac{[\top \rightarrow \bullet \mathbf{R}^\alpha, 0, 0 \mid -, -]}{[\mathbf{R}^\alpha \rightarrow \bullet M^\gamma \nu, 0, 0 \mid -, -]}$$

$$\frac{[M^\gamma \rightarrow \bullet O^\gamma \nu, 0, 0 \mid -, -]}{[O^\gamma \rightarrow \bullet P^\gamma \delta, 0, 0 \mid -, -]}$$

y de un paso  $\mathcal{D}_E^{Scan}$

$$\frac{[O^\gamma \rightarrow \bullet P^\gamma \nu, 0, 0 \mid -, -], [a, 0, 1]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, 0, 1 \mid -, -]}$$

- Un paso  $\mathcal{D}_{LC}^{LI_\varepsilon}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_E^{Ini}$

$$\overline{[\top \rightarrow \bullet \mathbf{R}^\alpha, 0, 0 \mid -, -]}$$

seguido de una secuencia de pasos  $\mathcal{D}_E^{Pred}$

$$\frac{[\top \rightarrow \bullet \mathbf{R}^\alpha, 0, 0 \mid -, -]}{[\mathbf{R}^\alpha \rightarrow \bullet M^\gamma \nu, 0, 0 \mid -, -]}$$

$$\frac{[M^\gamma \rightarrow \bullet O^\gamma \nu, 0, 0 \mid -, -]}{[O^\gamma \rightarrow \bullet P^\gamma \delta, 0, 0 \mid -, -]}$$

y de un paso  $\mathcal{D}_E^\varepsilon$

$$\frac{[O^\gamma \rightarrow \bullet P^\gamma \nu, 0, 0 \mid -, -]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, 0, 0 \mid -, -]}$$

- Un paso  $\mathcal{D}_{LC}^{LI_{pre}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_E^{Ini}$

$$\overline{[\top \rightarrow \bullet \mathbf{R}^\alpha, 0, 0 \mid -, -]}$$

seguido de una secuencia de pasos  $\mathcal{D}_E^{Pred}$

$$\frac{[\top \rightarrow \bullet \mathbf{R}^\alpha, 0, 0 \mid -, -]}{[\mathbf{R}^\alpha \rightarrow \bullet M^\gamma \nu, 0, 0 \mid -, -]}$$

$$\frac{[M^\gamma \rightarrow \bullet O^\gamma \nu, 0, 0 \mid -, -]}{[O^\gamma \rightarrow \bullet P^\gamma \delta, 0, 0 \mid -, -]}$$

- Un paso  $\mathcal{D}_{LC}^{LC_t}$  es equivalente a una secuencia de pasos  $\mathcal{D}_E^{Pred}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[M^\gamma \rightarrow \bullet O^\gamma \omega, j, j \mid -, -]}$$

$$\frac{[M^\gamma \rightarrow \bullet O^\gamma \omega, j, j \mid -, -]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j \mid -, -]}$$

seguida de un paso  $D_E^{\text{Scan}}$

$$\frac{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j \mid -, -], [a, j, j+1]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, j+1 \mid -, -]}$$

- Un paso  $D_{LC}^{\text{LC}_e}$  es equivalente a una secuencia de pasos  $D_E^{\text{Pred}}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[M^\gamma \rightarrow \bullet O^\gamma \omega, j, j \mid -, -]}$$

$$\frac{[M^\gamma \rightarrow \bullet O^\gamma \omega, j, j \mid -, -]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j \mid -, -]}$$

seguida de un paso  $D_E^e$

$$\frac{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j \mid -, -]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, j \mid -, -]}$$

- Un paso  $D_{LC}^{\text{LC}_{pre}}$  es equivalente a una secuencia de pasos  $D_E^{\text{Pred}}$ :

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[M^\gamma \rightarrow \bullet O^\gamma \omega, j, j \mid -, -]}$$

$$\frac{[M^\gamma \rightarrow \bullet O^\gamma \omega, j, j \mid -, -]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j \mid -, -]}$$

- Un paso  $D_{LC}^{\text{LC}_n}$  es equivalente a una secuencia de pasos  $D_E^{\text{Pred}}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[M^\gamma \rightarrow \bullet O^\gamma \omega, j, j \mid -, -]}$$

$$\frac{[M^\gamma \rightarrow \bullet O^\gamma \omega, j, j \mid -, -]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j \mid -, -]}$$

seguida de un paso  $D_E^{\text{Comp}}$

$$\frac{[P^\gamma \rightarrow \delta \bullet, j, k \mid p, q]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j \mid -, -]}$$

$$\frac{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, k \mid p, q]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, k \mid p, q]}$$

- Un paso  $D_{LC}^{\text{LA}_t}$  es equivalente a la secuencia formada por un paso  $D_E^{\text{AdjPred}}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[\top \rightarrow \bullet \mathbf{R}^\beta, j, j \mid -, -]}$$

una secuencia de pasos  $D_E^{\text{Pred}}$

$$\frac{[\top \rightarrow \bullet \mathbf{R}^\beta, j, j \mid -, -]}{[\mathbf{R}^\beta \rightarrow \bullet M^\beta \omega, j, j \mid -, -]}$$

$$\frac{[M^\beta \rightarrow \bullet O^\beta \omega, j, j \mid -, -]}{[O^\beta \rightarrow \bullet P^\beta \nu, j, j \mid -, -]}$$

y un paso  $D_E^{\text{Scan}}$

$$\frac{[O^\beta \rightarrow \bullet P^\beta \nu, j, j \mid -, -], [a, j, j + 1]}{[O^\beta \rightarrow P^\beta \bullet \nu, j, j + 1 \mid -, -]}$$

- Un paso  $D_{LC}^{\text{LA}\epsilon}$  es equivalente a la secuencia formada por un paso  $D_E^{\text{AdjPred}}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[\top \rightarrow \bullet R^\beta, j, j \mid -, -]}$$

una secuencia de pasos  $D_E^{\text{Pred}}$

$$\frac{[\top \rightarrow \bullet R^\beta, j, j \mid -, -]}{[R^\beta \rightarrow \bullet M^\beta \omega, j, j \mid -, -]}$$

$$\frac{[M^\beta \rightarrow \bullet O^\beta \omega, j, j \mid -, -]}{[O^\beta \rightarrow \bullet P^\beta \nu, j, j \mid -, -]}$$

y un paso  $D_E^\epsilon$

$$\frac{[O^\beta \rightarrow \bullet P^\beta \nu, j, j \mid -, -]}{[O^\beta \rightarrow P^\beta \bullet \nu, j, j \mid -, -]}$$

- Un paso  $D_{LC}^{\text{LApre}}$  es equivalente a la secuencia formada por un paso  $D_E^{\text{AdjPred}}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q]}{[\top \rightarrow \bullet R^\beta, j, j \mid -, -]}$$

seguido de una secuencia de pasos  $D_E^{\text{Pred}}$

$$\frac{[\top \rightarrow \bullet R^\beta, j, j \mid -, -]}{[R^\beta \rightarrow \bullet M^\beta \omega, j, j \mid -, -]}$$

$$\frac{[M^\beta \rightarrow \bullet O^\beta \omega, j, j \mid -, -]}{[O^\beta \rightarrow \bullet P^\beta \nu, j, j \mid -, -]}$$

- Un paso  $D_{LC}^{\text{LFt}}$  es equivalente a la secuencia formada por un paso  $D_E^{\text{FootPred}}$

$$\frac{[F^\beta \rightarrow \bullet \perp, k, k \mid -, -]}{[N^\gamma \rightarrow \bullet M^\gamma \nu, k, k \mid -, -]}$$

una secuencia de pasos  $D_E^{\text{Pred}}$

$$\frac{[N^\gamma \rightarrow \bullet M^\gamma \nu, k, k \mid -, -]}{[M^\gamma \rightarrow \bullet O^\gamma \omega, k, k \mid -, -]}$$

$$\frac{[M^\gamma \rightarrow \bullet O^\gamma \omega, k, k \mid -, -]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, k, k \mid -, -]}$$

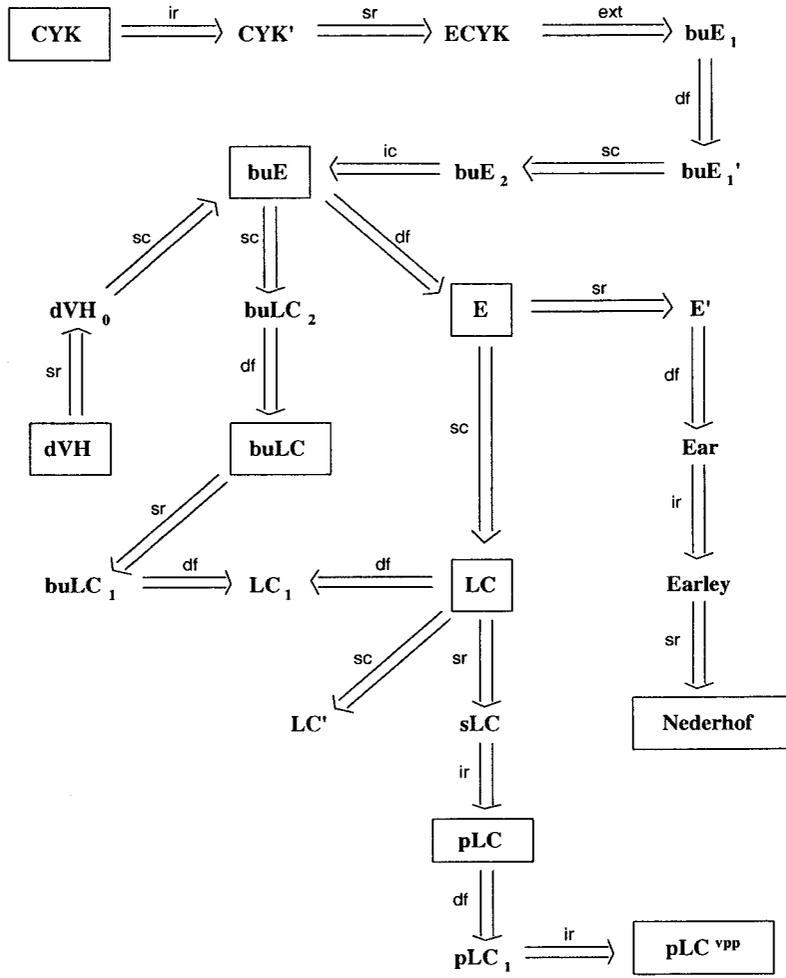


Figura 6.1: Red ampliada de analizadores para TAG

y un paso  $D_E^{\text{Scan}}$

$$\frac{[O^\gamma \rightarrow \bullet P^\gamma \nu, k, k \mid -, -], [a, k, k + 1]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, k, k + 1 \mid -, -]}$$

- Un paso  $D_{LC}^{\text{LF}^\epsilon}$  es equivalente a la secuencia formada por un paso  $D_E^{\text{FootPred}}$

$$\frac{[F^\beta \rightarrow \bullet \perp, k, k \mid -, -]}{[N^\gamma \rightarrow \bullet M^\gamma \nu, k, k \mid -, -]}$$

una secuencia de pasos  $D_E^{\text{Pred}}$

$$\frac{[N^\gamma \rightarrow \bullet M^\gamma \nu, k, k \mid -, -]}{[M^\gamma \rightarrow \bullet O^\gamma \omega, k, k \mid -, -]}$$

$$\frac{[M^\gamma \rightarrow \bullet O^\gamma \omega, k, k \mid -, -]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, k, k \mid -, -]}$$

y un paso  $\mathcal{D}_E^\epsilon$

$$\frac{[O^\gamma \rightarrow \bullet P^\gamma \nu, k, k \mid -, -]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, k, k \mid -, -]}$$

- Un paso  $\mathcal{D}_{LC}^{LF\text{pre}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_E^{\text{FootPred}}$

$$\frac{[F^\beta \rightarrow \bullet \perp, k, k \mid -, -]}{[N^\gamma \rightarrow \bullet M^\gamma \nu, k, k \mid -, -]}$$

seguido por una secuencia de pasos  $\mathcal{D}_E^{\text{Pred}}$

$$\frac{[N^\gamma \rightarrow \bullet M^\gamma \nu, k, k \mid -, -]}{[M^\gamma \rightarrow \bullet O^\gamma \omega, k, k \mid -, -]}$$

$$\frac{[M^\gamma \rightarrow \bullet O^\gamma \omega, k, k \mid -, -]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, k, k \mid -, -]}$$

- Los pasos deductivos de reconocimiento del esquema LC están incluidos en las inferencias de sus homónimos del esquema E, ya que estos últimos contemplan las operaciones en cualquier posición, mientras que los pertenecientes a LC sólo son aplicables en nodos que no sean hijos izquierdos. Por tanto, se cumple:

$$\mathcal{D}_{LC}^{\text{Scan}} \subset \mathcal{D}_E^{\text{Scan}}$$

$$\mathcal{D}_{LC}^\epsilon \subset \mathcal{D}_E^\epsilon$$

- El resto pasos son iguales en ambos esquemas:

$$\mathcal{D}_{LC}^{\text{Comp}} = \mathcal{D}_E^{\text{Comp}}$$

$$\mathcal{D}_{LC}^{\text{AdjComp}} = \mathcal{D}_E^{\text{AdjComp}}$$

$$\mathcal{D}_{LC}^{\text{FootComp}} = \mathcal{D}_E^{\text{FootComp}}$$

A modo de conclusión, en la figura 6.1 se presenta la red de analizadores para TAG descrita en el capítulo 4 ampliada con las relaciones mostradas en este capítulo.



## Parte III

# Gramáticas de inserción de árboles

## Capítulo 7

# Gramáticas de inserción de árboles

Las TAG ofrecen muchas ventajas a la hora de describir gramáticas para el procesamiento del lenguaje natural, como contrapartida, el coste computacional teórico derivado del uso de este formalismo es elevado. Recordemos que las TAG se analizan con un coste  $\mathcal{O}(n^6)$  en tiempo para el peor de los casos, donde  $n$  es la longitud de la cadena de entrada, frente a  $\mathcal{O}(n^3)$  para gramáticas incontextuales.

En este capítulo vamos a describir un nuevo formalismo gramatical, las gramáticas de inserción de árboles (TIG) [114], que es un refinamiento de las gramáticas incontextuales lexicalizadas introducidas por los mismos autores en [113]. TIG es un compromiso entre CFG y TAG, que combina la eficiencia en el análisis de la primera con las características lexicalizadas de la segunda. Las TIG pueden ser analizadas en  $\mathcal{O}(n^3)$ , pero al tratarse de un formalismo que se obtiene mediante la aplicación de restricciones al formalismo TAG, tanto el conjunto de árboles como el lenguaje que podemos derivar son más limitados.

El capítulo se estructura de la siguiente forma. En la sección 1 definiremos las gramáticas de inserción de árboles. En la sección 2 mostramos las relaciones entre las CFG, TAG y TIG. En la sección 3 se describe el analizador para TIG basado en el analizador Earley para CFG presentado en [114]. En la sección 4 mostramos la equivalencia entre el analizador Earley para CFG y el descrito en la sección 3, para ello introduciremos una nueva notación para representar los árboles elementales, la cual nos permitirá definir un nuevo analizador equivalente al de Earley para CFG.

### 7.1 Gramáticas de inserción de árboles

Las gramáticas de inserción de árboles (TIG) [114] es un formalismo gramatical que es una variante restringida de las gramáticas de adjunción de árboles. Al igual que en TAG, una gramática TIG consta de un conjunto de árboles iniciales y otro de árboles auxiliares, que

se combinan mediante operaciones de sustitución y adjunción. Sin embargo, tanto la forma de los árboles auxiliares como la adjunción se encuentran restringidas.

Una gramática de inserción de árboles es una 5-tupla  $(V_T, V_N, S, \mathbf{I}, \mathbf{A})$ , donde  $V_T$  es un conjunto de símbolos terminales,  $V_N$  es un conjunto de símbolos no terminales,  $S \in V_N$  es el axioma,  $\mathbf{I}$  es un conjunto finito de árboles iniciales y  $\mathbf{A}$  es un conjunto finito de árboles auxiliares. El conjunto  $\mathbf{I} \cup \mathbf{A}$  se denomina árboles elementales.

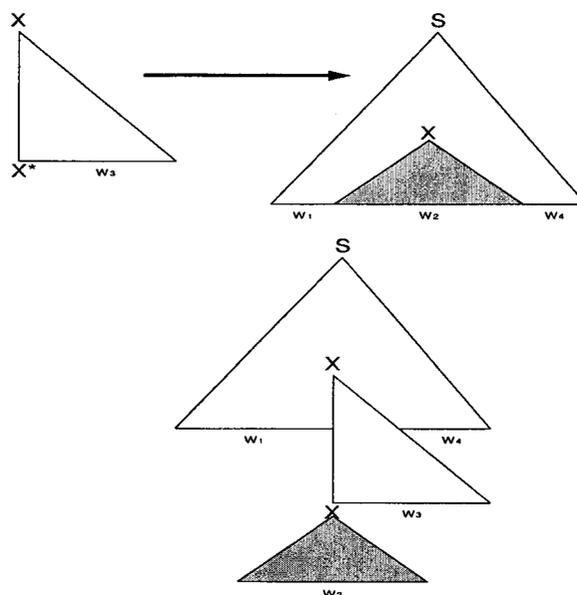
En los árboles iniciales los nodos interiores se etiquetan con símbolos no terminales. Los nodos de la frontera se etiquetan con símbolos terminales, la cadena vacía o símbolos no terminales. Los nodos de la frontera que están etiquetados con no terminales deben estar marcados para sustitución. Por convención, si un nodo está marcado para sustitución lo denotamos siguiéndolo de un símbolo  $\downarrow$ . En TIG se prohíbe la adjunción en nodos marcados para sustitución, porque los mismos árboles se pueden crear adjuntando en la raíces de los árboles sustituidos.

En los árboles auxiliares los nodos interiores se etiquetan con símbolos no terminales. Los nodos de la frontera se etiquetan con símbolos terminales, la cadena vacía o símbolos no terminales. Los nodos de la frontera que están etiquetados con no terminales deben estar marcados para sustitución, excepto el nodo pie.

Hasta este punto la definición de TIG es idéntica a la de TAG. Pero en lo que sigue difiere de ella. Antes de indicar las restricciones que presenta el formalismo TIG respecto a el TAG, vamos a definir algunos elementos nuevos. Los árboles auxiliares donde todos los nodos frontera no vacíos se encuentran a la derecha del nodo pie se denominan *árboles auxiliares derechos*. Los árboles auxiliares donde todos los nodos frontera no vacíos se encuentran a la izquierda del nodo pie se denominan *árboles auxiliares izquierdos*. El resto de árboles auxiliares se denominan *árboles auxiliares wrapping*. La adjunción de un árbol auxiliar derecho se denomina *adjunción derecha* (figura 7.1). La adjunción de un árbol auxiliar izquierdo se denomina *adjunción izquierda* (figura 7.2).

TIG presenta las siguientes características distintivas respecto al formalismo TAG:

- Prohíbe los árboles auxiliares wrapping y los árboles auxiliares nulos. Por tanto, solo permiten árboles auxiliares izquierdos y derechos.
- No permite que un árbol auxiliar sea adjuntado en ningún nodo a la izquierda (resp. derecha) de la espina de un árbol auxiliar derecho (resp. izquierdo).
- No permite que un árbol auxiliar izquierdo (resp. derecho) sea adjuntado en la espina de un árbol auxiliar derecho (resp. izquierdo).
- Permite *adjunción simultánea* de varios árboles auxiliares izquierdos y/o derechos sobre un mismo nodo, de manera similar a la derivación para TAG descrita en [111]. La adjunción simultánea se especifica mediante dos secuencias, una de árboles auxiliares izquierdos y otra de árboles auxiliares derechos, que indican el orden en que se combinan los árboles. En la figura 7.3 se muestran los árboles que se generan mediante

Figura 7.1: *Adjunción derecha*

la adjunción de dos árboles auxiliares, uno izquierdo y otro derecho, sobre el mismo nodo.

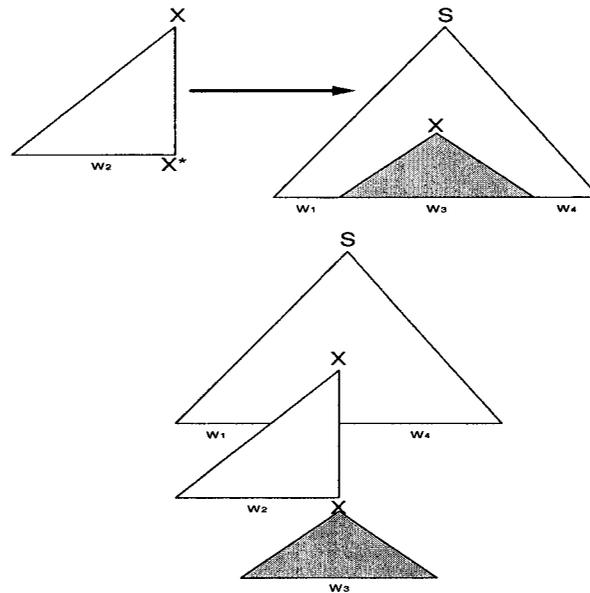
- Prohíbe la adjunción en nodos raíces y pies de árboles auxiliares, porque los mismos árboles se pueden generar mediante adjunción simultánea.

Es importante señalar que la adjunción de un árbol auxiliar de wrapping (figura 7.4) introduce entre dos cadenas a la cadena que produce el subárbol escindido. Esto es precisamente lo que hace de TAG un formalismo sensible al contexto [135] e incrementa considerablemente el coste computacional de su análisis. Por este motivo en las TIG se prohíbe el empleo de árboles auxiliares wrapping. No obstante, la eliminación de este tipo de árboles auxiliares no es suficiente para prevenir la aparición de esta sensibilidad al contexto, basta observar que una adjunción de un árbol auxiliar derecho en la espina de un árbol auxiliar izquierdo, y viceversa, produce el mismo efecto que una operación de adjunción wrapping, como se muestra en la figura 7.5. Este es el motivo por el cual también en TIG se prohíbe la adjunción de un árbol auxiliar izquierdo (resp. derecho) en la espina de un árbol auxiliar derecho (resp. izquierdo).

La adjunción simultánea es una operación en esencia ambigua, ya que aunque la secuencia fijada de árboles auxiliares izquierdos y derechos establece el orden en que se combinan las cadenas, el número de árboles que se puede producir es variable.

En la literatura nos podemos encontrar variantes de la definición original de TIG, como las siguientes:



Figura 7.2: *Adjunción izquierda***TIG lexicalizadas (LTIG, *Lexicalized Tree Insertion Grammar*)**

Presentada en [114]. Obligan a que al menos un nodo frontera de cada árbol elemental esté etiquetado con un símbolo terminal. Una LTIG se dice que está *anclada por la izquierda* si todos sus árboles elementales están anclados por la izquierda. Un árbol elemental está anclado por la izquierda si el primer elemento de la frontera, distinto del nodo pie, es un ítem léxico.

**LTIG estocásticas (SLTIG, *Stochastic Lexicalized Tree Insertion Grammar*)**

Presentada en [115]. Introducen parámetros estocásticos en LTIG para controlar la probabilidad de que una adjunción o sustitución ocurra.

**CFG lexicalizadas (LCFG, *Lexicalized Context Free Grammar*)**

Es el formalismo precursor de las TIG y fue presentada en [113]. A diferencia de las TIG, solo se permite que se adjunte en un mismo nodo como máximo un árbol auxiliar izquierdo y otro derecho. Esta restricción también se usa en [115] con el fin de simplificar el analizador sintáctico al estilo CYK que muestran estos autores. Es evidente que el objetivo de esta restricción sobre la adjunción simultánea es el de evitar la ambigüedad en los árboles derivados sin disminuir la potencia generadora de las TIG.

En este capítulo trabajaremos con la definición original de TIG sin ninguna variante y, por tanto, se permitirá la adjunción simultánea de múltiples árboles auxiliares izquierdos

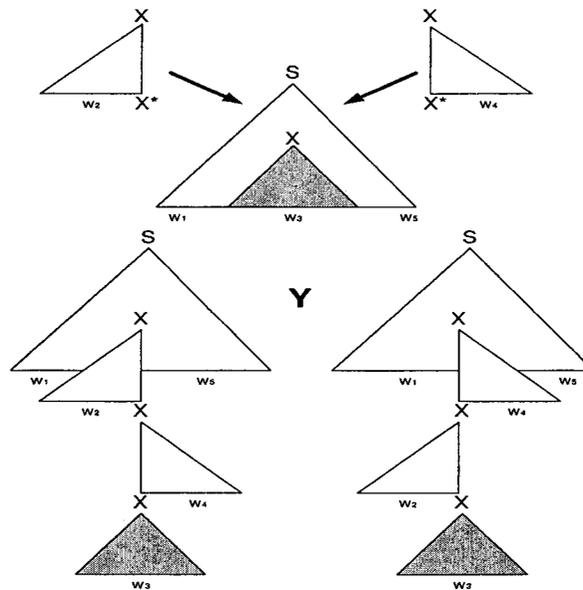


Figura 7.3: Adjunción simultánea

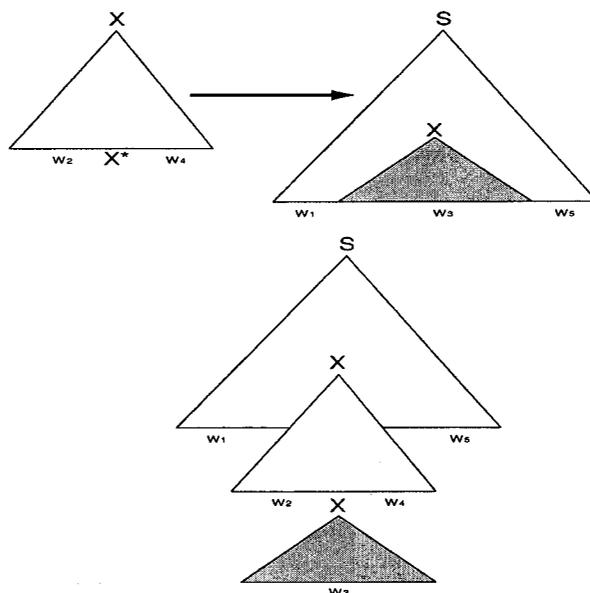
y derechos sobre un mismo nodo. Ésto va a reducir el número de estados del *chart*, ya que no se almacena información sobre si una adjunción ha tenido lugar en un nodo. No obstante, en el capítulo dedicado a los analizadores para TIG usaremos la restricción sobre la adjunción simultánea propuesta en [113], con objeto de no incrementar la ambigüedad de las derivaciones obtenidas en los análisis.

### 7.1.1 Notación

Puesto que TIG es un conjunto restringido TAG, vamos a mantener para ella toda la notación empleada en TAG. Es decir, denotamos con  $R^\gamma$  y  $F^\gamma$  la raíz y el pie de un árbol elemental  $\gamma$ . También añadimos a cada árbol elemental los dos nuevos nodos  $\top$  y  $\perp$  con restricciones de adjunción y sustitución nulas, de la misma forma en que lo hicimos para las TAG. Si un árbol auxiliar izquierdo o derecho  $\beta$  es adjuntable en un nodo  $M^\gamma$  lo indicamos como  $\beta \in adj(M^\gamma)$ . Si la adjunción derecha e izquierda en dicho nodo no es obligatoria, de nuevo usamos el símbolo **nil**, y lo denotamos como  $nil \in adj(M^\gamma)$ .

Sin embargo, en las TIG vamos a incluir alguna notación nueva.  $A_L$  es el conjunto de árboles auxiliares izquierdos de una gramática, mientras  $A_R$  es el conjunto de árboles auxiliares derechos. Si un árbol auxiliar  $\beta$  es un árbol auxiliar izquierdo, entonces  $\beta \in A_L$ . Si un árbol auxiliar  $\beta$  es un árbol auxiliar derecho, entonces  $\beta \in A_R$ . Si un árbol auxiliar izquierdo  $\beta \in A_L$  es adjuntable en un nodo  $M^\gamma$ , lo denotamos como  $\beta \in ladj(M^\gamma)$ . Si un árbol auxiliar derecho  $\beta \in A_R$  es adjuntable en un nodo  $M^\gamma$ , lo denotamos como  $\beta \in radj(M^\gamma)$ .



Figura 7.4: *Adjunción wrapping*

$\text{radj}(M^\gamma)$ . Cuando la adjunción izquierda no es obligatoria en dicho nodo, lo denotamos como  $\text{nil} \in \text{ladj}(M^\gamma)$  y, del mismo modo, cuando la adjunción derecha no es obligatoria, lo indicamos  $\text{nil} \in \text{radj}(M^\gamma)$ . Y, por último, si un árbol inicial  $\alpha \in I$  se puede sustituir en un nodo  $M^\gamma$  lo denotamos como  $\alpha \in \text{subst}(M^\gamma)$ .

## 7.2 Relaciones entre TIG, CFG y TAG

En esta sección vamos a comparar el formalismo TIG con los formalismos TAG y CFG. En primer lugar veremos como una gramática TIG sin ningún tipo de modificación es, en general, una gramática TAG. A continuación estableceremos la equivalencia generativa entre CFG y TIG, para lo cual mostraremos como cualquier CFG se puede convertir en una TIG y cualquier TIG se puede convertir a una CFG que genera el mismo lenguaje. De esta equivalencia se deduce que las TIG solo pueden generar lenguajes incontextuales.

Cualquier TIG que no haga uso de las restricciones de adjunción se puede convertir de manera trivial a una TAG que derive los mismos árboles. Veamos como afecta a esta conversión cada una de las diferencias entre ambos formalismos.

TIG prohíbe los árboles auxiliares de wrapping. Desde la perspectiva de esta diferencia, una TIG es una TAG sin ninguna modificación.

TIG prohíbe la adjunción en las raíces de los árboles auxiliares y permite la adjunción simultánea, mientras TAG permite la adjunción en las raíces de los árboles auxiliares y prohíbe la adjunción simultánea. Desde la perspectiva de esta diferencia, una TIG es una

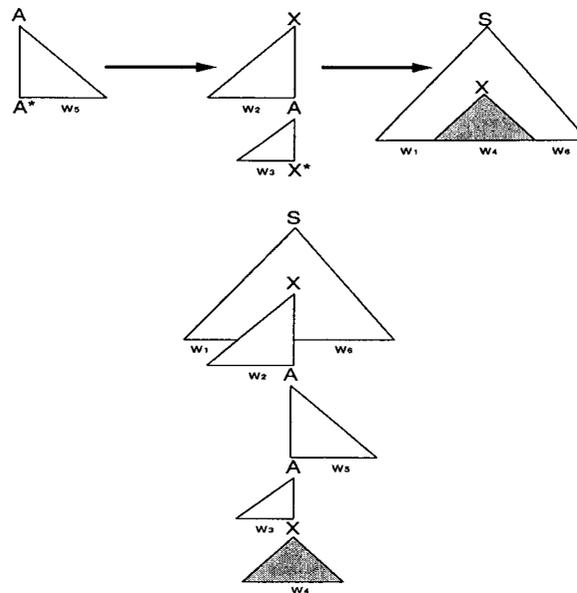


Figura 7.5: *Adjunción de un auxiliar derecho en la espina de un auxiliar izquierdo*

TAG sin ninguna modificación. Para demostrar ésto consideremos un conjunto de árboles auxiliares  $T \in \mathcal{A}$  que pueden adjuntarse en un nodo  $M^\gamma$ . Mediante adjunción simultánea en TIG podemos concatenar los árboles de  $T$  de todas las formas posibles en  $M^\gamma$ . En TAG podemos hacer lo mismo adjuntando los árboles de  $T$  sobre las raíces de los otros árboles auxiliares de  $T$ .

TIG prohíbe la adjunción de un árbol auxiliar izquierdo (resp. derecho) en la espina de un árbol auxiliar derecho (resp. izquierdo). Para convertir una TIG a una TAG que derive los mismos árboles y que contemple esta restricción será, en la mayoría de los casos, necesario añadir restricciones de adjunción en la TAG. Por tanto, si la TIG no tiene restricciones de adjunción, aparte de las propias de su definición, la conversión es factible. El problema se presenta cuando la TIG presenta restricciones de adjunción. En este caso la conversión puede ser difícil o incluso imposible.

Cualquier CFG se puede convertir de manera trivial a una TIG que derive los mismos árboles mediante el siguiente procedimiento: convertir cada regla de producción de la CFG en un árbol inicial de un nivel, donde la raíz del árbol es el símbolo de la parte izquierda de la producción y los símbolos de la parte derecha formarán la frontera del árbol; si la parte derecha de la regla está vacía, el único elemento de la frontera del árbol será  $\varepsilon$ ; los símbolos no terminales de la frontera serán marcados para sustitución. De ésto se concluye que cualquier lenguaje incontextual se puede generar mediante una TIG.

Para probar que las TIG solo pueden generar lenguajes incontextuales, solo nos queda probar que cualquier TIG se puede convertir en una CFG que genera el mismo lenguaje,

aunque no los mismos árboles. Mostraremos el procedimiento descrito en [115] para indicar como se puede llevar a cabo esta conversión.

Dada una TIG  $G$ , en primer lugar la convertimos a una *gramática de sustitución de árboles* (TSG)  $G'$  que genere el mismo lenguaje. Una TSG es lo mismo que una TIG salvo que no tienen árboles auxiliares y, por tanto, no requiere de la operación de adjunción. A continuación, esta gramática  $G'$  se convierte a una CFG  $G''$ .

Para crear  $G'$ , primero hacemos que cada árbol inicial de  $G$  sea un árbol inicial en  $G'$ . Después es necesario convertir los árboles auxiliares de  $G$  en árboles iniciales para poder combinarlos mediante la operación de sustitución. Para ello seguimos los siguientes pasos:

1. Etiquetamos todos los nodos pie con la cadena vacía  $\varepsilon$ .
2. Si la etiqueta de la raíz de un árbol auxiliar izquierdo es  $R$ , la renombramos como  $R_l$ . De forma análoga, si la raíz de un árbol auxiliar derecho es  $R$ , la renombramos como  $R_r$ .
3. Si un árbol auxiliar izquierdo es adjuntable en un nodo etiquetado  $R$ , añadimos un nuevo hijo a la izquierda con la etiqueta  $R_l$  y lo marcamos para sustitución. Además, si la adjunción no es obligatoria y no existe ya, añadimos un nuevo árbol asociado a la regla  $R_l \rightarrow \varepsilon$ . Análogamente, si un árbol auxiliar derecho es adjuntable en un nodo etiquetado  $R$ , añadimos un nuevo hijo a la derecha marcado para sustitución y etiquetado con  $R_r$ , así como el correspondiente árbol asociado a la regla  $R_r \rightarrow \varepsilon$ , si no existe y no presenta adjunción obligatoria.

La gramática TSG  $G'$  genera el mismo lenguaje que la TIG  $G$ , ya que todos los casos de adjunción son cubiertos por sustituciones equivalentes. Para transformar la TSG  $G'$  en una CFG  $G''$  equivalente, se aplanan cada árbol de  $G'$  en una regla CFG en  $G''$ , de forma que la etiqueta de la raíz se convierte en la parte izquierda de la regla y las etiquetas de la frontera en la parte derecha.

Aunque el lenguaje que genera una TIG es el mismo que el generado por una CFG, el conjunto de árboles que puede generar una TIG es mayor. Esto se comprueba observando el conjunto de *caminos* de los árboles derivados. El conjunto de caminos de una gramática es el conjunto de todos los caminos desde la raíz a la frontera de todos los árboles generados por la gramática. Así, mientras el conjunto de caminos para una CFG y una TSG es un lenguaje regular [129], el conjunto de caminos para una TIG y una TAG es un lenguaje incontextual. Este se puede intuir por el hecho de que una operación de adjunción sustituye el nodo adjuntado por una secuencia de nodos (la espina del árbol auxiliar adjuntado) en el camino del mismo y, luego los árboles auxiliares se comportan como las producciones de una CFG.

La figura 7.6 muestra un resumen de todo lo dicho en esta sección. El eje horizontal representa la complejidad de lenguajes que pueden generar los distintos formalismos: lenguajes regulares (RL), lenguajes incontextuales (CFL) y lenguajes de adjunción de árboles (TAL).

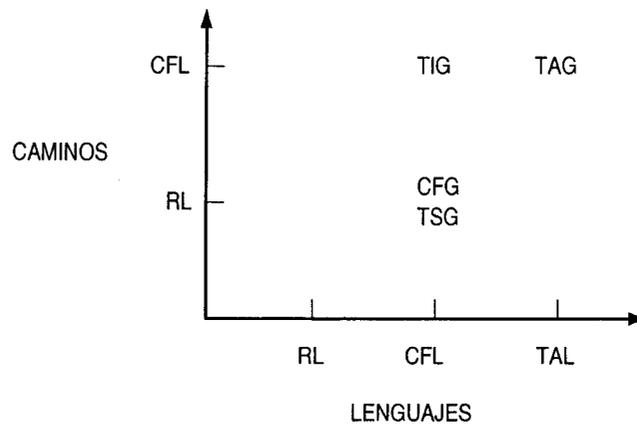


Figura 7.6: Complejidad de árboles y lenguajes en TIG, TAG, CFG y TSG

Frente al eje vertical, que representa la complejidad de los conjuntos de caminos que pueden generar dichos formalismos. Como se puede observar, TIG es un compromiso entre CFG y TAG, ya que al igual que CFG, solo genera lenguajes incontextuales, pero al igual que TAG, el conjunto de caminos que puede generar es también un lenguaje incontextual.

### 7.3 Un analizador tipo Earley para TIG

Como TIG es un caso restringido de TAG, es obvio que se pueden usar los analizadores existentes en la literatura para TAG. No obstante, como vimos en capítulos previos, estos analizadores presentan un coste computacional elevado. Es evidente que el objetivo fundamental del formalismo TIG es reducir los costes de análisis, aunque para ello se reduzca también la capacidad generativa de lenguajes a CFL. Se puede intuir que si solo produce CFL, el tiempo de análisis sea similar al de las gramáticas incontextuales.

En esta sección vamos a ver el analizador basado en Earley para TIG mostrado en [114], el cual presenta una complejidad de  $\mathcal{O}(n^3)$  en tiempo, análoga a la de CFG. Estos autores describen el algoritmo usando reglas de inferencia bajo las directrices de analizadores deductivos de [122]. Aunque la concepción es muy similar, pero con el objeto de mantener el marco de descripción de analizadores, nosotros reescribiremos este analizador usando los esquemas de análisis de [127]. Al igual que los autores, y de la misma forma que vimos en el capítulo dedicado a las TAG, vamos a utilizar la notación multicapa para transformar los árboles elementales en reglas de producción, asociando a cada árbol elemental  $\gamma$  un conjunto de producciones  $\mathcal{P}(\gamma)$ .

El dominio del esquema  $\mathbf{E}^i$  es similar al dominio del analizador Earley para CFG:

$$\mathcal{I}_{\mathbf{E}^i} = \{[M^\gamma \rightarrow \delta \bullet \nu, i, j]\}$$

donde  $M^\gamma \rightarrow \delta\nu \in \mathcal{P}(\gamma)$  y  $\gamma \in I \cup A$ . Los índices  $0 \leq i \leq j$  establecen las posiciones de la cadena de entrada que delimitan el fragmento reconocido por  $\delta$ .

Los pasos deductivos del esquema son una adaptación a las TIG de los pasos de inicio, reconocimiento, completión y predicción del analizador Earley para las CFG:

$$\begin{aligned} \mathcal{D}_{E^i} = & \mathcal{D}_{E^i}^{\text{Ini}} \cup \mathcal{D}_{E^i}^{\text{Scan}} \cup \mathcal{D}_{E^i}^\varepsilon \cup \mathcal{D}_{E^i}^{\text{Pred}} \cup \mathcal{D}_{E^i}^{\text{Comp}} \cup \mathcal{D}_{E^i}^{\text{Foot}} \cup \mathcal{D}_{E^i}^{\text{LAdjPred}} \cup \\ & \mathcal{D}_{E^i}^{\text{LAdjComp}} \cup \mathcal{D}_{E^i}^{\text{RAAdjPred}} \cup \mathcal{D}_{E^i}^{\text{RAAdjComp}} \cup \mathcal{D}_{E^i}^{\text{SubsPred}} \cup \mathcal{D}_{E^i}^{\text{SubsComp}} \end{aligned}$$

### Inicio

El análisis se inicia lanzando las predicciones de todos los árboles iniciales cuyas raíces estén etiquetadas con el axioma de la gramática.

$$\mathcal{D}_{E^i}^{\text{Ini}} = \frac{\alpha \in I}{[\top \rightarrow \bullet \mathbf{R}^\alpha, 0, 0]} \quad \text{label}(\mathbf{R}^\alpha) = S$$

### Reconocimiento

Los pasos  $\mathcal{D}_{E^i}^{\text{Scan}}$  llevan a cabo el reconocimiento de un símbolo de la cadena de entrada cuando el análisis lo alcanza. El paso  $\mathcal{D}_{E^i}^\varepsilon$  se aplica cuando dicho símbolo es la cadena vacía. Estas pasos son análogos a los pasos de reconocimiento del analizador Earley para CFG.

$$\begin{aligned} \mathcal{D}_{E^i}^{\text{Scan}} &= \frac{\begin{array}{c} [a, j, j+1] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j] \end{array}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, j+1]} \quad \text{label}(M^\gamma) = a \\ \mathcal{D}_{E^i}^\varepsilon &= \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j]}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, j]} \quad \text{label}(M^\gamma) = \varepsilon \end{aligned}$$

La forma que tienen los árboles auxiliares en las TIG nos permite ignorar los nodos pies de los árboles auxiliares cuando el reconocimiento los alcanza. Esto se debe al hecho de que en las TIG no existen árboles auxiliares de wrapping, que son los que provocan la necesidad de reconocer los subárboles escindidos.

$$\mathcal{D}_{E^i}^{\text{Foot}} = \frac{[\mathbf{F}^\beta \rightarrow \bullet \perp, j, j]}{[\mathbf{F}^\beta \rightarrow \perp \bullet, j, j]} \quad \beta \in A$$

### Predicción de subárbol

El paso  $\mathcal{D}_{E^i}^{\text{Pred}}$ , al igual que en el analizador Earley para CFG, inicia el reconocimiento del subárbol que pende de un nodo cuando dicho nodo no presente una restricción de adjunción obligatoria. Sin embargo, este paso si presenta una diferencia importante con el paso predictivo de las CFG, y es que también vamos a lanzar la predicción del subárbol cuando el nodo que domina el árbol sea un nodo adjuntable. Esto se justifica porque este esquema se adapta para "simular" las adjunciones izquierdas y derechas mediante sustituciones de los árboles auxiliares izquierdos y derechos antes y después de reconocer el subárbol, respectivamente.

$$\mathcal{D}_{E^i}^{\text{Pred}} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j]}{[M^\gamma \rightarrow \bullet \nu, j, j]} \quad \text{nil} \in \text{adj}(M^\gamma) \text{ o } \exists \beta \in \text{adj}(M^\gamma)$$

### Completación de subárbol

Cuando el subárbol dominado por un nodo se ha reconocido completamente, aplicamos el paso  $\mathcal{D}_{E^i}^{\text{Comp}}$  para continuar el reconocimiento ascendente.

$$\mathcal{D}_{E^i}^{\text{Comp}} = \frac{[M^\gamma \rightarrow \nu \bullet, j, k] \quad [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j]}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k]}$$

### Predicción y completación de adjunción izquierda

Los siguientes dos pasos reconocen la adjunción de un árbol auxiliar izquierdo sobre un nodo. La regla  $\mathcal{D}_{E^i}^{\text{LAdjPred}}$  es iniciada por ítems de la forma  $[M^\gamma \rightarrow \bullet \nu, i, j]$  cuando  $M^\gamma$  es un nodo adjuntable por la izquierda, mientras el reconocimiento ascendente lo realiza el paso  $\mathcal{D}_{E^i}^{\text{LAdjComp}}$ . No se almacena ninguna información de si ha tenido lugar alguna adjunción izquierda, luego los ítems de la forma  $[M^\gamma \rightarrow \bullet \nu, i, j]$  representan ambas situaciones: cuando una adjunción izquierda puede ocurrir y cuando ha ocurrido una o varias veces. Esto nos va a permitir realizar adjunciones izquierdas simultáneas sobre un mismo nodo.

$$\mathcal{D}_{E^i}^{\text{LAdjPred}} = \frac{[M^\gamma \rightarrow \bullet \nu, i, j]}{[\top \rightarrow \bullet \mathbf{R}^\beta, j, j]} \quad \beta \in \text{ladj}(M^\gamma)$$

$$\mathcal{D}_{E^i}^{\text{LAdjComp}} = \frac{[\top \rightarrow \mathbf{R}^\beta \bullet, j, k] \quad [M^\gamma \rightarrow \bullet \nu, i, j]}{[M^\gamma \rightarrow \bullet \nu, i, k]} \quad \beta \in \text{ladj}(M^\gamma)$$

### Predicción y completación de adjunción derecha

Los siguientes dos pasos reconocen la adjunción de un árbol auxiliar derecho sobre un nodo. La regla  $\mathcal{D}_{E^i}^{\text{RAAdjPred}}$  es iniciada por ítems de la forma  $[M^\gamma \rightarrow \nu \bullet, i, j]$  cuando  $M^\gamma$  es un nodo adjuntable por la derecha, mientras el reconocimiento ascendente lo realiza el paso  $\mathcal{D}_{E^i}^{\text{RAAdjComp}}$ . Al igual que en el caso de la adjunción izquierda, tampoco se almacena ninguna información de si ha tenido lugar alguna adjunción derecha, lo cual nos va a permitir realizar adjunciones derechas simultáneas sobre un mismo nodo.

$$\mathcal{D}_{E^i}^{\text{RAAdjPred}} = \frac{[M^\gamma \rightarrow \nu \bullet, i, j]}{[\top \rightarrow \bullet \mathbf{R}^\beta, j, j]} \quad \beta \in \text{radj}(M^\gamma)$$

$$\mathcal{D}_{E^i}^{\text{RAAdjComp}} = \frac{[\top \rightarrow \mathbf{R}^\beta \bullet, j, k] \quad [M^\gamma \rightarrow \nu \bullet, i, j]}{[M^\gamma \rightarrow \nu \bullet, i, k]} \quad \beta \in \text{radj}(M^\gamma)$$

### Predicción y completación de sustitución



Las reglas de sustitución se disparan por ítems de la forma  $[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j]$  cuando  $M^\gamma$  es un nodo de sustitución.  $\mathcal{D}_{E^i}^{\text{SubsPred}}$  lanza la predicción del árbol inicial y  $\mathcal{D}_{E^i}^{\text{SubsComp}}$  completa la sustitución.

$$\mathcal{D}_{E^i}^{\text{SubsPred}} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j]}{[\top \rightarrow \bullet \mathbf{R}^\alpha, j, j]} \quad \alpha \in \text{subst}(M^\gamma)$$

$$\mathcal{D}_{E^i}^{\text{SubsComp}} = \frac{[\top \rightarrow \mathbf{R}^\alpha \bullet, j, k]}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k]} \quad \alpha \in \text{subst}(M^\gamma)$$

El conjunto de ítems finales viene dado por:

$$\mathcal{F}_{E^i} = \{[\top \rightarrow \mathbf{R}^\alpha \bullet, 0, n] \mid \alpha \in I, \text{label}(\mathbf{R}^\alpha) = S\}$$

La complejidad espacial del algoritmo es de  $\mathcal{O}(|G|n^2)$  y la complejidad temporal en el caso peor es de  $\mathcal{O}(|G|^2n^3)$ , debida a los pasos deductivos de completación  $\mathcal{D}_{E^i}^{\text{Comp}}$ ,  $\mathcal{D}_{E^i}^{\text{LAdjComp}}$ ,  $\mathcal{D}_{E^i}^{\text{RAAdjComp}}$  y  $\mathcal{D}_{E^i}^{\text{SubsComp}}$ .

En [114] se proponen diversas alternativas para mejorar la eficiencia del analizador anterior:

#### Análisis con tamaño de la gramática lineal

La complejidad temporal del analizador se puede reducir de  $\mathcal{O}(|G|^2n^3)$  a  $\mathcal{O}(|G|n^3)$  usando las técnicas descritas en [56]. Esta mejora es muy importante, porque  $|G|$  normalmente es mucho mayor que  $n$  en aplicaciones de lenguaje natural.

#### Eliminación de estados equivalentes

Si observamos los pasos deductivos  $\mathcal{D}_{E^i}^\varepsilon$  y  $\mathcal{D}_{E^i}^{\text{Foot}}$ , cambian de estado pero no reconocen ningún nuevo elemento de la cadena de entrada. Por tanto, pueden ser precompiladas anotando los estados equivalentes.

$$[N^\gamma \rightarrow \bullet P^\gamma \nu, i, j] \equiv [N^\gamma \rightarrow P^\gamma \bullet \nu, i, j]$$

si  $(\text{label}(P^\gamma) = \varepsilon \text{ o } P^\gamma = \mathbf{F}^\gamma)$  y  $\neg \exists \beta \mid \beta \in \text{ladj}(N^\gamma)$ ;

$$[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j] \equiv [N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, j]$$

si  $\text{label}(P^\gamma) = \varepsilon \text{ o } P^\gamma = \mathbf{F}^\gamma$ .

Durante el análisis, el analizador salta directamente del primero al último estado en el conjunto de estados equivalentes.

#### Compartición de nodos en una TIG

Una importante característica del analizador es que los hijos de un nodo no tienen que ser únicos y un subárbol no necesita tener solo un padre. Esto quiere decir que un subárbol o un superárbol pueden aparecer en distintos lugares de la gramática. El

único requisito para compartir nodos es que cada posible forma de construir un árbol que sea consistente con las relaciones padre-hijo debe ser un árbol elemental de la gramática. La compartición de nodos puede reducir de forma drástica el tamaño de  $|G|$  y, por consiguiente, el coste computacional de análisis.

#### Análisis de LTIG

El algoritmo puede ser extendido para sacar partido del hecho de que los árboles elementales en las LTIG están lexicalizados. La complejidad teórica en el peor caso del algoritmo permanece en la misma cota, pero se reduce de manera significativa el tamaño de la gramática que se va a considerar cuando se analiza una entrada en particular.

Evidentemente, estas no son las únicas modificaciones que se pueden aplicar a este analizador para mejorar su eficiencia. De hecho, nuestra propuesta va en la línea de usar un filtro basado en la esquina izquierda, como veremos más adelante en el capítulo dedicado a los analizadores para las TIG.

## 7.4 Revisando el analizador tipo Earley para TIG

La representación de los árboles elementales TIG se puede modificar con objeto de aprovechar la definición de analizadores clásicos para CFG. Cuando usamos una representación multicapa, debemos incluir en los analizadores reglas que nos permitan navegar dentro de cada árbol elemental y predicados que garanticen las restricciones de adjunción.

En esta sección vamos a ver una nueva manera de representar los árboles elementales de una TAG y, por tratarse de un subconjunto de las mismas, de una TIG. Esta representación fue introducida en [46] y formalizada en [38], y la denominamos *representación plana de los árboles elementales*. A continuación reescribiremos el analizador de la sección anterior para mostrar que es totalmente equivalente al analizador Earley para CFG [37].

### 7.4.1 Representación plana de los árboles elementales

Inicialmente describiremos una notación alternativa para árboles, la cual hace uso de una representación basada en palabras en lugar de la tradicional representación gráfica. La notación es la siguiente:  $a$  para  $a \in V_T$ ;  $X(t_1 \dots t_n)$  para los árboles elementales que cuya raíz sea  $X$  y sus descendientes directos los subárboles  $t_1 \dots t_n$ . De forma recursiva haremos lo mismo con cada uno de los subárboles  $t_i$ . Cuando  $X$  no tenga hijos, usaremos la notación  $X$  en lugar de  $X()$ .

A continuación, transformaremos de manera trivial la representación de  $X(t_1 \dots t_n)$  en una forma equivalente  $X_L t_1 \dots t_n X_R$  para cada símbolo no terminal  $X$ . En otras palabras, un símbolo no terminal  $X$  se divide en dos nuevos símbolos no terminales,  $X_L$  y  $X_R$ , que dominarán los contextos izquierdo y derecho del símbolo.

Por ejemplo, la representación plana de un árbol inicial con la forma  $S(e)$ , será  $S_L e S_R$ . Y  $S_L e S_L^* S_R^* e S_R$  corresponde a un árbol auxiliar con la forma  $S(e, S^*, e)$  siendo  $S^*$  el nodo pie.

En general, la representación de un árbol auxiliar tendrá la forma  $X_L r_1 X_L^* X_R^* r_2 X_R$ , donde  $r_1$  y  $r_2$  son secuencias de símbolos y  $X$  y  $X^*$  los símbolos de la raíz y el pie respectivamente. Si observamos atentamente, podemos establecer que  $X_L r_1 X_L^*$  es el árbol contextual izquierdo dominado por la raíz con respecto al nodo pie. Análogamente,  $X_R^* r_2 X_R$  será el árbol contextual derecho.

La operación de adjunción también puede dividirse en dos partes con respecto a la espina de un árbol auxiliar. Supongamos que  $\beta$  es un árbol auxiliar con raíz  $X$  y frontera  $w_L X^* w_R$ , donde  $w_L$  y  $w_R$  son secuencias de símbolos. Sea  $\alpha$  un árbol inicial que contiene un nodo  $X$  con frontera  $r_1 w r_2$ , donde  $r_1, w, r_2$  son secuencias de símbolos y  $w$  es la cadena que domina  $X$ . Cuando adjuntamos  $\beta$  en el nodo  $X$  de  $\alpha$ , tendremos la frontera  $r_1 w_L w w_R r_2$ . Podemos ver que  $w_L (w_R)$  es la cadena que genera el árbol contextual izquierdo (derecho) dominado por  $X$  en  $\beta$ .

En resumen, los árboles precedentes se pueden representar usando la notación plana de la siguiente manera:

$$\alpha = S_L r_1 X_L w X_R r_2 S_R$$

$$\beta = X_L w_L X_L^* X_R^* w_R X_R$$

Con estas consideraciones, en general, establecemos las siguientes tres reglas CFG para traducir los árboles elementales:

$$S \rightarrow S_L r_1 X_L w X_R r_2 S_R$$

$$X_L \rightarrow w_L X_L^*$$

$$X_R \rightarrow X_R^* w_R$$

$S$  es la raíz de un árbol inicial  $\alpha$ . La representación plana de un árbol auxiliar  $\beta$  se divide en dos reglas que representan los contextos izquierdo y derecho. Eliminamos la referencia asociada a la raíces de los árboles auxiliares, porque la adjunción en la raíz o pie de un árbol auxiliar generan el mismo árbol.

Como dijimos previamente, en el formalismo TIG solo existen árboles auxiliares izquierdos y derechos. Si el árbol es un auxiliar izquierdo, la regla asociada al contexto derecho solo produce la cadena vacía. Además, no se permite adjuntar un árbol auxiliar derecho ni en la espina ni en ningún nodo a la derecha de la espina de uno izquierdo. Esto quiere decir que la regla asociada al contexto derecho no es necesaria para la definición de un árbol auxiliar izquierdo. Otra restricción que presentan las TIG es la prohibición de adjunción sobre los nodos raíces y pie de los árboles auxiliares, sin embargo, no podemos eliminar el nodo pie porque su presencia es necesaria para llevar a cabo adjunciones simultáneas sobre el mismo

nodo. Realmente solo necesitamos el lado izquierdo del nodo pie para garantizar este tipo de adjunciones. Por consiguiente, los árboles auxiliares izquierdos tendrán la forma:

$$X_L \rightarrow w_L X_L^*$$

Análogamente, los árboles auxiliares derechos se definen con reglas de la forma:

$$X_R \rightarrow X_R^* w_R$$

Los símbolos no terminales marcados para sustitución no pueden ser adjuntados. Estos símbolos no siguen la norma general de dividirse en dos nuevos símbolos no terminales (contextos izquierdo y derecho). De esta forma prevenimos sustituir una regla asociada a un árbol auxiliar.

### 7.4.2 Analizador tipo Earley para TIG con notación plana

Vamos a ver ahora como todas las operaciones del analizador visto en la sección anterior se pueden simplificar usando la representación plana para representar los árboles elementales. Después de examinar todos los pasos, concluiremos que el analizador obtenido es igual al analizador clásico para CFG.

La adjunción izquierda es equivalente a la sustitución en un símbolo no terminal de contexto izquierdo (subíndice  $L$ ). Se puede observar que las restricciones que existen sobre la adjunción izquierda se capturan mediante mecanismos de sustitución. No se permite sustituir una regla asociada a un árbol auxiliar derecho en los símbolos asociados con el contexto derecho de la espina (subíndice  $R$ ) de un árbol auxiliar izquierdo, porque estos símbolos no existen. Luego, la restricción de adjunción de un árbol auxiliar derecho en la espina de uno izquierdo está asegurada. Tampoco se permite sustituir una regla asociada a un árbol auxiliar derecho en los símbolos del contexto derecho (símbolos a la derecha de la espina) de un árbol auxiliar izquierdo, porque estos símbolos no existen. Esto garantiza la restricción de adjunción de un árbol auxiliar derecho en el contexto derecho de un árbol auxiliar izquierdo.

La adjunción derecha es equivalente a la sustitución en un símbolo no terminal de contexto derecho (subíndice  $R$ ). Al igual que en la adjunción izquierda, todas las restricciones sobre la adjunción derecha se mantienen.

Esta representación va a permitir la adjunción en los nodos pies de los árboles auxiliares. Aunque incumpla la restricción de adjunción sobre los nodos pies, es precisamente esta relajación la que permite realizar adjunciones múltiples sobre un mismo nodo. De esta forma, la adjunción simultánea es equivalente a sustituir en símbolos contextuales izquierdos y después en sus equivalentes derechos.

Con respecto a la sustitución, las reglas asociadas con los árboles iniciales son las únicas que pueden ser sustituidas en los nodos marcados para sustitución, ya que las otras reglas representan árboles auxiliares (etiquetadas con contextos izquierdo y derecho).



Después de todas estas consideraciones, vamos a revisar desde la perspectiva de una representación plana cada uno de los pasos deductivos del analizador presentados previamente:

#### Reconocimiento .

Los pasos deductivos  $\mathcal{D}_{E^i}^{Sc}$  y  $\mathcal{D}_{E^i}^{\epsilon}$  se mantienen. El paso  $\mathcal{D}_{E^i}^{Foot}$  ya no es necesario, porque el nodo pie puede ser sustituido en múltiples adjunciones. Este proceso es equivalente a la operación de reconocimiento en el analizador Earley para CFG.

#### Sustitución .

El predicado *subst* ya no es necesario, debido a que la operación de sustitución ya está filtrada. Por tanto, los pasos  $\mathcal{D}_{E^i}^{SubsPred}$  y  $\mathcal{D}_{E^i}^{SubsCmp}$  son equivalentes a las operaciones de predicción y compleción en el analizador Earley para CFG, respectivamente.

#### Predicción/compleción de subárbol .

La navegación a través de los árboles elementales ya no es necesaria, y los pasos  $\mathcal{D}_{E^i}^{Pred}$  y  $\mathcal{D}_{E^i}^{Cmp}$  son equivalentes a las operaciones de predicción y compleción en el analizador Earley para CFG, respectivamente.

#### Adjunciones izquierda y derecha .

Los predicados *ladj* y *radj* ya no son necesarios, ya que los árboles auxiliares han sido etiquetados con nuevos símbolos no terminales. Por tanto, los pasos  $\mathcal{D}_{E^i}^{LAdjPred}$  y  $\mathcal{D}_{E^i}^{RAAdjPred}$  son equivalentes a la operación de predicción en el analizador Earley para CFG, mientras  $\mathcal{D}_{E^i}^{LAdjCmp}$  y  $\mathcal{D}_{E^i}^{RAAdjCmp}$  lo son a la de compleción.

## Capítulo 8

# Algoritmos de análisis sintáctico para TIG

En este capítulo se describen en detalle algunos de los analizadores para TIG presentados en [25] y las relaciones formales existentes entre ellos. El objetivo es crear la espina dorsal del núcleo de una taxonomía de analizadores basados en el algoritmo de Earley, similar a las ya existentes para CFG y TAG.

Como punto de partida definiremos un esquema basado en el conocido algoritmo CYK para CFG. A continuación definiremos un esquema ascendente basado en Earley que nos permita ampliar la clase de gramáticas sobre la que pueda actuar. Y concluiremos el conjunto de esquemas de estrategia ascendente con la presentación de un analizador con recorrido bidireccional de la cadena de entrada al estilo del propuesto por de Vreught y Honig para CFG. Como punto final de esta red de analizadores, introduciremos una variante del esquema basado en el algoritmo tipo Earley para TIG, descrito en el tema anterior, y estableceremos las relaciones formales existentes entre ellos.

Para la definición de todos los esquemas suponemos que nos ceñimos a la definición restringida de las TIG [115], que como máximo permiten una adjunción izquierda y una derecha en un nodo adjuntable.

### 8.1 Esquema tipo CYK

El primer esquema que veremos, que denominaremos  $\mathbf{CYK}^i$ , presenta una estrategia de análisis ascendente con lectura unidireccional, de izquierda a derecha, de la cadena de entrada. Se trata de una extensión del algoritmo CYK definido originalmente para gramáticas independientes del contexto. El esquema  $\mathbf{CYK}^i$  fue introducido en [25] y se basa en el presentado en forma algorítmica para las SLTIG en [115].

El esquema  $\mathbf{CYK}^i$  sólo es aplicable a la clase de gramáticas de inserción de árboles  $\mathcal{CNF}_{TIG}$  cuyos árboles elementales presentan las siguientes restricciones: (i) un nodo in-

terno, salvo el nodo pie, dominará directamente un máximo de dos nodos y (ii) los nodos etiquetados con símbolos terminales, la palabra vacía o el nodo *bottom* no tendrán nodos hermanos.

El dominio del esquema  $\text{CYK}^i$  viene dado por:

$$\mathcal{I}_{\text{CYK}^i} = \{[M^\gamma, i, j, \text{code}]\}$$

donde  $\text{label}(M^\gamma) \in V_N$  y  $\gamma \in \mathbf{I} \cup \mathbf{A}$ . Los índices  $0 \leq i \leq j$  establecen las posiciones dentro de la cadena de entrada que delimitan el fragmento reconocido por  $M^\gamma$ . La función del parámetro  $\text{code} \subseteq \{L, R\}$  es indicar si sobre el nodo  $M^\gamma$  se ha completado una adjunción izquierda y/o derecha o no se ha completado ninguna adjunción, y puede tomar uno de los siguientes valores:

- $\emptyset$  si no se ha completado ninguna adjunción en el nodo  $M^\gamma$ ,
- $\{L\}$  si se ha completado una adjunción de un árbol auxiliar izquierdo en el nodo  $M^\gamma$ ,
- $\{R\}$  si se ha completado una adjunción de un árbol auxiliar derecho en el nodo  $M^\gamma$ ,
- $\{L, R\}$  si se ha completado las adjunciones de un árbol auxiliar izquierdo y otro derecho en el nodo  $M^\gamma$ .

Los pasos deductivos del esquema son:

$$\mathcal{D}_{\text{CYK}^i} = \mathcal{D}_{\text{CYK}^i}^{\text{Scan}} \cup \mathcal{D}_{\text{CYK}^i}^\varepsilon \cup \mathcal{D}_{\text{CYK}^i}^{\text{CompUna}} \cup \mathcal{D}_{\text{CYK}^i}^{\text{CompBin}} \cup \mathcal{D}_{\text{CYK}^i}^{\text{LAdj}} \cup \mathcal{D}_{\text{CYK}^i}^{\text{Radj}} \cup \mathcal{D}_{\text{CYK}^i}^{\text{Subs}}$$

$$\mathcal{D}_{\text{CYK}^i}^{\text{Scan}} = \frac{[a, j, j+1]}{[N^\gamma, j, j+1, \emptyset]} \quad N^\gamma \rightarrow a \in \mathcal{P}(\gamma)$$

$$\mathcal{D}_{\text{CYK}^i}^\varepsilon = \frac{}{[N^\gamma, j, j, \emptyset]} \quad N^\gamma \rightarrow \varepsilon \in \mathcal{P}(\gamma) \text{ o } N^\gamma \rightarrow \perp$$

$$\mathcal{D}_{\text{CYK}^i}^{\text{CompUna}} = \frac{[O^\gamma, i, j, \text{code}]}{[M^\gamma, i, j, \emptyset]} \quad M^\gamma \rightarrow O^\gamma \in \mathcal{P}(\gamma)$$

donde se debe cumplir:

- (i) ( $\text{nil} \in \text{ladj}(O^\gamma)$  y  $L \notin \text{code}$ ) o ( $\beta \in \text{ladj}(O^\gamma)$  y  $L \in \text{code}$ ),  
(ii) ( $\text{nil} \in \text{radj}(O^\gamma)$  y  $R \notin \text{code}$ ) o ( $\beta \in \text{radj}(O^\gamma)$  y  $R \in \text{code}$ ).

$$\mathcal{D}_{\text{CYK}^i}^{\text{CompBin}} = \frac{\begin{matrix} [O_1^\gamma, i, j, \text{code}] \\ [O_2^\gamma, j, k, \text{code}'] \end{matrix}}{[M^\gamma, i, k, \emptyset]} \quad M^\gamma \rightarrow O_1^\gamma O_2^\gamma \in \mathcal{P}(\gamma)$$

donde se debe cumplir:

- (i) ( $\text{nil} \in \text{ladj}(O_1^\gamma)$  y  $L \notin \text{code}$ ) o ( $\beta \in \text{ladj}(O_1^\gamma)$  y  $L \in \text{code}$ ),

- (ii) ( $\text{nil} \in \text{radj}(O_1^\gamma)$  y  $R \notin \text{code}$ ) o ( $\beta \in \text{radj}(O_1^\gamma)$  y  $R \in \text{code}$ ),
- (iii) ( $\text{nil} \in \text{ladj}(O_2^\gamma)$  y  $L \notin \text{code}'$ ) o ( $\beta \in \text{ladj}(O_2^\gamma)$  y  $L \in \text{code}'$ ),
- (iv) ( $\text{nil} \in \text{radj}(O_2^\gamma)$  y  $R \notin \text{code}'$ ) o ( $\beta \in \text{radj}(O_2^\gamma)$  y  $R \in \text{code}'$ ).

$$\mathcal{D}_{\text{CYK}^i}^{\text{LAdj}} = \frac{\begin{array}{l} [P^\beta, i, j, \emptyset] \\ [M^\gamma, j, k, \text{code}] \end{array}}{[M^\gamma, i, k, \{L\} \cup \text{code}]} \quad \begin{array}{l} \text{label}(P^\beta) = \top \\ \beta \in \text{ladj}(M^\gamma) \\ L \notin \text{code} \end{array}$$

$$\mathcal{D}_{\text{CYK}^i}^{\text{RAdj}} = \frac{\begin{array}{l} [P^\beta, j, k, \emptyset] \\ [M^\gamma, i, j, \text{code}] \end{array}}{[M^\gamma, i, k, \{R\} \cup \text{code}]} \quad \begin{array}{l} \text{label}(P^\beta) = \top \\ \beta \in \text{radj}(M^\gamma) \\ R \notin \text{code} \end{array}$$

$$\mathcal{D}_{\text{CYK}^i}^{\text{Subs}} = \frac{[P^\alpha, i, j, \emptyset]}{[M^\gamma, i, j, \emptyset]} \quad \begin{array}{l} \text{label}(P^\alpha) = \top \\ \alpha \in \text{subst}(M^\gamma) \end{array}$$

Los pasos deductivos  $\mathcal{D}_{\text{CYK}^i}^{\text{Scan}}$  y  $\mathcal{D}_{\text{CYK}^i}^\epsilon$  son los que inician el reconocimiento ascendente. También se incluye en este caso los nodos pies de los árboles auxiliares, ya que en los árboles auxiliares TIG no es necesario transmitir la información del subárbol escindido.

Una vez reconocido el subárbol dominado por un nodo, los pasos  $\mathcal{D}_{\text{CYK}^i}^{\text{CompUna}}$  y  $\mathcal{D}_{\text{CYK}^i}^{\text{CompBin}}$  permiten continuar el reconocimiento ascendente. La condición que acompaña a ambos pasos de completación comprueba que el nodo que domina el subárbol que se va a completar no presente adjunción (izquierda y/o derecha) obligatoria y, en el caso de presentarla, que se haya efectuado. Cuando se ha reconocido un árbol auxiliar izquierdo (resp. derecho), el paso  $\mathcal{D}_{\text{CYK}^i}^{\text{LAdj}}$  (resp.  $\mathcal{D}_{\text{CYK}^i}^{\text{RAdj}}$ ) efectúa la adjunción en un nodo de adjunción izquierda (resp. derecha), siempre que el reconocimiento lo haya alcanzado y no haya sido ya adjuntado por la izquierda (resp. derecha).

Por último, la operación  $\mathcal{D}_{\text{CYK}^i}^{\text{Subs}}$  sustituye un árbol inicial que ha sido completamente reconocido en todos los nodos de sustitución donde se puede sustituir dicho árbol.

El conjunto de ítems finales se define como:

$$\mathcal{F}_{\text{CYK}^i} = \{[P^\alpha, 0, n, \emptyset] \mid \alpha \in I, \text{label}(P^\alpha) = \top, \text{label}(R^\alpha) = S\}$$

## 8.2 Esquema tipo Earley ascendente

Este esquema, al que vamos a denominar **buE**<sup>i</sup>, es una adaptación para TIG del esquema **buE** para CFG. Fue presentado en [25] y se puede derivar mediante una generalización del esquema **CYK**<sup>i</sup>. El interés de este esquema radica en que se trata de un reconocedor con estrategia ascendente que elimina la restricción impuesta por el esquema anterior sobre la forma que deben tener los árboles elementales.

Al igual que el esquema **CYK**<sup>i</sup>, no vamos a permitir que se adjunte sobre un nodo hasta que el subárbol que domina haya sido completamente reconocido. Ésto hace que el dominio



del esquema  $\text{buE}^i$  disponga de dos tipos de ítems: los que presentan el punto en el interior de una regla, que no pueden haber sido adjuntados, y los que lo presentan al final, pueden haber sido adjuntados por la izquierda y/o derecha.

$$\mathcal{I}_{\text{buE}^i} = \mathcal{I}_{\text{buE}^i}^{(i)} \cup \mathcal{I}_{\text{buE}^i}^{(ii)}$$

$$\mathcal{I}_{\text{buE}^i}^{(i)} = \{[M^\gamma \rightarrow \delta \bullet \nu, i, j, \emptyset]\}$$

donde  $M^\gamma \rightarrow \delta \nu \in \mathcal{P}(\gamma)$  y  $\gamma \in \mathbf{I} \cup \mathbf{A}$ . Los índices  $0 \leq i \leq j$  establecen las posiciones de la cadena de entrada que delimitan el fragmento reconocido por  $\delta$ . Para evitar que el punto aparezca a la derecha de la producción, se exige que  $\nu \neq \varepsilon$ . El valor  $\emptyset$  indica que no se ha completado ninguna adjunción en el nodo  $M^\gamma$ .

$$\mathcal{I}_{\text{buE}^i}^{(ii)} = \{[M^\gamma \rightarrow \nu \bullet, i, j, \text{code}]\}$$

donde  $M^\gamma \rightarrow \nu \in \mathcal{P}(\gamma)$  y  $\gamma \in \mathbf{I} \cup \mathbf{A}$ . Los índices  $0 \leq i \leq j$  establecen las posiciones de la cadena de entrada que delimitan el fragmento reconocido por  $\nu$ . Y  $\text{code} = \emptyset$  si no se completó ninguna adjunción sobre  $M^\gamma$ ,  $\text{code} = \{L\}$  si se completó una adjunción izquierda sobre  $M^\gamma$ ,  $\text{code} = \{R\}$  si se completó una adjunción derecha sobre  $M^\gamma$  y  $\text{code} = \{L, R\}$  si se completó una adjunción izquierda y otra derecha sobre  $M^\gamma$ .

Los pasos deductivos del esquema son:

$$\mathcal{D}_{\text{buE}^i} = \mathcal{D}_{\text{buE}^i}^{\text{Ini}} \cup \mathcal{D}_{\text{buE}^i}^{\text{Scan}} \cup \mathcal{D}_{\text{buE}^i}^\varepsilon \cup \mathcal{D}_{\text{buE}^i}^{\text{Comp}} \cup \mathcal{D}_{\text{buE}^i}^{\text{LAdj}} \cup \mathcal{D}_{\text{buE}^i}^{\text{Radj}} \cup \mathcal{D}_{\text{buE}^i}^{\text{Subs}}$$

$$\mathcal{D}_{\text{buE}^i}^{\text{Ini}} = \frac{}{[N^\gamma \rightarrow \bullet \nu, i, i, \emptyset]} \quad \gamma \in \mathbf{I} \cup \mathbf{A}$$

$$\mathcal{D}_{\text{buE}^i}^{\text{Scan}} = \frac{[a, j, j+1]}{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \emptyset]} \quad \text{label}(M^\gamma) = a$$

$$\mathcal{D}_{\text{buE}^i}^\varepsilon = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \emptyset]}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, j, \emptyset]} \quad \text{label}(M^\gamma) = \varepsilon \text{ o } \text{label}(M^\gamma) = \perp$$

$$\mathcal{D}_{\text{buE}^i}^{\text{Comp}} = \frac{[M^\gamma \rightarrow \omega \bullet, j, k, \text{code}]}{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \emptyset]} \quad \frac{}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k, \emptyset]}$$

donde se debe cumplir:

- (i) ( $\mathbf{nil} \in \text{ladj}(M^\gamma)$  y  $L \notin \text{code}$ ) o ( $\beta \in \text{ladj}(M^\gamma)$  y  $L \in \text{code}$ ),
- (ii) ( $\mathbf{nil} \in \text{radj}(M^\gamma)$  y  $R \notin \text{code}$ ) o ( $\beta \in \text{radj}(M^\gamma)$  y  $R \in \text{code}$ ).

$$\mathcal{D}_{\text{buE}^i}^{\text{LAdj}} = \frac{[\top \rightarrow \mathbf{R}^\beta \bullet, i, j, \emptyset]}{[M^\gamma \rightarrow \nu \bullet, j, k, \text{code}]} \quad \beta \in \text{ladj}(M^\gamma)$$

$$[M^\gamma \rightarrow \nu \bullet, i, k, \{L\} \cup \text{code}] \quad L \notin \text{code}$$

$$\mathcal{D}_{\text{buE}^i}^{\text{RAdj}} = \frac{[\top \rightarrow \mathbf{R}^\beta \bullet, j, k, \emptyset]}{[M^\gamma \rightarrow \nu \bullet, i, j, \text{code}]} \quad \beta \in \text{radj}(M^\gamma)$$

$$[M^\gamma \rightarrow \nu \bullet, i, k, \{R\} \cup \text{code}] \quad R \notin \text{code}$$

$$\mathcal{D}_{\text{buE}^i}^{\text{Subs}} = \frac{[\top \rightarrow \mathbf{R}^\alpha \bullet, j, k, \emptyset]}{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \emptyset]} \quad \alpha \in \text{subst}(M^\gamma)$$

$$[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k, \emptyset]$$

El paso  $\mathcal{D}_{\text{buE}^i}^{\text{Ini}}$  inicia el reconocimiento desde todos los subárboles de los árboles elementales. El paso  $\mathcal{D}_{\text{buE}^i}^{\text{Scan}}$  reconoce la presencia de un símbolo terminal en la cadena de entrada. Mientras  $\mathcal{D}_{\text{buE}^i}^\varepsilon$  refleja el hecho de que se puede saltar sobre nodos etiquetados con  $\varepsilon$  y nodos pie sin tener que reconocer nada. El paso  $\mathcal{D}_{\text{buE}^i}^{\text{Comp}}$  continúa el reconocimiento ascendente cuando se ha completado el reconocimiento de un subárbol. El resto de pasos funcionan de forma análoga a los homónimos del esquema anterior.

El conjunto de ítems finales es:

$$\mathcal{F}_{\text{buE}^i} = \{[\top \rightarrow \mathbf{R}^\alpha \bullet, 0, n, \emptyset] \mid \alpha \in I, \text{label}(\mathbf{R}^\alpha) = S\}$$

### 8.2.1 Una variante del esquema tipo Earley ascendente

Si estudiamos con atención el esquema  $\text{buE}^i$  podemos observar ciertas ineficiencias, las cuales son debidas esencialmente a que el paso  $\mathcal{D}_{\text{buE}^i}^{\text{Ini}}$  inicia el reconocimiento de todos los subárboles, sin comprobar que la raíz del mismo sea un nodo que presente una restricción de adjunción izquierda obligatoria. Ello provoca un doble problema, el primero es de prestaciones del analizador, ya que puede reconocer subárboles que no son correctos y cuya incorrección no se detecta hasta que se lleva a cabo un paso de completación de subárbol. Y de ésto se deriva el segundo problema, debido a que el paso  $\mathcal{D}_{\text{buE}^i}^{\text{Comp}}$  debe conocer las adjunciones que se han efectuado antes de realizar la completación, lo que obliga al analizador a arrastrar esta información. Si se controla el inicio del reconocimiento de cada subárbol obtendríamos un doble beneficio: (1) aumentaríamos la eficiencia eliminando del reconocimiento subárboles que de partida sabemos que son incorrectos y (2) el parámetro *code* se simplificaría, ya que sólo debe indicar si tuvo lugar una adjunción derecha sobre el nodo.

El esquema que proponemos, al que denominaremos  $\text{buE}^k$ , pretende paliar estos problemas. Se obtiene modificando las condiciones laterales de los pasos  $\mathcal{D}_{\text{buE}^i}^{\text{Ini}}$  y  $\mathcal{D}_{\text{buE}^i}^{\text{Comp}}$ , el paso  $\mathcal{D}_{\text{buE}^i}^{\text{LAdj}}$  y la forma del parámetro *code* del esquema  $\text{buE}^i$ .

El dominio del esquema  $\text{buE}^k$  es:

$$\mathcal{I}_{\text{buE}^k} = \mathcal{I}_{\text{buE}^k}^{(i)} \cup \mathcal{I}_{\text{buE}^k}^{(ii)}$$

$$\mathcal{I}_{\text{buE}^k}^{(i)} = \{[M^\gamma \rightarrow \delta \bullet \nu, i, j, false]\}$$

donde  $M^\gamma \rightarrow \delta \nu \in \mathcal{P}(\gamma)$  y  $\gamma \in \mathbf{I} \cup \mathbf{A}$ . Los índices  $0 \leq i \leq j$  establecen las posiciones de la cadena de entrada que delimitan el fragmento reconocido por  $\delta$ . Para evitar que el punto aparezca a la derecha de la producción, se exige que  $\nu \neq \varepsilon$ . El valor *false* indica que no se ha completado ninguna adjunción en el nodo  $M^\gamma$ .

$$\mathcal{I}_{\text{buE}^k}^{(ii)} = \{[M^\gamma \rightarrow \nu \bullet, i, j, radj]\}$$

donde  $M^\gamma \rightarrow \nu \in \mathcal{P}(\gamma)$  y  $\gamma \in \mathbf{I} \cup \mathbf{A}$ . Los índices  $0 \leq i \leq j$  establecen las posiciones de la cadena de entrada que delimitan el fragmento reconocido por  $\nu$ . Y *radj* = *false* si no se completó ninguna adjunción derecha sobre el nodo  $M^\gamma$  y *radj* = *true* si se completó una adjunción derecha sobre  $M^\gamma$ .

Los pasos deductivos del esquema son:

$$\begin{aligned} \mathcal{D}_{\text{buE}^k} &= \mathcal{D}_{\text{buE}^k}^{\text{Ini}} \cup \mathcal{D}_{\text{buE}^k}^{\text{Scan}} \cup \mathcal{D}_{\text{buE}^k}^\varepsilon \cup \mathcal{D}_{\text{buE}^k}^{\text{Comp}} \cup \mathcal{D}_{\text{buE}^k}^{\text{LAdj}} \cup \mathcal{D}_{\text{buE}^k}^{\text{RAdj}} \cup \mathcal{D}_{\text{buE}^k}^{\text{Subs}} \\ \mathcal{D}_{\text{buE}^k}^{\text{Ini}} &= \frac{}{[N^\gamma \rightarrow \bullet \nu, i, i, false]} \quad \begin{array}{l} \gamma \in \mathbf{I} \cup \mathbf{A} \\ \text{nil} \in \text{ladj}(N^\gamma) \end{array} \\ \mathcal{D}_{\text{buE}^k}^{\text{Scan}} &= \frac{\begin{array}{l} [a, j, j+1] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false] \end{array}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, j+1, false]} \quad \text{label}(M^\gamma) = a \\ \mathcal{D}_{\text{buE}^k}^\varepsilon &= \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, j, false]} \quad \text{label}(M^\gamma) = \varepsilon \text{ o } \text{label}(M^\gamma) = \perp \\ \mathcal{D}_{\text{buE}^k}^{\text{Comp}} &= \frac{\begin{array}{l} [M^\gamma \rightarrow \nu \bullet, j, k, radj] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false] \end{array}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k, false]} \end{aligned}$$

donde se debe cumplir que si *radj* = *false* entonces  $\text{nil} \in \text{radj}(M^\gamma)$ .

$$\begin{aligned} \mathcal{D}_{\text{buE}^k}^{\text{LAdj}} &= \frac{[\top \rightarrow \mathbf{R}^\beta \bullet, i, j, false]}{[M^\gamma \rightarrow \bullet \nu, i, j, false]} \quad \beta \in \text{ladj}(M^\gamma) \\ \mathcal{D}_{\text{buE}^k}^{\text{RAdj}} &= \frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\beta \bullet, j, k, false] \\ [M^\gamma \rightarrow \nu \bullet, i, j, false] \end{array}}{[M^\gamma \rightarrow \nu \bullet, i, k, true]} \quad \beta \in \text{radj}(M^\gamma) \\ \mathcal{D}_{\text{buE}^k}^{\text{Subs}} &= \frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\alpha \bullet, j, k, false] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false] \end{array}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k, false]} \quad \alpha \in \text{subst}(M^\gamma) \end{aligned}$$

El conjunto de ítems finales se define mediante:

$$\mathcal{F}_{\text{buE}^k} = \{[\top \rightarrow \mathbf{R}^\alpha \bullet, 0, n, false] \mid \alpha \in \mathbf{I}, \text{label}(\mathbf{R}^\alpha) = S\}$$

### 8.3 Esquema tipo de Vreught y Honig

En esta sección definimos el esquema  $\mathbf{dVH}^i$  con estrategia ascendente y recorrido bidireccional de la cadena de entrada, que está basado en el algoritmo para gramáticas independientes del contexto definido por De Vreught y Honig. La ventaja fundamental que ofrece este esquema es que permite obtener mayor información parcial para entradas incorrectas, lo cual es muy ventajoso para su empleo como analizador sintáctico fragmental.

Modificamos la forma general de regla punteada usada en los esquemas de tipo Earley e introducimos un punto adicional para delimitar la parte reconocida dentro de una regla. De nuevo no permitiremos que se adjunte sobre un nodo hasta que el subárbol que domina haya sido completamente reconocido. Ésto hace que el dominio del esquema  $\mathbf{buE}^i$  se divida en ítems que representan el reconocimiento parcial o completo de la regla.

$$\mathcal{I}_{\mathbf{dVH}^i} = \mathcal{I}_{\mathbf{dVH}^i}^{(i)} \cup \mathcal{I}_{\mathbf{dVH}^i}^{(ii)}$$

$$\mathcal{I}_{\mathbf{dVH}^i}^{(i)} = \{[M^\gamma \rightarrow \nu \bullet \delta \bullet \omega, i, j, \emptyset]\}$$

donde  $M^\gamma \rightarrow \nu \delta \omega \in \mathcal{P}(\gamma)$  y  $\gamma \in \mathbf{I} \cup \mathbf{A}$ . Los índices  $0 \leq i \leq j$  establecen las posiciones de la cadena de entrada que delimitan el fragmento reconocido por  $\delta$ . Para evitar el reconocimiento completo de la regla se requiere que  $\nu \neq \varepsilon$  o  $\omega \neq \varepsilon$ . El valor  $\emptyset$  indica que no se ha completado ninguna adjunción en el nodo  $M^\gamma$ .

$$\mathcal{I}_{\mathbf{dVH}^i}^{(ii)} = \{[M^\gamma \rightarrow \bullet \delta \bullet, i, j, code]\}$$

donde  $M^\gamma \rightarrow \delta \in \mathcal{P}(\gamma)$  y  $\gamma \in \mathbf{I} \cup \mathbf{A}$ . Los índices  $0 \leq i \leq j$  establecen las posiciones de la cadena de entrada que delimitan el fragmento reconocido por  $\delta$ . Y  $code = \emptyset$  si no se completó ninguna adjunción sobre  $M^\gamma$ ,  $code = \{L\}$  si se completó una adjunción izquierda sobre  $M^\gamma$ ,  $code = \{R\}$  si se completó una adjunción derecha sobre  $M^\gamma$  y  $code = \{L, R\}$  si se completó una adjunción izquierda y otra derecha sobre  $M^\gamma$ .

El conjunto de pasos deductivos viene dado por:

$$\mathcal{D}_{\mathbf{dVH}^i} = \mathcal{D}_{\mathbf{dVH}^i}^{\text{Scan}} \cup \mathcal{D}_{\mathbf{dVH}^i}^\varepsilon \cup \mathcal{D}_{\mathbf{dVH}^i}^{\text{Comp}} \cup \mathcal{D}_{\mathbf{dVH}^i}^{\text{Con}} \cup \mathcal{D}_{\mathbf{dVH}^i}^{\text{LAdj}} \cup \mathcal{D}_{\mathbf{dVH}^i}^{\text{RAAdj}} \cup \mathcal{D}_{\mathbf{dVH}^i}^{\text{Subs}}$$

$$\mathcal{D}_{\mathbf{dVH}^i}^{\text{Scan}} = \frac{[a, j, j+1]}{[N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, j, j+1, \emptyset]} \quad \text{label}(M^\gamma) = a$$

$$\mathcal{D}_{\mathbf{dVH}^i}^\varepsilon = \frac{}{[N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, j, j, \emptyset]} \quad \text{label}(M^\gamma) = \varepsilon \text{ o } \text{label}(M^\gamma) = \perp$$

$$\mathcal{D}_{\mathbf{dVH}^i}^{\text{Comp}} = \frac{[M^\gamma \rightarrow \bullet \delta \bullet, i, j, code]}{[N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, i, j, \emptyset]}$$

donde se debe cumplir:

- (i) ( $\mathbf{nil} \in \text{ladj}(M^\gamma)$  y  $L \notin \text{code}$ ) o ( $\beta \in \text{ladj}(M^\gamma)$  y  $L \in \text{code}$ ),
- (ii) ( $\mathbf{nil} \in \text{radj}(M^\gamma)$  y  $R \notin \text{code}$ ) o ( $\beta \in \text{radj}(M^\gamma)$  y  $R \in \text{code}$ ).

$$\mathcal{D}_{\text{dVH}^i}^{\text{Con}} = \frac{\begin{array}{l} [N^\gamma \rightarrow \nu \bullet \delta \bullet \delta' \omega, i, j', \emptyset] \\ [N^\gamma \rightarrow \nu \delta \bullet \delta' \bullet \omega, j', j, \emptyset] \end{array}}{[N^\gamma \rightarrow \nu \bullet \delta \delta' \bullet \omega, i, j, \emptyset]}$$

$$\mathcal{D}_{\text{dVH}^i}^{\text{LAdj}} = \frac{\begin{array}{l} [\top \rightarrow \bullet \mathbf{R}^\beta \bullet, i, j, \emptyset] \\ [M^\gamma \rightarrow \bullet \delta \bullet, j, k, \text{code}] \end{array}}{[M^\gamma \rightarrow \bullet \delta \bullet, i, k, \{L\} \cup \text{code}]} \quad \begin{array}{l} \beta \in \text{ladj}(M^\gamma) \\ L \notin \text{code} \end{array}$$

$$\mathcal{D}_{\text{dVH}^i}^{\text{RAdj}} = \frac{\begin{array}{l} [\top \rightarrow \bullet \mathbf{R}^\beta \bullet, j, k, \emptyset] \\ [M^\gamma \rightarrow \bullet \delta \bullet, i, j, \text{code}] \end{array}}{[M^\gamma \rightarrow \bullet \delta \bullet, i, k, \{R\} \cup \text{code}]} \quad \begin{array}{l} \beta \in \text{radj}(M^\gamma) \\ R \notin \text{code} \end{array}$$

$$\mathcal{D}_{\text{dVH}^i}^{\text{Subs}} = \frac{[\top \rightarrow \bullet \mathbf{R}^\alpha \bullet, j, k, \emptyset]}{[N^\gamma \rightarrow \delta \bullet M^\gamma \bullet \nu, j, k, \emptyset]} \quad \alpha \in \text{subst}(M^\gamma)$$

Los pasos deductivos  $\mathcal{D}_{\text{dVH}^i}^{\text{Scan}}$  y  $\mathcal{D}_{\text{dVH}^i}^\epsilon$  son los que inician el reconocimiento ascendente desde los nodos etiquetados con símbolos terminales que coincida con algún símbolo de la cadena de entrada y nodos etiquetados con la palabra vacía o nodos *bottom*, respectivamente.

La función del paso  $\mathcal{D}_{\text{dVH}^i}^{\text{Comp}}$  es continuar el reconocimiento del superárbol respecto a un nodo una vez que el subárbol dominado por él ha sido completamente reconocido. Se acompaña una condición lateral a este paso para garantizar que el nodo que domina el subárbol no presente adjunción obligatoria y, en el caso de presentarla, que se haya efectuado.

El paso  $\mathcal{D}_{\text{dVH}^i}^{\text{Con}}$  concatena dos fragmentos adyacentes en una regla que reconocen segmentos colindantes en la cadena de entrada. Obsérvese como en ambos antecedentes el componente *adj* se encuentra a *false*, ésto es debido a que ninguno de ellos representa el reconocimiento completo de un subárbol dominado por un nodo, y tal como se establece en la definición del dominio de este esquema, ésta es una condición necesaria para completar una adjunción en un nodo.

Los pasos deductivos  $\mathcal{D}_{\text{dVH}^i}^{\text{LAdj}}$  y  $\mathcal{D}_{\text{dVH}^i}^{\text{RAdj}}$  funcionan de forma análoga a sus homónimos en el esquema  $\mathbf{buE}^i$ .

La función del paso  $\mathcal{D}_{\text{dVH}^i}^{\text{Subs}}$  es continuar el reconocimiento del superárbol respecto a un nodo de sustitución una vez completado el análisis de un árbol inicial que se pueda sustituir en él.

El conjunto de ítems finales se define mediante:

$$\mathcal{F}_{\text{dVH}^i} = \{[\top \rightarrow \bullet \mathbf{R}^\alpha \bullet, 0, n, \emptyset] \mid \alpha \in I, \text{label}(\mathbf{R}^\alpha) = S\}$$

Si somos totalmente minuciosos, los pasos  $\mathcal{D}_{\text{dVH}^i}^{\text{LAdj}}$  y  $\mathcal{D}_{\text{dVH}^i}^{\text{RAdj}}$  no están formulados desde una perspectiva totalmente bidireccional, ya que una adjunción izquierda y/o derecha no

se efectúa sobre un nodo hasta que no se completa el subárbol que domina el mismo. Sin embargo, este tipo de algoritmo nos permite expandir mediante una adjunción izquierda (resp. derecha) una regla cuando se haya reconocido un prefijo (resp. sufijo) de la misma, es decir, el punto izquierdo (resp. derecho) que delimita la parte reconocida está al comienzo (resp. final) de la regla. Teniendo esto en cuenta, podríamos plantear un nuevo esquema  $dVH^k$  cuyas reglas sean iguales a las del esquema anterior, excepto las dos mencionadas.

El dominio del esquema  $dVH^k$  se ve alterado respecto al de  $dVH^i$  por el hecho de que las adjunciones sobre un nodo se pueden efectuar sin tener que reconocer la regla completa, por lo que es necesario introducir nuevos subdominios en el esquema anterior para distinguir posibles adjunciones izquierdas o derechas. El conjunto de ítems para el nuevo esquema se define mediante:

$$\mathcal{I}_{dVH^k} = \mathcal{I}_{dVH^k}^{(i)} \cup \mathcal{I}_{dVH^k}^{(ii)} \cup \mathcal{I}_{dVH^k}^{(iii)} \cup \mathcal{I}_{dVH^k}^{(iv)}$$

$$\mathcal{I}_{dVH^k}^{(ii)} = \mathcal{I}_{dVH^i}^{(ii)}$$

$$\mathcal{I}_{dVH^k}^{(i)} = \{[M^\gamma \rightarrow \nu \bullet \delta \bullet \omega, i, j, \emptyset]\}$$

donde  $M^\gamma \rightarrow \nu\delta\omega \in \mathcal{P}(\gamma)$  y  $\gamma \in I \cup A$ . Los índices  $0 \leq i \leq j$  establecen las posiciones de la cadena de entrada que delimitan el fragmento reconocido por  $\delta$ . Para evitar el reconocimiento completo de la regla se requiere que  $\nu \neq \varepsilon$  y  $\omega \neq \varepsilon$ . El valor  $\emptyset$  indica que no se ha completado ninguna adjunción en el nodo  $M^\gamma$ .

$$\mathcal{I}_{dVH^k}^{(iii)} = \{[M^\gamma \rightarrow \bullet \delta \bullet \nu, i, j, code]\}$$

donde  $M^\gamma \rightarrow \delta\nu \in \mathcal{P}(\gamma)$  y  $\gamma \in I \cup A$ . Los índices  $0 \leq i \leq j$  establecen las posiciones de la cadena de entrada que delimitan el fragmento reconocido por  $\delta$ . Para evitar el reconocimiento completo de la regla se requiere que  $\nu \neq \varepsilon$ . Y  $code = \emptyset$  si no se completó ninguna adjunción sobre  $M^\gamma$  y  $code = \{L\}$  si se completó una adjunción izquierda sobre  $M^\gamma$ .

$$\mathcal{I}_{dVH^k}^{(iv)} = \{[M^\gamma \rightarrow \nu \bullet \delta \bullet, i, j, code]\}$$

donde  $M^\gamma \rightarrow \nu\delta \in \mathcal{P}(\gamma)$  y  $\gamma \in I \cup A$ . Los índices  $0 \leq i \leq j$  establecen las posiciones de la cadena de entrada que delimitan el fragmento reconocido por  $\delta$ . Para evitar el reconocimiento completo de la regla se requiere que  $\nu \neq \varepsilon$ . Y  $code = \emptyset$  si no se completó ninguna adjunción sobre  $M^\gamma$  y  $code = \{R\}$  si se completó una adjunción derecha sobre  $M^\gamma$ .

El conjunto de pasos deductivos viene dado por:

$$\mathcal{D}_{dVH^k} = \mathcal{D}_{dVH^k}^{Scan} \cup \mathcal{D}_{dVH^k}^\varepsilon \cup \mathcal{D}_{dVH^k}^{Comp} \cup \mathcal{D}_{dVH^k}^{Con} \cup \mathcal{D}_{dVH^k}^{LAdj} \cup \mathcal{D}_{dVH^k}^{RAdj} \cup \mathcal{D}_{dVH^k}^{Subs}$$

$$\mathcal{D}_{dVH^k}^{Scan} = \mathcal{D}_{dVH^i}^{Scan}$$

$$\mathcal{D}_{dVH^k}^\varepsilon = \mathcal{D}_{dVH^i}^\varepsilon$$

$$\mathcal{D}_{dVH^k}^{\text{Comp}} = \mathcal{D}_{dVH^i}^{\text{Comp}}$$

$$\mathcal{D}_{dVH^k}^{\text{Subs}} = \mathcal{D}_{dVH^i}^{\text{Subs}}$$

$$\mathcal{D}_{dVH^k}^{\text{Con}} = \frac{\begin{array}{l} [N^\gamma \rightarrow \nu \bullet \delta \bullet \delta' \omega, i, j', \text{code}] \\ [N^\gamma \rightarrow \nu \delta \bullet \delta' \bullet \omega, j', j, \text{code}'] \end{array}}{[N^\gamma \rightarrow \nu \bullet \delta \delta' \bullet \omega, i, j, \text{code} \cup \text{code}']}$$

donde  $\text{code} \subseteq \{L\}$  y  $\text{code}' \subseteq \{R\}$ .

$$\mathcal{D}_{dVH^k}^{\text{LAdj}} = \frac{\begin{array}{l} [\top \rightarrow \bullet R^\beta \bullet, i, j, \emptyset] \\ [M^\gamma \rightarrow \bullet \delta \bullet \nu, j, k, \text{code}] \end{array}}{[M^\gamma \rightarrow \bullet \delta \bullet \nu, i, k, \{L\} \cup \text{code}]} \quad \begin{array}{l} \beta \in \text{ladj}(M^\gamma) \\ L \notin \text{code} \end{array}$$

$$\mathcal{D}_{dVH^k}^{\text{RAdj}} = \frac{\begin{array}{l} [\top \rightarrow \bullet R^\beta \bullet, j, k, \emptyset] \\ [M^\gamma \rightarrow \nu \bullet \delta \bullet, i, j, \text{code}] \end{array}}{[M^\gamma \rightarrow \nu \bullet \delta \bullet, i, k, \{R\} \cup \text{code}]} \quad \begin{array}{l} \beta \in \text{radj}(M^\gamma) \\ R \notin \text{code} \end{array}$$

Y su conjunto de ítems finales también es igual al del esquema  $dVH^i$ :

$$\mathcal{F}_{dVH^k} = \mathcal{F}_{dVH^i}$$

## 8.4 Esquema tipo Earley

Aquí presentamos un algoritmo de análisis unidireccional para TIG que combina predicciones descendentes con reconocimiento ascendente, al estilo del algoritmo de Earley para CFG. Una variante de este esquema, al que denominaremos **Earley<sup>i</sup>**, la introducimos en [25] y se obtiene mediante una adaptación del esquema **E<sup>i</sup>** a la restricción impuesta en este capítulo que obliga a que como máximo se efectúe una adjunción izquierda y otra derecha sobre un nodo. Este algoritmo satisface la propiedad del prefijo válido.

El dominio del esquema **Earley<sup>i</sup>** es igual al del esquema **buE<sup>k</sup>** :

$$\mathcal{I}_{\text{Earley}^i} = \mathcal{I}_{\text{buE}^k}$$

El conjunto de pasos deductivos del esquema viene dado por:

$$\begin{aligned} \mathcal{D}_{\text{Earley}^i} = & \mathcal{D}_{\text{Earley}^i}^{\text{Ini}} \cup \mathcal{D}_{\text{Earley}^i}^{\text{Scan}} \cup \mathcal{D}_{\text{Earley}^i}^\varepsilon \cup \mathcal{D}_{\text{Earley}^i}^{\text{Pred}} \cup \mathcal{D}_{\text{Earley}^i}^{\text{Comp}} \cup \mathcal{D}_{\text{Earley}^i}^{\text{LAdjPred}} \cup \\ & \mathcal{D}_{\text{Earley}^i}^{\text{LAdjComp}} \cup \mathcal{D}_{\text{Earley}^i}^{\text{RAdjPred}} \cup \mathcal{D}_{\text{Earley}^i}^{\text{RAdjComp}} \cup \mathcal{D}_{\text{Earley}^i}^{\text{SubsPred}} \cup \mathcal{D}_{\text{Earley}^i}^{\text{SubsComp}} \end{aligned}$$

**Inicio**

El reconocimiento comienza con la predicción de todo árbol inicial cuya raíz sea el axioma.

$$\mathcal{D}_{\text{Earley}^i}^{\text{Ini}} = \frac{[\top \rightarrow \bullet \mathbf{R}^\alpha, 0, 0, \text{false}]}{\text{label}(\mathbf{R}^\alpha) = S} \quad \alpha \in I$$

### Reconocimiento

Los pasos deductivos de reconocimiento son iguales a los del esquema  $\text{buE}^k$ :

$$\mathcal{D}_{\text{Earley}^i}^{\text{Scan}} = \mathcal{D}_{\text{buE}^k}^{\text{Scan}}$$

$$\mathcal{D}_{\text{Earley}^i}^{\varepsilon} = \mathcal{D}_{\text{buE}^k}^{\varepsilon}$$

Los pasos deductivos que establecen la estrategia ascendente predictiva del algoritmo de Earley son los correspondientes a predicciones y compleciones. Para el caso de las TIG, vamos a distinguir cuatro tipos de predicciones con sus correspondientes pasos de compleción asociados: subárbol, adjunción izquierda, adjunción derecha y sustitución.

### Predicción de subárbol

Si el reconocimiento alcanza un nodo que no presenta adjunción izquierda obligatoria, el análisis debe continuar el reconocimiento descendente del subárbol dominado por dicho nodo.

$$\mathcal{D}_{\text{Earley}^i}^{\text{Pred}} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{false}]}{[M^\gamma \rightarrow \bullet \nu, j, j, \text{false}]} \quad \text{nil} \in \text{ladj}(M^\gamma)$$

### Compleción de subárbol

Este paso deductivo de compleción de subárbol es igual al del esquema  $\text{buE}^k$ :

$$\mathcal{D}_{\text{Earley}^i}^{\text{Comp}} = \mathcal{D}_{\text{buE}^k}^{\text{Comp}}$$

### Predicción de adjunción izquierda

Cuando el reconocimiento alcanza un nodo adjuntable por la izquierda, el análisis debe lanzar el reconocimiento de todos los árboles auxiliares que se pueden adjuntar en él.

$$\mathcal{D}_{\text{Earley}^i}^{\text{LAdjPred}} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{false}]}{[\top \rightarrow \bullet \mathbf{R}^\beta, j, j, \text{false}]} \quad \beta \in \text{ladj}(M^\gamma)$$

### Compleción de adjunción izquierda

Una vez que se ha completado el reconocimiento de un árbol auxiliar izquierdo, debemos continuar el reconocimiento del árbol donde se ha efectuado la adjunción.

$$\mathcal{D}_{\text{Earley}^i}^{\text{LAdjComp}} = \frac{[\top \rightarrow \mathbf{R}^\beta \bullet, j, k, \text{false}]}{[M^\gamma \rightarrow \bullet \nu, j, k, \text{false}]} \quad \beta \in \text{ladj}(M^\gamma)$$



### Predicción de adjunción derecha

Cuando se completa el reconocimiento de un subárbol dominado por un nodo adjuntable por la derecha, el análisis debe lanzar el reconocimiento de todos los árboles auxiliares que se pueden adjuntar en él.

$$\mathcal{D}_{\text{Earley}^i}^{\text{RAjPred}} = \frac{[M^\gamma \rightarrow \nu \bullet, i, j, false]}{[\top \rightarrow \bullet \mathbf{R}^\beta, j, j, false]} \quad \beta \in \text{radj}(M^\gamma)$$

### Compleción de adjunción derecha

El paso deductivo de completación de adjunción derecha es igual al del esquema  $\text{buE}^k$ :

$$\mathcal{D}_{\text{Earley}^i}^{\text{RAjComp}} = \mathcal{D}_{\text{buE}^k}^{\text{RAj}}$$

### Predicción de sustitución

Cuando el reconocimiento alcanza un nodo marcado para sustitución, el análisis debe lanzar el reconocimiento de todos los árboles iniciales que se pueden sustituir en él:

$$\mathcal{D}_{\text{Earley}^i}^{\text{SubsPred}} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[\top \rightarrow \bullet \mathbf{R}^\alpha, j, j, false]} \quad \alpha \in \text{subst}(M^\gamma)$$

### Compleción de sustitución

El paso deductivo de completación de sustitución es igual al del esquema  $\text{buE}^k$ :

$$\mathcal{D}_{\text{Earley}^i}^{\text{SubsComp}} = \mathcal{D}_{\text{buE}^k}^{\text{Subs}}$$

El conjunto de ítems finales es igual al del esquema  $\text{buE}^k$ :

$$\mathcal{F}_{\text{Earley}^i} = \mathcal{F}_{\text{buE}^k}$$

## 8.5 Red de analizadores para TIG

En esta sección vamos a establecer las relaciones formales entre los distintos esquemas de análisis que hemos presentado en este capítulo.

**Teorema 8.1** *Relación entre los esquemas  $\text{CYK}^i$  y  $\text{buE}^i$*

*Se mantienen las siguientes relaciones entre esquemas:*

$$\text{CYK}^i \xrightarrow{\text{sc}} \text{CYK}_1^i \xrightarrow{\text{ir}} \text{CYK}_2^i \xrightarrow{\text{sr}} \text{ECYK}^i \xrightarrow{\text{ext}} \text{buE}^i$$

*Prueba*

*El esquema  $\text{CYK}_1^i$  se obtiene desdoblado el paso de sustitución del esquema  $\text{CYK}^i$  en cuatro.*

El dominio del nuevo esquema es igual al del  $CYK^i$ :

$$\mathcal{I}_{CYK_1^i} = \mathcal{I}_{CYK^i}$$

El conjunto de pasos deductivos es:

$$\mathcal{D}_{CYK_1^i} = \mathcal{D}_{CYK_1^i}^{Scan} \cup \mathcal{D}_{CYK_1^i}^\epsilon \cup \mathcal{D}_{CYK_1^i}^{CompUna} \cup \mathcal{D}_{CYK_1^i}^{CompBin} \cup \mathcal{D}_{CYK_1^i}^{LAdj} \cup \mathcal{D}_{CYK_1^i}^{Radj} \cup \\ \mathcal{D}_{CYK_1^i}^{SubsUna} \cup \mathcal{D}_{CYK_1^i}^{SubsBin_1} \cup \mathcal{D}_{CYK_1^i}^{SubsBin_2} \cup \mathcal{D}_{CYK_1^i}^{SubsBin_3}$$

$$\mathcal{D}_{CYK_1^i}^{Scan} = \mathcal{D}_{CYK^i}^{Scan}$$

$$\mathcal{D}_{CYK_1^i}^\epsilon = \mathcal{D}_{CYK^i}^\epsilon$$

$$\mathcal{D}_{CYK_1^i}^{CompUna} = \mathcal{D}_{CYK^i}^{CompUna}$$

$$\mathcal{D}_{CYK_1^i}^{CompBin} = \mathcal{D}_{CYK^i}^{CompBin}$$

$$\mathcal{D}_{CYK_1^i}^{LAdj} = \mathcal{D}_{CYK^i}^{LAdj}$$

$$\mathcal{D}_{CYK_1^i}^{Radj} = \mathcal{D}_{CYK^i}^{Radj}$$

$$\mathcal{D}_{CYK_1^i}^{SubsUna} = \frac{[P^\alpha, i, j, \emptyset]}{[M^\gamma, i, j, \emptyset]} \quad \begin{array}{l} label(P^\alpha) = \top \\ M^\gamma \rightarrow O^\gamma \in \mathcal{P}(\gamma) \\ \alpha \in \text{subst}(O^\gamma) \end{array}$$

$$\mathcal{D}_{CYK_1^i}^{SubsBin_1} = \frac{[P^\alpha, i, j, \emptyset]}{[M^\gamma, i, k, \emptyset]} \quad \begin{array}{l} label(P^\alpha) = \top \\ M^\gamma \rightarrow O_1^\gamma O_2^\gamma \in \mathcal{P}(\gamma) \\ \alpha \in \text{subst}(O_1^\gamma) \end{array}$$

donde se debe cumplir:

- (i) ( $\text{nil} \in \text{ladj}(O_2^\gamma)$  y  $L \notin \text{code}$ ) o ( $\beta \in \text{ladj}(O_2^\gamma)$  y  $L \in \text{code}$ ),
- (ii) ( $\text{nil} \in \text{radj}(O_2^\gamma)$  y  $R \notin \text{code}$ ) o ( $\beta \in \text{radj}(O_2^\gamma)$  y  $R \in \text{code}$ ).

$$\mathcal{D}_{CYK_1^i}^{SubsBin_2} = \frac{[O_1^\gamma, i, j, \text{code}]}{[M^\gamma, i, k, \emptyset]} \quad \begin{array}{l} label(P^\alpha) = \top \\ M^\gamma \rightarrow O_1^\gamma O_2^\gamma \in \mathcal{P}(\gamma) \\ \alpha \in \text{subst}(O_2^\gamma) \end{array}$$

donde se debe cumplir:

- (i) ( $\text{nil} \in \text{ladj}(O_1^\gamma)$  y  $L \notin \text{code}$ ) o ( $\beta \in \text{ladj}(O_1^\gamma)$  y  $L \in \text{code}$ ),
- (ii) ( $\text{nil} \in \text{radj}(O_1^\gamma)$  y  $R \notin \text{code}$ ) o ( $\beta \in \text{radj}(O_1^\gamma)$  y  $R \in \text{code}$ ).

$$\mathcal{D}_{CYK_1^i}^{SubsBin_3} = \frac{[P^\alpha, i, j, \emptyset]}{[M^\gamma, i, k, \emptyset]} \quad \begin{array}{l} label(P^\alpha) = \top \\ label(P^{\alpha 1}) = \top \\ M^\gamma \rightarrow O_1^\gamma O_2^\gamma \in \mathcal{P}(\gamma) \\ \alpha \in \text{subst}(O_1^\gamma) \\ \alpha 1 \in \text{subst}(O_2^\gamma) \end{array}$$

El conjunto de ítems finales es idéntico al del esquema  $\text{CYK}^i$ :

$$\mathcal{F}_{\text{CYK}_1^i} = \mathcal{F}_{\text{CYK}^i}$$

Para demostrar que  $\text{CYK}^i \xrightarrow{\text{sc}} \text{CYK}_1^i$  hay que probar que:

1.  $\mathcal{I}_{\text{CYK}_1^i} \subseteq \mathcal{I}_{\text{CYK}^i}$
2.  $\vdash_{\text{CYK}_1^i}^* \subseteq \vdash_{\text{CYK}^i}^*$

Lo primero es cierto por definición, ya que los dominios de ambos esquemas son iguales:

$$\mathcal{I}_{\text{CYK}_1^i} = \mathcal{I}_{\text{CYK}^i}$$

Para demostrar que  $\vdash_{\text{CYK}_1^i}^* \subseteq \vdash_{\text{CYK}^i}^*$  nos basta con probar que  $\mathcal{D}_{\text{CYK}_1^i} \subseteq \mathcal{D}_{\text{CYK}^i}$ . Vamos a ver sólo los pasos de sustitución, el resto de pasos son idénticos en ambos esquemas:

- Un paso deductivo  $\mathcal{D}_{\text{CYK}_1^i}^{\text{SubsUna}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{CYK}^i}^{\text{Subs}}$

$$\frac{[P^\alpha, i, j, \emptyset]}{[O^\gamma, i, j, \emptyset]}$$

seguido de otro  $\mathcal{D}_{\text{CYK}^i}^{\text{CompUna}}$

$$\frac{[O^\gamma, i, j, \text{code}]}{[M^\gamma, i, j, \emptyset]}$$

- Un paso deductivo  $\mathcal{D}_{\text{CYK}_1^i}^{\text{SubsBin}_1}$  es equivalente a una secuencia formada por un paso  $\mathcal{D}_{\text{CYK}^i}^{\text{Subs}}$

$$\frac{[P^\alpha, i, j, \emptyset]}{[O_1^\gamma, i, j, \emptyset]}$$

seguido de otro  $\mathcal{D}_{\text{CYK}^i}^{\text{CompBin}}$

$$\frac{\begin{array}{l} [O_1^\gamma, i, j, \emptyset] \\ [O_2^\gamma, j, k, \text{code}] \end{array}}{[M^\gamma, i, k, \emptyset]}$$

- Un paso deductivo  $\mathcal{D}_{\text{CYK}_1^i}^{\text{SubsBin}_2}$  es equivalente a una secuencia formada por un paso  $\mathcal{D}_{\text{CYK}^i}^{\text{Subs}}$

$$\frac{[P^\alpha, j, k, \emptyset]}{[O_2^\gamma, j, k, \emptyset]}$$

seguido de otro  $\mathcal{D}_{\text{CYK}^i}^{\text{CompBin}}$

$$\frac{\begin{array}{l} [O_1^\gamma, i, j, \text{code}] \\ [O_2^\gamma, j, k, \emptyset] \end{array}}{[M^\gamma, i, k, \emptyset]}$$

- Un paso deductivo  $\mathcal{D}_{\text{CYK}_1^i}^{\text{SubsBin}_3}$  es equivalente a una secuencia formada por dos pasos

$$\mathcal{D}_{\text{CYK}_1^i}^{\text{Subs}}$$

$$\frac{[P^\alpha, i, j, \emptyset]}{[O_1^\gamma, i, j, \emptyset]} \\ \frac{[P^{\alpha 1}, j, k, \emptyset]}{[O_2^\gamma, j, k, \emptyset]}$$

seguidos de otro  $\mathcal{D}_{\text{CYK}_1^i}^{\text{CompBin}}$

$$\frac{[O_1^\gamma, i, j, \emptyset]}{[O_2^\gamma, j, k, \emptyset]} \\ \frac{[O_2^\gamma, j, k, \emptyset]}{[M^\gamma, i, k, \emptyset]}$$

El esquema  $\text{CYK}_2^i$  se obtiene incluyendo reglas de producción en los ítems del esquema  $\text{CYK}_1^i$ .

El dominio de  $\text{CYK}_2^i$  se define mediante:

$$\mathcal{I}_{\text{CYK}_2^i} = \{[M^\gamma \rightarrow \nu \bullet, i, j, \text{code}] \mid M^\gamma \rightarrow \nu \in \mathcal{P}(\gamma), \gamma \in I \cup A, \\ 0 \leq i \leq j, \text{code} \subseteq \{L, R\}\}$$

El conjunto de pasos deductivos es:

$$\mathcal{D}_{\text{CYK}_2^i} = \mathcal{D}_{\text{CYK}_2^i}^{\text{Scan}} \cup \mathcal{D}_{\text{CYK}_2^i}^\varepsilon \cup \mathcal{D}_{\text{CYK}_2^i}^{\text{CompUna}} \cup \mathcal{D}_{\text{CYK}_2^i}^{\text{CompBin}} \cup \mathcal{D}_{\text{CYK}_2^i}^{\text{LAdj}} \cup \mathcal{D}_{\text{CYK}_2^i}^{\text{Radj}} \cup \\ \mathcal{D}_{\text{CYK}_2^i}^{\text{SubsUna}} \cup \mathcal{D}_{\text{CYK}_2^i}^{\text{SubsBin}_1} \cup \mathcal{D}_{\text{CYK}_2^i}^{\text{SubsBin}_2} \cup \mathcal{D}_{\text{CYK}_2^i}^{\text{SubsBin}_3}$$

$$\mathcal{D}_{\text{CYK}_2^i}^{\text{Scan}} = \frac{[a, j, j+1]}{[N^\gamma \rightarrow P^\gamma \bullet, j, j+1, \emptyset]} \quad \text{label}(P^\gamma) = a$$

$$\mathcal{D}_{\text{CYK}_2^i}^\varepsilon = \frac{}{[N^\gamma \rightarrow P^\gamma \bullet, j, j, \emptyset]} \quad \text{label}(P^\gamma) = \varepsilon \text{ o } \text{label}(P^\gamma) = \perp$$

$$\mathcal{D}_{\text{CYK}_2^i}^{\text{CompUna}} = \frac{[O^\gamma \rightarrow \nu \bullet, i, j, \text{code}]}{[M^\gamma \rightarrow O^\gamma \bullet, i, j, \emptyset]}$$

donde se debe cumplir:

- (i) ( $\text{nil} \in \text{ladj}(O^\gamma)$  y  $L \notin \text{code}$ ) o ( $\beta \in \text{ladj}(O^\gamma)$  y  $L \in \text{code}$ ),
- (ii) ( $\text{nil} \in \text{radj}(O^\gamma)$  y  $R \notin \text{code}$ ) o ( $\beta \in \text{radj}(O^\gamma)$  y  $R \in \text{code}$ ).

$$\mathcal{D}_{\text{CYK}_2^i}^{\text{CompBin}} = \frac{[O_1^\gamma \rightarrow \nu \bullet, i, j, \text{code}]}{[O_2^\gamma \rightarrow \delta \bullet, j, k, \text{code}']} \\ \frac{[O_2^\gamma \rightarrow \delta \bullet, j, k, \text{code}']}{[M^\gamma \rightarrow O_1^\gamma O_2^\gamma \bullet, i, k, \emptyset]}$$

donde se debe cumplir:

- (i) ( $\text{nil} \in \text{ladj}(O_1^\gamma)$  y  $L \notin \text{code}$ ) o ( $\beta \in \text{ladj}(O_1^\gamma)$  y  $L \in \text{code}$ ),

- (ii) ( $\mathbf{nil} \in \text{radj}(O_1^\gamma)$  y  $R \notin \text{code}$ ) o ( $\beta \in \text{radj}(O_1^\gamma)$  y  $R \in \text{code}$ ),  
 (iii) ( $\mathbf{nil} \in \text{ladj}(O_2^\gamma)$  y  $L \notin \text{code}'$ ) o ( $\beta \in \text{ladj}(O_2^\gamma)$  y  $L \in \text{code}'$ ),  
 (iv) ( $\mathbf{nil} \in \text{radj}(O_2^\gamma)$  y  $R \notin \text{code}'$ ) o ( $\beta \in \text{radj}(O_2^\gamma)$  y  $R \in \text{code}'$ ).

$$\mathcal{D}_{\text{CYK}_2^{\text{LAdj}}} = \frac{[\top \rightarrow \mathbf{R}^{\beta \bullet}, i, j, \emptyset] \quad [M^\gamma \rightarrow \nu \bullet, j, k, \text{code}]}{[M^\gamma \rightarrow \nu \bullet, i, k, \{L\} \cup \text{code}]} \quad \begin{array}{l} \beta \in \text{ladj}(M^\gamma) \\ L \notin \text{code} \end{array}$$

$$\mathcal{D}_{\text{CYK}_2^{\text{RAAdj}}} = \frac{[\top \rightarrow \mathbf{R}^{\beta \bullet}, j, k, \emptyset] \quad [M^\gamma \rightarrow \nu \bullet, i, j, \text{code}]}{[M^\gamma \rightarrow \nu \bullet, i, k, \{R\} \cup \text{code}]} \quad \begin{array}{l} \beta \in \text{radj}(M^\gamma) \\ R \notin \text{code} \end{array}$$

$$\mathcal{D}_{\text{CYK}_2^{\text{SubsUna}}} = \frac{[\top \rightarrow \mathbf{R}^{\alpha \bullet}, i, j, \emptyset]}{[M^\gamma \rightarrow O^\gamma \bullet, i, j, \emptyset]} \quad \alpha \in \text{subst}(O^\gamma)$$

$$\mathcal{D}_{\text{CYK}_2^{\text{SubsBin}_1}} = \frac{[\top \rightarrow \mathbf{R}^{\alpha \bullet}, i, j, \emptyset] \quad [O_2^\gamma \rightarrow \nu \bullet, j, k, \text{code}]}{[M^\gamma \rightarrow O_1^\gamma O_2^\gamma \bullet, i, k, \emptyset]} \quad \alpha \in \text{subst}(O_1^\gamma)$$

donde se debe cumplir:

- (i) ( $\mathbf{nil} \in \text{ladj}(O_2^\gamma)$  y  $L \notin \text{code}$ ) o ( $\beta \in \text{ladj}(O_2^\gamma)$  y  $L \in \text{code}$ ),  
 (ii) ( $\mathbf{nil} \in \text{radj}(O_2^\gamma)$  y  $R \notin \text{code}$ ) o ( $\beta \in \text{radj}(O_2^\gamma)$  y  $R \in \text{code}$ ).

$$\mathcal{D}_{\text{CYK}_2^{\text{SubsBin}_2}} = \frac{[O_1^\gamma \rightarrow \nu \bullet, i, j, \text{code}] \quad [\top \rightarrow \mathbf{R}^{\alpha \bullet}, j, k, \emptyset]}{[M^\gamma \rightarrow O_1^\gamma O_2^\gamma \bullet, i, k, \emptyset]} \quad \alpha \in \text{subst}(O_2^\gamma)$$

donde se debe cumplir:

- (i) ( $\mathbf{nil} \in \text{ladj}(O_1^\gamma)$  y  $L \notin \text{code}$ ) o ( $\beta \in \text{ladj}(O_1^\gamma)$  y  $L \in \text{code}$ ),  
 (ii) ( $\mathbf{nil} \in \text{radj}(O_1^\gamma)$  y  $R \notin \text{code}$ ) o ( $\beta \in \text{radj}(O_1^\gamma)$  y  $R \in \text{code}$ ).

$$\mathcal{D}_{\text{CYK}_2^{\text{SubsBin}_3}} = \frac{[\top \rightarrow \mathbf{R}^{\alpha \bullet}, i, j, \emptyset] \quad [\top \rightarrow \mathbf{R}^{\alpha_1 \bullet}, j, k, \emptyset]}{[M^\gamma \rightarrow O_1^\gamma O_2^\gamma \bullet, i, k, \emptyset]} \quad \begin{array}{l} \alpha \in \text{subst}(O_1^\gamma) \\ \alpha_1 \in \text{subst}(O_2^\gamma) \end{array}$$

El conjunto de ítems finales se define como:

$$\mathcal{F}_{\text{CYK}_2^i} = \{[\top \rightarrow \mathbf{R}^\alpha, 0, n, \text{code}] \mid \alpha \in I, \text{label}(\mathbf{R}^\alpha) = S\}$$

Probemos ahora la relación  $\text{CYK}_1^i \xrightarrow{\text{ir}} \text{CYK}_2^i$ . Para ello debemos mostrar que existe una función regular  $f : \mathcal{I}_{\text{CYK}_2^i} \rightarrow \mathcal{I}_{\text{CYK}_1^i}$  tal que:

1.  $\mathcal{I}_{\text{CYK}_1^i} = f(\mathcal{I}_{\text{CYK}_2^i})$
2.  $\Delta_{\text{CYK}_1^i} = f(\Delta_{\text{CYK}_2^i})$

Una función regular de contracción de ítems que cumple estas condiciones es la siguiente:

$$f([N^\gamma \rightarrow \delta \bullet, i, j, code]) = [N^\gamma, i, j, code]$$

De  $f$  se sigue inmediatamente que  $\mathcal{I}_{\text{CYK}_1^i} = f(\mathcal{I}_{\text{CYK}_2^i})$  y  $\Delta_{\text{CYK}_1^i} = f(\Delta_{\text{CYK}_2^i})$  por inducción en la longitud de las secuencias de derivación.

El esquema  $\text{ECYK}^i$  se obtiene a partir del esquema  $\text{buE}^i$ , restringiendo la clase de gramáticas sobre las que se define. De forma que el esquema  $\text{ECYK}^i$  sólo está definido para la clase de gramáticas de inserción de árboles en las cuales ningún nodo puede tener más de dos descendientes y los nodos etiquetados con símbolos terminales,  $\epsilon$  o  $\perp$  no tienen nodos hermanos.

Los conjuntos de ítems, pasos deductivos y finales del esquema  $\text{ECYK}^i$  son iguales a los del esquema  $\text{buE}^i$ .

$$\mathcal{I}_{\text{ECYK}^i} = \mathcal{I}_{\text{buE}^i}$$

$$\mathcal{D}_{\text{ECYK}^i} = \mathcal{D}_{\text{ECYK}^i}^{\text{Ini}} \cup \mathcal{D}_{\text{ECYK}^i}^{\text{Scan}} \cup \mathcal{D}_{\text{ECYK}^i}^{\epsilon} \cup \mathcal{D}_{\text{ECYK}^i}^{\text{Comp}} \cup \mathcal{D}_{\text{ECYK}^i}^{\text{LAdj}} \cup \mathcal{D}_{\text{ECYK}^i}^{\text{Radj}} \cup \mathcal{D}_{\text{ECYK}^i}^{\text{Subs}}$$

$$\mathcal{D}_{\text{ECYK}^i}^{\text{Ini}} = \mathcal{D}_{\text{buE}^i}^{\text{Ini}}$$

$$\mathcal{D}_{\text{ECYK}^i}^{\text{Scan}} = \mathcal{D}_{\text{buE}^i}^{\text{Scan}}$$

$$\mathcal{D}_{\text{ECYK}^i}^{\epsilon} = \mathcal{D}_{\text{buE}^i}^{\epsilon}$$

$$\mathcal{D}_{\text{ECYK}^i}^{\text{Comp}} = \mathcal{D}_{\text{buE}^i}^{\text{Comp}}$$

$$\mathcal{D}_{\text{ECYK}^i}^{\text{LAdj}} = \mathcal{D}_{\text{buE}^i}^{\text{LAdj}}$$

$$\mathcal{D}_{\text{ECYK}^i}^{\text{Radj}} = \mathcal{D}_{\text{buE}^i}^{\text{Radj}}$$

$$\mathcal{D}_{\text{ECYK}^i}^{\text{Subs}} = \mathcal{D}_{\text{buE}^i}^{\text{Subs}}$$

$$\mathcal{F}_{\text{ECYK}^i} = \mathcal{F}_{\text{buE}^i}$$

Para demostrar la relación  $\text{ECYK}^i \xrightarrow{\text{ext}} \text{buE}^i$  hay que probar:

1.  $CG_{\text{ECYK}^i} \subseteq CG_{\text{buE}^i}$
2.  $\text{ECYK}^i(G)(a_1 \dots a_n) = \text{buE}^i(G)(a_1 \dots a_n)$   
para toda  $G \in \mathcal{CG}_{\text{ECYK}^i}$  y cadena de entrada  $a_1 \dots a_n$

Lo primero es obvio, ya que la subclase de gramáticas sobre la que se define  $\text{ECYK}^i$  es un subconjunto de la clase de gramáticas de inserción de árboles sobre la cual está definido el esquema  $\text{buE}^i$ . Lo segundo es cierto por definición, ya que  $\text{ECYK}^i = \text{buE}^i$ .

A continuación probaremos la relación  $\text{CYK}_2^i \xrightarrow{\text{st}} \text{ECYK}^i$ . Para ello tenemos que demostrar que:

1.  $\mathcal{I}_{\text{CYK}_2^i} \subseteq \mathcal{I}_{\text{ECYK}^i}$

$$2. \vdash_{\text{CYK}_2}^* \subseteq \vdash_{\text{ECYK}^i}^*$$

Lo primero es cierto porque el dominio del esquema  $\text{CYK}_2^i$ , que sólo contempla ítems con el punto al final de la regla, está contenido en el dominio de  $\text{ECYK}^i$ :

$$\mathcal{I}_{\text{CYK}_2^i} \subset \mathcal{I}_{\text{ECYK}^i}$$

Para demostrar que  $\vdash_{\text{CYK}_2^i}^* \subseteq \vdash_{\text{ECYK}^i}^*$  nos basta con probar que  $\mathcal{D}_{\text{CYK}_2^i} \subseteq \mathcal{D}_{\text{ECYK}^i}^*$ . Veamos cada paso:

- Un paso deductivo  $\mathcal{D}_{\text{CYK}_2^i}^{\text{Scan}}$  es equivalente a una secuencia formada por un paso  $\mathcal{D}_{\text{ECYK}^i}^{\text{Ini}}$

$$\overline{[N^\gamma \rightarrow \bullet M^\gamma, j, j, \emptyset]}$$

seguido de otro  $\mathcal{D}_{\text{ECYK}^i}^{\text{Scan}}$

$$\frac{\begin{array}{c} [a, j, j+1] \\ [N^\gamma \rightarrow \bullet M^\gamma, j, j, \emptyset] \end{array}}{[N^\gamma \rightarrow M^\gamma \bullet, j, j+1, \emptyset]}$$

- Un paso deductivo  $\mathcal{D}_{\text{CYK}_2^i}^{\text{e}}$  es equivalente a una secuencia formada por un paso  $\mathcal{D}_{\text{ECYK}^i}^{\text{Ini}}$

$$\overline{[N^\gamma \rightarrow \bullet M^\gamma, j, j, \emptyset]}$$

seguido de otro  $\mathcal{D}_{\text{ECYK}^i}^{\text{e}}$

$$\frac{[N^\gamma \rightarrow \bullet M^\gamma, j, j, \emptyset]}{[N^\gamma \rightarrow M^\gamma \bullet, j, j, \emptyset]}$$

- Un paso deductivo  $\mathcal{D}_{\text{CYK}_2^i}^{\text{CompUna}}$  es equivalente a una secuencia formada por un paso  $\mathcal{D}_{\text{ECYK}^i}^{\text{Ini}}$

$$\overline{[N^\gamma \rightarrow \bullet M^\gamma, j, j, \emptyset]}$$

seguido de otro  $\mathcal{D}_{\text{ECYK}^i}^{\text{Comp}}$

$$\frac{\begin{array}{c} [M^\gamma \rightarrow \nu \bullet, j, k, \text{code}] \\ [N^\gamma \rightarrow \bullet M^\gamma, j, j, \emptyset] \end{array}}{[N^\gamma \rightarrow M^\gamma \bullet, j, k, \text{code}]}$$

- Un paso deductivo  $\mathcal{D}_{\text{CYK}_2^i}^{\text{CompBin}}$  es equivalente a una secuencia formada por un paso  $\mathcal{D}_{\text{ECYK}^i}^{\text{Ini}}$

$$\overline{[N^\gamma \rightarrow \bullet M^\gamma P^\gamma, i, i, \emptyset]}$$

seguido de una secuencia de dos pasos  $\mathcal{D}_{\text{ECYK}^i}^{\text{Comp}}$

$$\frac{\begin{array}{c} [M^\gamma \rightarrow \nu \bullet, i, j, \text{code}] \\ [N^\gamma \rightarrow \bullet M^\gamma P^\gamma, i, i, \emptyset] \end{array}}{[N^\gamma \rightarrow M^\gamma \bullet P^\gamma, i, j, \emptyset]}$$

$$\frac{\begin{array}{c} [P^\gamma \rightarrow \omega \bullet, j, k, \text{code}'] \\ [N^\gamma \rightarrow M^\gamma \bullet P^\gamma, i, j, \emptyset] \end{array}}{[N^\gamma \rightarrow M^\gamma P^\gamma \bullet, i, k, \emptyset]}$$

- Un paso deductivo  $\mathcal{D}_{\text{CYK}_2^{\text{SubsUna}}}^{\text{SubsUna}}$  es equivalente a una secuencia formada por un paso  $\mathcal{D}_{\text{ECYK}^i}^{\text{Ini}}$

$$\overline{[N^\gamma \rightarrow \bullet M^\gamma, j, j, \emptyset]}$$

seguido de otro  $\mathcal{D}_{\text{ECYK}^i}^{\text{Subs}}$

$$\frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\alpha \bullet, j, k, \emptyset] \\ [N^\gamma \rightarrow \bullet M^\gamma, j, j, \emptyset] \end{array}}{[N^\gamma \rightarrow M^\gamma \bullet, j, k, \emptyset]}$$

- Un paso deductivo  $\mathcal{D}_{\text{CYK}_2^{\text{SubsBin1}}}^{\text{SubsBin1}}$  es equivalente a una secuencia formada por un paso  $\mathcal{D}_{\text{ECYK}^i}^{\text{Ini}}$

$$\overline{[N^\gamma \rightarrow \bullet M^\gamma P^\gamma, i, i, \emptyset]}$$

seguido de uno  $\mathcal{D}_{\text{ECYK}^i}^{\text{Subs}}$

$$\frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\alpha \bullet, i, j, \emptyset] \\ [N^\gamma \rightarrow \bullet M^\gamma P^\gamma, i, i, \emptyset] \end{array}}{[N^\gamma \rightarrow M^\gamma \bullet P^\gamma, i, j, \emptyset]}$$

y de otro  $\mathcal{D}_{\text{ECYK}^i}^{\text{Comp}}$

$$\frac{\begin{array}{l} [P^\gamma \rightarrow \nu \bullet, j, k, \text{code}] \\ [N^\gamma \rightarrow M^\gamma \bullet P^\gamma, i, j, \emptyset] \end{array}}{[N^\gamma \rightarrow M^\gamma P^\gamma \bullet, i, k, \emptyset]}$$

- Un paso deductivo  $\mathcal{D}_{\text{CYK}_2^{\text{SubsBin2}}}^{\text{SubsBin2}}$  es equivalente a una secuencia formada por un paso  $\mathcal{D}_{\text{ECYK}^i}^{\text{Ini}}$

$$\overline{[N^\gamma \rightarrow \bullet M^\gamma P^\gamma, i, i, \emptyset]}$$

seguido de uno  $\mathcal{D}_{\text{ECYK}^i}^{\text{Comp}}$

$$\frac{\begin{array}{l} [M^\gamma \rightarrow \nu \bullet, i, j, \text{code}] \\ [N^\gamma \rightarrow \bullet M^\gamma P^\gamma, i, i, \emptyset] \end{array}}{[N^\gamma \rightarrow M^\gamma \bullet P^\gamma, i, j, \emptyset]}$$

y de otro  $\mathcal{D}_{\text{ECYK}^i}^{\text{Subs}}$

$$\frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\alpha \bullet, j, k, \emptyset] \\ [N^\gamma \rightarrow M^\gamma \bullet P^\gamma, i, j, \emptyset] \end{array}}{[N^\gamma \rightarrow M^\gamma P^\gamma \bullet, i, k, \emptyset]}$$

- Un paso deductivo  $\mathcal{D}_{\text{CYK}_2^{\text{SubsBin3}}}^{\text{SubsBin3}}$  es equivalente a una secuencia formada por un paso  $\mathcal{D}_{\text{ECYK}^i}^{\text{Ini}}$

$$\overline{[N^\gamma \rightarrow \bullet M^\gamma P^\gamma, i, i, \emptyset]}$$

seguido de dos  $\mathcal{D}_{\text{ECYK}^i}^{\text{Subs}}$

$$\frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\alpha \bullet, i, j, \emptyset] \\ [N^\gamma \rightarrow \bullet M^\gamma P^\gamma, i, i, \emptyset] \end{array}}{[N^\gamma \rightarrow M^\gamma \bullet P^\gamma, i, j, \emptyset]}$$

$$\frac{\begin{array}{l} [\Gamma \rightarrow \mathbf{R}^{\alpha_1 \bullet, j, k, \emptyset}] \\ [N^\gamma \rightarrow M^\gamma \bullet P^\gamma, i, j, \emptyset] \end{array}}{[N^\gamma \rightarrow M^\gamma P^\gamma \bullet, i, k, \emptyset]}$$

- El resto de pasos son iguales en ambos esquemas:

$$\mathcal{D}_{\text{CYK}_2^{\text{LAdj}}} = \mathcal{D}_{\text{ECYK}^{\text{LAdj}}}$$

$$\mathcal{D}_{\text{CYK}_2^{\text{RAdj}}} = \mathcal{D}_{\text{ECYK}^{\text{RAdj}}}$$

**Teorema 8.2** *Relación entre los esquemas  $\text{dVH}^i$  y  $\text{buE}^i$*

Se mantienen las siguientes relaciones de refinamiento y contracción de secuencias deductivas entre esquemas:

$$\text{dVH}^i \xrightarrow{\text{sr}} \text{dVH}_1^i \xrightarrow{\text{sc}} \text{buE}^i$$

*Prueba*

Primero vamos a definir el esquema intermedio  $\text{dVH}_1^i$  que nos servirá de enlace en esta relación. Se trata de un analizador con un recorrido de la cadena de entrada bidireccional, pero el reconocimiento ascendente comienza con un paso de inicio al igual que el esquema  $\text{buE}^i$ , aunque, a diferencia de éste, desde todas las posibles posiciones dentro de cada regla.

Su dominio es igual al del esquema  $\text{dVH}^i$ :

$$\mathcal{I}_{\text{dVH}_1^i} = \mathcal{I}_{\text{dVH}^i}$$

El conjunto de pasos deductivos viene dado por:

$$\mathcal{D}_{\text{dVH}_1^i} = \mathcal{D}_{\text{dVH}_1^{\text{Ini}}} \cup \mathcal{D}_{\text{dVH}_1^{\text{Scan}}} \cup \mathcal{D}_{\text{dVH}_1^\epsilon} \cup \mathcal{D}_{\text{dVH}_1^{\text{Comp}}} \cup \mathcal{D}_{\text{dVH}_1^{\text{Con}}} \cup \mathcal{D}_{\text{dVH}_1^{\text{LAdj}}} \cup \mathcal{D}_{\text{dVH}_1^{\text{RAdj}}} \cup \mathcal{D}_{\text{dVH}_1^{\text{Subs}}}$$

$$\mathcal{D}_{\text{dVH}_1^{\text{Comp}}} = \mathcal{D}_{\text{dVH}^{\text{Comp}}}$$

$$\mathcal{D}_{\text{dVH}_1^{\text{Con}}} = \mathcal{D}_{\text{dVH}^{\text{Con}}}$$

$$\mathcal{D}_{\text{dVH}_1^{\text{LAdj}}} = \mathcal{D}_{\text{dVH}^{\text{LAdj}}}$$

$$\mathcal{D}_{\text{dVH}_1^{\text{RAdj}}} = \mathcal{D}_{\text{dVH}^{\text{RAdj}}}$$

$$\mathcal{D}_{\text{dVH}_1^{\text{Subs}}} = \mathcal{D}_{\text{dVH}^{\text{Subs}}}$$

$$\mathcal{D}_{\text{dVH}_1^{\text{Ini}}} = \overline{[N^\gamma \rightarrow \nu \bullet \bullet \omega, j, j, \emptyset]}$$

$$\mathcal{D}_{\text{dVH}_1^{\text{Scan}}} = \frac{\begin{array}{l} [a, j, j+1] \\ [N^\gamma \rightarrow \nu \bullet \bullet M^\gamma \omega, j, j, \emptyset] \end{array}}{[N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, j, j+1, \emptyset]} \quad \text{label}(M^\gamma) = a$$

$$\mathcal{D}_{dVH_1}^\varepsilon = \frac{[N^\gamma \rightarrow \nu \bullet \bullet M^\gamma \omega, j, j, \emptyset]}{[N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, j, j, \emptyset]} \text{ label}(M^\gamma) = \varepsilon \text{ o } \text{label}(M^\gamma) = \perp$$

El conjunto de ítems finales es igual al de  $\mathcal{F}_{dVH_1}$ :

$$\mathcal{F}_{dVH_1} = \mathcal{F}_{dVH_1}$$

Ahora probaremos la relación  $dVH_1 \xrightarrow{sr} dVH_1^i$ . Para ello tenemos que demostrar que:

1.  $\mathcal{I}_{dVH_1} \subseteq \mathcal{I}_{dVH_1^i}$
2.  $\vdash_{dVH_1}^* \subseteq \vdash_{dVH_1^i}^*$

Lo primero es cierto porque los dominios de ambos esquemas son iguales:

$$\mathcal{I}_{dVH_1} = \mathcal{I}_{dVH_1^i}$$

Para demostrar que  $\vdash_{dVH_1}^* \subseteq \vdash_{dVH_1^i}^*$  hay que probar que  $\mathcal{D}_{dVH_1} \subseteq \mathcal{D}_{dVH_1^i}^*$ . Veamos cada paso:

- Un paso deductivo  $\mathcal{D}_{dVH_1}^{\text{Scan}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{dVH_1^i}^{\text{Ini}}$

$$\overline{[N^\gamma \rightarrow \nu \bullet \bullet M^\gamma \omega, j, j, \emptyset]}$$

seguido de otro  $\mathcal{D}_{dVH_1^i}^{\text{Scan}}$

$$\frac{\frac{[a, j, j + 1]}{[N^\gamma \rightarrow \nu \bullet \bullet M^\gamma \omega, j, j, \emptyset]}}{[N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, j, j + 1, \emptyset]}$$

- Un paso deductivo  $\mathcal{D}_{dVH_1}^\varepsilon$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{dVH_1^i}^{\text{Ini}}$

$$\overline{[N^\gamma \rightarrow \nu \bullet \bullet M^\gamma \omega, j, j, \emptyset]}$$

seguido de otro  $\mathcal{D}_{dVH_1^i}^\varepsilon$

$$\frac{[N^\gamma \rightarrow \nu \bullet \bullet M^\gamma \omega, j, j, \emptyset]}{[N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, j, j, \emptyset]}$$

- El resto de pasos son iguales en ambos esquemas:

$$\mathcal{D}_{dVH_1}^{\text{Comp}} = \mathcal{D}_{dVH_1^i}^{\text{Comp}}$$

$$\mathcal{D}_{dVH_1}^{\text{Con}} = \mathcal{D}_{dVH_1^i}^{\text{Con}}$$

$$\mathcal{D}_{dVH_1}^{\text{LAdj}} = \mathcal{D}_{dVH_1^i}^{\text{LAdj}}$$

$$\mathcal{D}_{dVH_1}^{\text{RAAdj}} = \mathcal{D}_{dVH_1^i}^{\text{RAAdj}}$$

$$\mathcal{D}_{dVH_1}^{\text{Subs}} = \mathcal{D}_{dVH_1^i}^{\text{Subs}}$$



Nos queda probar que  $dVH_1^i \xrightarrow{sc} buE^i$ , para lo cual tenemos que demostrar que:

1.  $\mathcal{I}_{buE^i} \subseteq \mathcal{I}_{dVH_1^i}$
2.  $\vdash_{buE^i}^* \subseteq \vdash_{dVH_1^i}^*$

Si establecemos la siguiente equivalencia de ítems:

$$[N^\gamma \rightarrow \nu \bullet \omega, i, j, code] \in \mathcal{I}_{buE^i} \equiv [N^\gamma \rightarrow \bullet \nu \bullet \omega, i, j, code] \in \mathcal{I}_{dVH_1^i},$$

entonces se cumple

$$\mathcal{I}_{buE^i} \subset \mathcal{I}_{dVH_1^i},$$

ya que, mientras el primero sólo reconoce fragmentos que comienzan al principio de las reglas, el segundo no presenta esta limitación.

Para probar que  $\vdash_{buE^i}^* \subseteq \vdash_{dVH_1^i}^*$  debemos demostrar que  $\mathcal{D}_{buE^i} \subseteq \mathcal{D}_{dVH_1^i}$ . Veamos cada uno de los pasos:

- Un paso deductivo  $\mathcal{D}_{buE^i}^{Ini}$  es equivalente a un paso  $\mathcal{D}_{dVH_1^i}^{Ini}$

$$\overline{[N^\gamma \rightarrow \bullet \bullet \nu, j, j, \emptyset]}$$

- Un paso deductivo  $\mathcal{D}_{buE^i}^{Scan}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{dVH_1^i}^{Ini}$

$$\overline{[N^\gamma \rightarrow \nu \bullet \bullet M^\gamma \omega, j, j, \emptyset]}$$

seguido de otro  $\mathcal{D}_{dVH_1^i}^{Scan}$

$$\frac{\begin{array}{c} [a, j, j+1] \\ [N^\gamma \rightarrow \nu \bullet \bullet M^\gamma \omega, j, j, \emptyset] \end{array}}{[N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, j, j+1, \emptyset]}$$

y de otro  $\mathcal{D}_{dVH_1^i}^{Con}$

$$\frac{\begin{array}{c} [N^\gamma \rightarrow \bullet \nu \bullet M^\gamma \omega, i, j, \emptyset] \\ [N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, j, j+1, \emptyset] \end{array}}{[N^\gamma \rightarrow \bullet \nu M^\gamma \bullet \omega, i, j+1, \emptyset]}$$

- Un paso deductivo  $\mathcal{D}_{buE^i}^\epsilon$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{dVH_1^i}^{Ini}$

$$\overline{[N^\gamma \rightarrow \nu \bullet \bullet M^\gamma \omega, j, j, \emptyset]}$$

seguido de otro  $\mathcal{D}_{dVH_1^i}^\epsilon$

$$\frac{[N^\gamma \rightarrow \nu \bullet \bullet M^\gamma \omega, j, j, \emptyset]}{[N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, j, j+1, \emptyset]}$$

y de otro  $\mathcal{D}_{dVH_1^i}^{Con}$

$$\frac{\begin{array}{c} [N^\gamma \rightarrow \bullet \nu \bullet M^\gamma \omega, i, j, \emptyset] \\ [N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, j, j, \emptyset] \end{array}}{[N^\gamma \rightarrow \bullet \nu M^\gamma \bullet \omega, i, j, \emptyset]}$$

- Un paso deductivo  $\mathcal{D}_{\text{buE}^i}^{\text{Comp}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{dVH}_1^{\text{Comp}}}$

$$\frac{[M^\gamma \rightarrow \bullet \delta \bullet, j, k, \text{code}]}{[N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, j, k, \emptyset]}$$

seguido de otro  $\mathcal{D}_{\text{dVH}_1^{\text{Con}}}$

$$\frac{\begin{array}{l} [N^\gamma \rightarrow \bullet \nu \bullet M^\gamma \omega, i, j, \emptyset] \\ [N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, j, k, \emptyset] \end{array}}{[N^\gamma \rightarrow \bullet \nu M^\gamma \bullet \omega, i, k, \emptyset]}$$

- Un paso deductivo  $\mathcal{D}_{\text{buE}^i}^{\text{LAdj}}$  es equivalente a un paso  $\mathcal{D}_{\text{dVH}_1^{\text{LAdj}}}$

$$\frac{\begin{array}{l} [\top \rightarrow \bullet \mathbf{R}^\beta \bullet, i, j, \emptyset] \\ [M^\gamma \rightarrow \bullet \delta \bullet, j, k, \text{code}] \end{array}}{[M^\gamma \rightarrow \bullet \delta \bullet, i, k, \{L\} \cup \text{code}]}$$

- Un paso deductivo  $\mathcal{D}_{\text{buE}^i}^{\text{RAAdj}}$  es equivalente a un paso  $\mathcal{D}_{\text{dVH}_1^{\text{RAAdj}}}$

$$\frac{\begin{array}{l} [\top \rightarrow \bullet \mathbf{R}^\beta \bullet, j, k, \emptyset] \\ [M^\gamma \rightarrow \bullet \delta \bullet, i, j, \text{code}] \end{array}}{[M^\gamma \rightarrow \bullet \delta \bullet, i, k, \{R\} \cup \text{code}]}$$

- Un paso deductivo  $\mathcal{D}_{\text{buE}^i}^{\text{Subs}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{dVH}_1^{\text{Subs}}}$

$$\frac{[\top \rightarrow \bullet \mathbf{R}^\alpha \bullet, j, k, \emptyset]}{[N^\gamma \rightarrow \delta \bullet M^\gamma \bullet \omega, j, k, \emptyset]}$$

seguido de otro  $\mathcal{D}_{\text{dVH}_1^{\text{Con}}}$

$$\frac{\begin{array}{l} [N^\gamma \rightarrow \bullet \nu \bullet M^\gamma \omega, i, j, \emptyset] \\ [N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, j, k, \emptyset] \end{array}}{[N^\gamma \rightarrow \bullet \nu M^\gamma \bullet \omega, i, k, \emptyset]}$$

**Teorema 8.3** Relación entre los esquemas  $\text{buE}^k$  y  $\text{buE}^i$

Se mantienen las siguientes relaciones entre esquemas:

$$\text{buE}^i \xrightarrow{\text{sr}} \text{buE}_1^i \xrightarrow{\text{sf}} \text{buE}_2^i \xrightarrow{\text{ic}} \text{buE}^k$$

*Prueba*

El esquema  $\text{buE}_1^i$  se obtiene desdoblado el paso de adjunción izquierda del esquema  $\text{buE}^i$  mediante una aplicación parcial.

El dominio del nuevo esquema se ve modificado por la introducción de un tipo de ítem intermedio que almacena el resultado del primer paso de la aplicación parcial.

$$\mathcal{I}_{\text{buE}_1^i} = \mathcal{I}_{\text{buE}_1^i}^{(i)} \cup \mathcal{I}_{\text{buE}_1^i}^{(ii)}$$

$$\mathcal{I}_{\text{buE}_1^i}^{(i)} = \{[M^\gamma \rightarrow \delta \bullet \nu, i, j, \text{code}] \mid M^\gamma \rightarrow \delta \nu \in \mathcal{P}(\gamma), \gamma \in I \cup A, \\ 0 \leq i \leq j, \nu \neq \varepsilon, \text{code} \subseteq \{L\}\}$$

$$\mathcal{I}_{\text{buE}_1^i}^{(ii)} = \mathcal{I}_{\text{buE}^i}^{(ii)}$$

El conjunto de pasos deductivos es:

$$\mathcal{D}_{\text{buE}_1^i} = \mathcal{D}_{\text{buE}_1^i}^{\text{Ini}} \cup \mathcal{D}_{\text{buE}_1^i}^{\text{Scan}} \cup \mathcal{D}_{\text{buE}_1^i}^\varepsilon \cup \mathcal{D}_{\text{buE}_1^i}^{\text{Comp}} \cup \mathcal{D}_{\text{buE}_1^i}^{\text{LAdj}^1} \cup \mathcal{D}_{\text{buE}_1^i}^{\text{LAdj}^2} \cup \mathcal{D}_{\text{buE}_1^i}^{\text{RAAdj}} \cup \mathcal{D}_{\text{buE}_1^i}^{\text{Subs}}$$

$$\mathcal{D}_{\text{buE}_1^i}^{\text{Ini}} = \mathcal{D}_{\text{buE}^i}^{\text{Ini}}$$

$$\mathcal{D}_{\text{buE}_1^i}^{\text{Scan}} = \frac{[a, j, j+1] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{code}]}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, j+1, \text{code}]} \quad \text{label}(M^\gamma) = a$$

$$\mathcal{D}_{\text{buE}_1^i}^\varepsilon = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{code}]}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, j, \text{code}]} \quad \begin{array}{l} \text{label}(M^\gamma) = \varepsilon \\ \text{ó label}(M^\gamma) = \perp \end{array}$$

$$\mathcal{D}_{\text{buE}_1^i}^{\text{Comp}} = \frac{[M^\gamma \rightarrow \omega \bullet, j, k, \text{code}] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{code}']}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k, \text{code}']}$$

donde se debe cumplir:

- (i) ( $\text{nil} \in \text{ladj}(M^\gamma)$  y  $L \notin \text{code}$ ) o ( $\beta \in \text{ladj}(M^\gamma)$  y  $L \in \text{code}$ ),
- (ii) ( $\text{nil} \in \text{radj}(M^\gamma)$  y  $R \notin \text{code}$ ) o ( $\beta \in \text{radj}(M^\gamma)$  y  $R \in \text{code}$ ).

$$\mathcal{D}_{\text{buE}_1^i}^{\text{RAAdj}} = \frac{[\top \rightarrow \mathbf{R}^\beta \bullet, j, k, \emptyset] \\ [M^\gamma \rightarrow \nu \bullet, i, j, \text{code}]}{[M^\gamma \rightarrow \nu \bullet, i, k, \{R\} \cup \text{code}]} \quad \begin{array}{l} \beta \in \text{radj}(M^\gamma) \\ R \notin \text{code} \end{array}$$

$$\mathcal{D}_{\text{buE}_1^i}^{\text{Subs}} = \frac{[\top \rightarrow \mathbf{R}^\alpha \bullet, j, k, \emptyset] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{code}]}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k, \text{code}]} \quad \alpha \in \text{subst}(M^\gamma)$$

$$\mathcal{D}_{\text{buE}_1^i}^{\text{LAdj}^1} = \frac{[\top \rightarrow \mathbf{R}^\beta \bullet, i, j, \emptyset]}{[M^\gamma \rightarrow \bullet \nu, i, j, \{L\}]} \quad \beta \in \text{ladj}(M^\gamma)$$

$$\mathcal{D}_{\text{buE}_1^i}^{\text{LAdj}^2} = \frac{[M^\gamma \rightarrow \bullet \nu, i, j, \{L\}] \\ [M^\gamma \rightarrow \nu \bullet, j, k, \text{code}]}{[M^\gamma \rightarrow \nu \bullet, i, k, \{L\} \cup \text{code}]} \quad L \notin \text{code}$$

El conjunto de ítems finales es idéntico al del esquema  $\text{buE}^i$ :

$$\mathcal{F}_{\text{buE}_1^i} = \mathcal{F}_{\text{buE}^i}$$

Para demostrar que  $\text{buE}^i \xrightarrow{\text{sr}} \text{buE}_1^i$  hay que probar que:

$$1. \mathcal{I}_{\text{buE}^i} \subseteq \mathcal{I}_{\text{buE}_1^i}$$

$$2. \vdash_{\text{buE}^i}^* \subseteq \vdash_{\text{buE}_1^i}^*$$

Como el dominio de  $\text{buE}_1^i$  incluye los ítems de  $\text{buE}^i$  y el tipo de ítem intermedio se cumple que:

$$\mathcal{I}_{\text{buE}^i} \subset \mathcal{I}_{\text{buE}_1^i}$$

Para demostrar que  $\vdash_{\text{buE}^i}^* \subseteq \vdash_{\text{buE}_1^i}^*$  nos basta con probar que  $\mathcal{D}_{\text{buE}^i} \subseteq \mathcal{D}_{\text{buE}_1^i}^*$ .

- Un paso deductivo  $\mathcal{D}_{\text{buE}^i}^{\text{LAdj}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{buE}_1^i}^{\text{LAdj}^1}$

$$\frac{[\text{T} \rightarrow \mathbf{R}^\beta \bullet, i, j, \emptyset]}{[M^\gamma \rightarrow \bullet \nu, i, j, \{L\}]}$$

seguido de otro  $\mathcal{D}_{\text{buE}_1^i}^{\text{LAdj}^2}$

$$\frac{\begin{array}{l} [M^\gamma \rightarrow \bullet \nu, i, j, \{L\}] \\ [M^\gamma \rightarrow \nu \bullet, j, k, \text{code}] \end{array}}{[M^\gamma \rightarrow \nu \bullet, i, k, \{L\} \cup \text{code}]}$$

- Para el resto de pasos se cumple:

$$\mathcal{D}_{\text{buE}^i}^{\text{Ini}} = \mathcal{D}_{\text{buE}_1^i}^{\text{Ini}}$$

$$\mathcal{D}_{\text{buE}^i}^{\text{Scan}} \subset \mathcal{D}_{\text{buE}_1^i}^{\text{Scan}}$$

$$\mathcal{D}_{\text{buE}^i}^\varepsilon \subset \mathcal{D}_{\text{buE}_1^i}^\varepsilon$$

$$\mathcal{D}_{\text{buE}^i}^{\text{Comp}} \subset \mathcal{D}_{\text{buE}_1^i}^{\text{Comp}}$$

$$\mathcal{D}_{\text{buE}^i}^{\text{RAAdj}} = \mathcal{D}_{\text{buE}_1^i}^{\text{RAAdj}}$$

$$\mathcal{D}_{\text{buE}^i}^{\text{Subs}} \subset \mathcal{D}_{\text{buE}_1^i}^{\text{Subs}}$$

El esquema  $\text{buE}_2^i$  se obtiene eliminando el paso  $\mathcal{D}_{\text{buE}_1^i}^{\text{LAdj}^2}$  del esquema  $\text{buE}_1^i$ . Este paso no es necesario, ya que los ítems que genera son derivados igualmente en el proceso de análisis a partir del consecuente de  $\mathcal{D}_{\text{buE}_1^i}^{\text{LAdj}^1}$ .

Para probar que  $\text{buE}_1^i \xrightarrow{\text{sf}} \text{buE}_2^i$  basta con que se cumpla:

$$1. \mathcal{I}_{\text{buE}_2^i} \subseteq \mathcal{I}_{\text{buE}_1^i}$$

$$2. \mathcal{D}_{\text{buE}_1^i} \subseteq \mathcal{D}_{\text{buE}_2^i}$$

Ambas condiciones son ciertas por definición, ya que  $\mathcal{I}_{\text{buE}_2^i} = \mathcal{I}_{\text{buE}_1^i}$  y  $\mathcal{D}_{\text{buE}_1^i} \subset \mathcal{D}_{\text{buE}_2^i}$ .

El esquema  $\text{buE}^k$  se obtiene cambiando la semántica y tipo del parámetro auxiliar code en el esquema  $\text{buE}_2^i$ , de manera que dicho parámetro tomará el valor true cuando se haya efectuado una adjunción derecha sobre el nodo y el valor false en cualquier otro caso.

Probemos ahora la relación  $\text{buE}_2^i \xrightarrow{\text{ic}} \text{buE}^k$ . Para ello debemos mostrar que existe una función regular  $f : \mathcal{I}_{\text{buE}_2^i} \rightarrow \mathcal{I}_{\text{buE}^k}$  tal que:



1.  $\mathcal{I}_{\text{buE}^k} = f(\mathcal{I}_{\text{buE}_2^i})$
2.  $\Delta_{\text{buE}^k} = f(\Delta_{\text{buE}_2^i})$

Una función regular de contracción de ítems que cumple estas condiciones es la siguiente:

$$f([N^\gamma \rightarrow \delta \bullet \nu, i, j, \text{code}]) = [N^\gamma \rightarrow \delta \bullet \nu, i, j, \text{radj}]$$

donde  $\text{radj} = \text{true}$  si  $R \in \text{code}$  y  $\text{radj} = \text{false}$  en cualquier otro caso.

De  $f$  se sigue inmediatamente que  $\mathcal{I}_{\text{buE}^k} = f(\mathcal{I}_{\text{buE}_2^i})$  y  $\Delta_{\text{buE}^k} = f(\Delta_{\text{buE}_2^i})$  por inducción en la longitud de las secuencias de derivación.

**Teorema 8.4** *Relación entre los esquemas  $\text{buE}^k$  y Earley<sup>i</sup>*

Se mantiene la siguiente relación de filtro dinámico entre esquemas:

$$\text{buE}^k \stackrel{\text{df}}{\Rightarrow} \text{Earley}^i$$

*Prueba*

Tenemos que demostrar que:

1.  $\mathcal{I}_{\text{Earley}^i} \subseteq \mathcal{I}_{\text{buE}^k}$
2.  $\vdash_{\text{Earley}^i} \subseteq \vdash_{\text{buE}^k}$

Lo primero es cierto por definición, ya que los dominios de ambos esquemas son iguales:

$$\mathcal{I}_{\text{Earley}^i} = \mathcal{I}_{\text{buE}^k}$$

Para probar que  $\vdash_{\text{Earley}^i} \subseteq \vdash_{\text{buE}^k}$  debemos demostrar que  $\mathcal{D}_{\text{Earley}^i} \subseteq \mathcal{D}_{\text{buE}^k}$ . Veamos cada uno de los pasos:

- Dado un paso

$$\frac{}{[\top \rightarrow \bullet \mathbf{R}^\alpha, 0, 0, \text{false}]} \in \mathcal{D}_{\text{Earley}^i}^{\text{Ini}}$$

existe un paso

$$\frac{}{[N^\gamma \rightarrow \bullet \nu, i, i, \text{false}]} \in \mathcal{D}_{\text{buE}^k}^{\text{Ini}}$$

y, por tanto, existe la inferencia

$$\vdash_{\text{buE}^k} [\top \rightarrow \bullet \mathbf{R}^\alpha, 0, 0, \text{false}]$$

- Dado un paso

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{false}]}{[M^\gamma \rightarrow \bullet \nu, j, j, \text{false}]} \in \mathcal{D}_{\text{Earley}^i}^{\text{Pred}}$$

existe un paso

$$\frac{}{[M^\gamma \rightarrow \bullet \nu, j, j, \text{false}]} \in \mathcal{D}_{\text{buE}^k}^{\text{Ini}}$$

y, por tanto, existe la inferencia

$$[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{false}] \vdash_{\text{buE}^k} [M^\gamma \rightarrow \bullet \nu, j, j, \text{false}]$$

- Dado un paso

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[\top \rightarrow \bullet \mathbf{R}^\beta, j, j, false]} \in \mathcal{D}_{\text{Earley}^i}^{\text{LAdjPred}}$$

existe un paso

$$\overline{[N^\gamma \rightarrow \bullet \nu, j, j, false]} \in \mathcal{D}_{\text{buE}^k}^{\text{Ini}}$$

y, por tanto, existe la inferencia

$$[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false] \vdash_{\text{buE}^k} [\top \rightarrow \bullet \mathbf{R}^\beta, j, j, false]$$

- Dado un paso

$$\frac{[\top \rightarrow \mathbf{R}^\beta \bullet, j, k, false] \quad [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[M^\gamma \rightarrow \bullet \nu, j, k, false]} \in \mathcal{D}_{\text{Earley}^i}^{\text{LAdjComp}}$$

existe un paso

$$\frac{[\top \rightarrow \mathbf{R}^\beta \bullet, j, k, false]}{[M^\gamma \rightarrow \bullet \nu, j, k, false]} \in \mathcal{D}_{\text{buE}^k}^{\text{LAdj}}$$

y, por tanto, existe la inferencia

$$[\top \rightarrow \mathbf{R}^\beta \bullet, j, k, false] \quad [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false] \vdash_{\text{buE}^k} [M^\gamma \rightarrow \bullet \nu, j, k, false]$$

- Dado un paso

$$\frac{[M^\gamma \rightarrow \nu \bullet, i, j, false]}{[\top \rightarrow \bullet \mathbf{R}^\beta, j, j, false]} \in \mathcal{D}_{\text{Earley}^i}^{\text{RAAdjPred}}$$

existe un paso

$$\overline{[N^\gamma \rightarrow \bullet \nu, j, j, false]} \in \mathcal{D}_{\text{buE}^k}^{\text{Ini}}$$

y, por tanto, existe la inferencia

$$[M^\gamma \rightarrow \nu \bullet, i, j, false] \vdash_{\text{buE}^k} [\top \rightarrow \bullet \mathbf{R}^\beta, j, j, false]$$

- Dado un paso

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[\top \rightarrow \bullet \mathbf{R}^\alpha, j, j, false]} \in \mathcal{D}_{\text{Earley}^i}^{\text{SubsPred}}$$

existe un paso

$$\overline{[N^\gamma \rightarrow \bullet \nu, j, j, false]} \in \mathcal{D}_{\text{buE}^k}^{\text{Ini}}$$

y, por tanto, existe la inferencia

$$[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false] \vdash_{\text{buE}^k} [\top \rightarrow \bullet \mathbf{R}^\alpha, j, j, false]$$



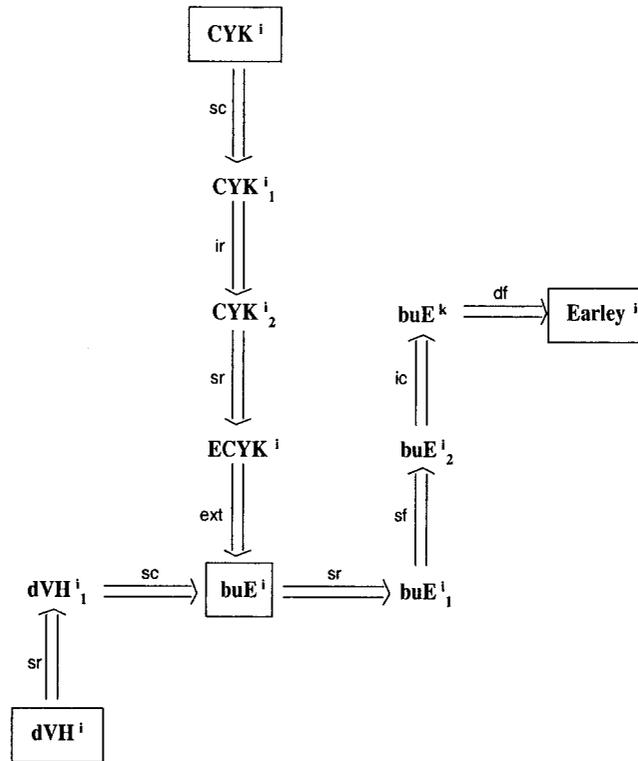


Figura 8.1: Red de analizadores para TIG

- El resto de pasos mantienen las siguientes equivalencias:

$$\mathcal{D}_{\text{Earley}^i}^{\text{Scan}} = \mathcal{D}_{\text{buE}^k}^{\text{Scan}}$$

$$\mathcal{D}_{\text{Earley}^i}^{\epsilon} = \mathcal{D}_{\text{buE}^k}^{\epsilon}$$

$$\mathcal{D}_{\text{Earley}^i}^{\text{Comp}} = \mathcal{D}_{\text{buE}^k}^{\text{Comp}}$$

$$\mathcal{D}_{\text{Earley}^i}^{\text{RAjComp}} = \mathcal{D}_{\text{buE}^k}^{\text{RAj}}$$

$$\mathcal{D}_{\text{Earley}^i}^{\text{SubsComp}} = \mathcal{D}_{\text{buE}^k}^{\text{Subs}}$$

En la figura 8.1 se presenta esquemáticamente todas las relaciones mostradas en esta sección.

## Capítulo 9

# Algoritmos de análisis basados en Left Corner para TIG

En este capítulo presentaremos un conjunto de analizadores para TIG contruidos mediante la aplicación de filtros basados en la relación de Left Corner para TIG a los esquemas tipo Earley descritos en el capítulo anterior. La relación de Left Corner para TIG se definirá como una modificación de la definida para TAG.

El capítulo se estructura de la siguiente forma. En la sección 1 se define un esquema tipo Left Corner ascendente, obtenido mediante la aplicación de un filtro basado en la esquina izquierda al analizador tipo Earley ascendente. En la sección 2 se presenta el concepto de relación de Left Corner en las TIG. En la sección 3 se define un esquema tipo Left Corner que hace uso de la relación de Left Corner para TIG definida en la sección 2. En la sección 4 se presenta un esquema tipo Left Corner que servirá de enlace para establecer las relaciones formales entre los esquemas de las secciones 3 y 5. En la sección 5 definimos un analizador de tipo Left Corner que se obtiene aplicando un filtro basado en la relación de Left Corner para TIG al algoritmo tipo Earley. En la última sección se muestran las relaciones formales existentes entre los esquemas introducidos en este capítulo, y entre éstos y los definidos en el capítulo anterior.

### 9.1 Esquema tipo Left Corner ascendente

De forma similar el esquema **buLC** para TAG, este esquema elimina del dominio de **buE<sup>k</sup>** aquellos ítems que no aportan nada significativo en el proceso de análisis, y que son los que tienen la forma:

$$[M^\gamma \rightarrow \bullet \nu, i, i, false]$$

Al posibilitar el esquema **buE<sup>k</sup>** este tipo de ítems, se provoca un aumento en el número de ítems deducidos y, por consiguiente, una merma en el comportamiento práctico del analizador. Por ello proponemos modificar el dominio y los pasos deductivos de dicho esquema

para evitar que se genere este tipo de ítems, obteniendo como resultado el esquema  $\mathbf{buLC}^i$ .

El dominio del esquema  $\mathbf{buLC}^i$  es:

$$\mathcal{I}_{\mathbf{buLC}^i} = \mathcal{I}_{\mathbf{buLC}^i}^{(i)} \cup \mathcal{I}_{\mathbf{buLC}^i}^{(ii)}$$

$$\mathcal{I}_{\mathbf{buLC}^i}^{(i)} = \{[M^\gamma \rightarrow P^\gamma \delta \bullet \nu, i, j, false]\}$$

donde  $M^\gamma \rightarrow P^\gamma \delta \nu \in \mathcal{P}(\gamma)$  y  $\gamma \in \mathbf{I} \cup \mathbf{A}$ . Los índices  $0 \leq i \leq j$  establecen las posiciones de la cadena de entrada que delimitan el fragmento reconocido por  $P^\gamma \delta$ . Para evitar que el punto aparezca a la derecha de la producción, se exige que  $\nu \neq \varepsilon$ . El valor *false* indica que no se ha completado ninguna adjunción en el nodo  $M^\gamma$ .

El subconjunto  $\mathcal{I}_{\mathbf{buLC}^i}^{(ii)}$  es igual al del esquema  $\mathbf{buE}^k$ :

$$\mathcal{I}_{\mathbf{buLC}^i}^{(ii)} = \mathcal{I}_{\mathbf{buE}^k}^{(ii)}$$

Los pasos deductivos del esquema son:

$$\mathcal{D}_{\mathbf{buLC}^i} = \mathcal{D}_{\mathbf{buLC}^i}^{\mathbf{LC}_t} \cup \mathcal{D}_{\mathbf{buLC}^i}^{\mathbf{LC}_\varepsilon} \cup \mathcal{D}_{\mathbf{buLC}^i}^{\mathbf{LC}_{\text{subs}}} \cup \mathcal{D}_{\mathbf{buLC}^i}^{\mathbf{LC}_n} \cup \mathcal{D}_{\mathbf{buLC}^i}^{\mathbf{LAdj}_t} \cup \mathcal{D}_{\mathbf{buLC}^i}^{\mathbf{LAdj}_\varepsilon} \cup \mathcal{D}_{\mathbf{buLC}^i}^{\mathbf{LAdj}_{\text{subs}}} \cup$$

$$\mathcal{D}_{\mathbf{buLC}^i}^{\mathbf{LAdj}_n} \cup \mathcal{D}_{\mathbf{buLC}^i}^\varepsilon \cup \mathcal{D}_{\mathbf{buLC}^i}^{\mathbf{Scan}} \cup \mathcal{D}_{\mathbf{buLC}^i}^{\mathbf{Comp}} \cup \mathcal{D}_{\mathbf{buLC}^i}^{\mathbf{RAAdj}} \cup \mathcal{D}_{\mathbf{buLC}^i}^{\mathbf{Subs}}$$

$$\mathcal{D}_{\mathbf{buLC}^i}^{\mathbf{LC}_t} = \frac{[a, j, j+1]}{[O^\gamma \rightarrow M^\gamma \bullet \nu, j, j+1, false]} \quad \begin{array}{l} \mathbf{nil} \in \text{ladj}(O^\gamma) \\ \text{label}(M^\gamma) = a \end{array}$$

$$\mathcal{D}_{\mathbf{buLC}^i}^{\mathbf{LC}_\varepsilon} = \frac{}{[O^\gamma \rightarrow M^\gamma \bullet \nu, j, j, false]} \quad \begin{array}{l} \mathbf{nil} \in \text{ladj}(O^\gamma) \\ \text{label}(M^\gamma) = \varepsilon \text{ o } \text{label}(M^\gamma) = \perp \end{array}$$

$$\mathcal{D}_{\mathbf{buLC}^i}^{\mathbf{LC}_{\text{subs}}} = \frac{[\top \rightarrow \mathbf{R}^\alpha \bullet, j, k, false]}{[O^\gamma \rightarrow M^\gamma \bullet \nu, j, k, false]} \quad \begin{array}{l} \mathbf{nil} \in \text{ladj}(O^\gamma) \\ \alpha \in \text{subst}(M^\gamma) \end{array}$$

$$\mathcal{D}_{\mathbf{buLC}^i}^{\mathbf{LC}_n} = \frac{[O^\gamma \rightarrow \delta \bullet, j, k, \text{radj}]}{[Q^\gamma \rightarrow O^\gamma \bullet \nu, j, k, false]} \quad \mathbf{nil} \in \text{ladj}(Q^\gamma)$$

donde se debe cumplir que si  $\text{radj} = false$  entonces  $\mathbf{nil} \in \text{radj}(O^\gamma)$ .

$$\mathcal{D}_{\mathbf{buLC}^i}^{\mathbf{LAdj}_t} = \frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\beta \bullet, i, j, false] \\ [a, j, j+1] \end{array}}{[Q^\gamma \rightarrow M^\gamma \bullet \nu, i, j+1, false]} \quad \begin{array}{l} \beta \in \text{ladj}(Q^\gamma) \\ \text{label}(M^\gamma) = a \end{array}$$

$$\mathcal{D}_{\mathbf{buLC}^i}^{\mathbf{LAdj}_\varepsilon} = \frac{[\top \rightarrow \mathbf{R}^\beta \bullet, i, j, false]}{[Q^\gamma \rightarrow M^\gamma \bullet \nu, i, j, false]} \quad \begin{array}{l} \beta \in \text{ladj}(Q^\gamma) \\ \text{label}(M^\gamma) = \varepsilon \text{ o } \text{label}(M^\gamma) = \perp \end{array}$$

$$\mathcal{D}_{\text{buLC}^i}^{\text{LAdj}_{\text{subs}}} = \frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^{\beta \bullet}, i, j, \text{false}] \\ [\top \rightarrow \mathbf{R}^{\alpha \bullet}, j, k, \text{false}] \end{array}}{[Q^\gamma \rightarrow M^\gamma \bullet \nu, i, k, \text{false}]} \quad \begin{array}{l} \beta \in \text{ladj}(Q^\gamma) \\ \alpha \in \text{subst}(M^\gamma) \end{array}$$

$$\mathcal{D}_{\text{buLC}^i}^{\text{LAdj}_n} = \frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^{\beta \bullet}, i, j, \text{false}] \\ [O^\gamma \rightarrow \delta \bullet, j, k, \text{radj}] \end{array}}{[Q^\gamma \rightarrow O^\gamma \bullet \nu, i, k, \text{false}]} \quad \beta \in \text{ladj}(Q^\gamma)$$

donde se debe cumplir que si  $\text{radj} = \text{false}$  entonces  $\text{nil} \in \text{radj}(O^\gamma)$ .

$$\mathcal{D}_{\text{buLC}^i}^{\text{Scan}} = \frac{\begin{array}{l} [a, j, j+1] \\ [N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \nu, i, j, \text{false}] \end{array}}{[N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, i, j+1, \text{false}]} \quad \text{label}(M^\gamma) = a$$

$$\mathcal{D}_{\text{buLC}^i}^\varepsilon = \frac{[N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \nu, i, j, \text{false}]}{[N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, i, j, \text{false}]} \quad \text{label}(M^\gamma) = \varepsilon \text{ o } \text{label}(M^\gamma) = \perp$$

$$\mathcal{D}_{\text{buLC}^i}^{\text{Comp}} = \frac{\begin{array}{l} [M^\gamma \rightarrow \nu \bullet, j, k, \text{radj}] \\ [N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \nu, i, j, \text{false}] \end{array}}{[N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, i, k, \text{false}]}$$

donde se debe cumplir que si  $\text{radj} = \text{false}$  entonces  $\text{nil} \in \text{radj}(M^\gamma)$ .

El paso de completación de adjunción derecha es igual a su homónimo del esquema  $\text{buE}^k$ :

$$\mathcal{D}_{\text{buLC}^i}^{\text{RAdj}} = \mathcal{D}_{\text{buE}^k}^{\text{RAdj}}$$

$$\mathcal{D}_{\text{buLC}^i}^{\text{Subs}} = \frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^{\alpha \bullet}, j, k, \text{false}] \\ [N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \nu, i, j, \text{false}] \end{array}}{[N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, i, k, \text{false}]} \quad \alpha \in \text{subst}(M^\gamma)$$

La eliminación del dominio de un determinado tipo de ítems provoca que tengamos que reescribir el paso  $\mathcal{D}_{\text{buE}^k}^{\text{Ini}}$ , obteniendo los pasos  $\mathcal{D}_{\text{buLC}^i}^{\text{LC}_t}$ ,  $\mathcal{D}_{\text{buLC}^i}^{\text{LC}_\varepsilon}$ ,  $\mathcal{D}_{\text{buLC}^i}^{\text{LC}_{\text{subs}}}$  y  $\mathcal{D}_{\text{buLC}^i}^{\text{LC}_n}$ , que se aplican cuando el símbolo que está a la izquierda de la regla (la esquina izquierda) es un terminal, la cadena vacía, un nodo de sustitución o un no terminal, respectivamente. Actuamos de forma análoga con la regla  $\mathcal{D}_{\text{buE}^k}^{\text{LAdj}}$  para obtener las cuatro reglas:  $\mathcal{D}_{\text{buLC}^i}^{\text{LAdj}_t}$ ,  $\mathcal{D}_{\text{buLC}^i}^{\text{LAdj}_\varepsilon}$ ,  $\mathcal{D}_{\text{buLC}^i}^{\text{LAdj}_{\text{subs}}}$  y  $\mathcal{D}_{\text{buLC}^i}^{\text{LAdj}_n}$  en este esquema. El resto de pasos funcionan de la misma forma que sus homónimos en el esquema  $\text{buE}^k$ , aunque los pasos  $\mathcal{D}_{\text{buLC}^i}^{\text{Scan}}$ ,  $\mathcal{D}_{\text{buLC}^i}^\varepsilon$ ,  $\mathcal{D}_{\text{buLC}^i}^{\text{Comp}}$  y  $\mathcal{D}_{\text{buLC}^i}^{\text{Subs}}$  sólo se aplican a nodos que no son hijos izquierdos.

El conjunto de ítems finales es idéntico al del esquema  $\text{buE}^k$ :

$$\mathcal{F}_{\text{buLC}^i} = \mathcal{F}_{\text{buE}^k}$$

## 9.2 La relación Left Corner en TIG

En las siguientes secciones vamos a definir esquemas que usan una extensión del concepto de relación de Left Corner en CFG para filtrar las predicciones en el analizador basado en el algoritmo de Earley para TIG. La complejidad temporal de todos estos algoritmos se mantienen en la cota de  $\mathcal{O}(n^3)$ , pero mejoran sus prestaciones mediante una reducción en el número de ítems deducidos. Antes de describir los nuevos analizadores necesitamos definir el concepto de relación de Left Corner en las TIG.

**Definición 9.1** *Relación de Left Corner (LC) en los árboles elementales de TIG*

La esquina izquierda de un nodo  $O^\gamma$  es su hijo izquierdo  $P^\gamma$  si y sólo si  $\text{ladj}(P^\gamma) = \{\text{nil}\}$ .

La relación Left Corner  $>_\ell$  en  $V_N \times \{V_N \cup V_T \cup \{\varepsilon, \perp\}\}$  se define como

$O^\gamma >_\ell P^\gamma$  si hay una producción  $O^\gamma \rightarrow P^\gamma \nu \in \mathcal{P}(\gamma)$  y  $\text{ladj}(P^\gamma) = \{\text{nil}\}$ .

La clausura reflexiva y transitiva de  $>_\ell$  la denotaremos como  $>_\ell^*$ .

En un abuso de notación, vamos a denotar como  $P^\gamma >_\ell \Delta$  si hay una producción  $O^\gamma \rightarrow P^\gamma \nu \in \mathcal{P}(\gamma)$  y existe un  $\beta$  tal que  $\beta \in \text{ladj}(P^\gamma)$ .

Al igual que en las TAG, la relación LC en las TIG no va más allá de los límites de un árbol elemental. Es importante señalar que toda relación de LC en las TIG comienza en un nodo etiquetado con un símbolo no terminal y finaliza en: (1) un nodo adjuntable por la izquierda; (2) un nodo etiquetado con un símbolo terminal,  $\varepsilon$  o  $\perp$ ; (3) un nodo marcado para sustitución.

Las diferencias de esta definición respecto a la que vimos para TAG son fundamentalmente dos:

- Al introducir la operación de sustitución, en las fronteras de los árboles elementales de las TIG pueden aparecer nodos etiquetados con símbolos no terminales marcados para sustitución. Ésto nos obliga a introducir este caso como una posibilidad de finalización en las relaciones LC.
- En las TIG distinguimos dos tipos de adjunciones: izquierda y derecha. Las segundas no afectan a las relaciones de esquina izquierda, puesto que sólo introducen contextos derechos en los subárboles. Sin embargo, las primeras si rompen las relaciones LC. Por esta causa, únicamente tenemos en cuenta los nodos adjuntables por la izquierda como fin de una relación LC.

## 9.3 Esquema tipo Left Corner con ítems predictivos

En este esquema, al que denominaremos  $\text{pLC}^i$ , los pasos predictivos en el esquema Earley<sup>i</sup> son reemplazados por objetivos que se intentan satisfacer de forma ascendente. La fase ascendente del proceso de reconocimiento es guiada hacia el correspondiente objetivo mediante la relación LC. En el dominio del esquema  $\text{pLC}^i$  vamos a distinguir dos tipos de

ítems: predictivos y *left corner*, cuya semántica es similar a la que vimos en los ítems de la misma denominación en el esquema **pLC** para TAG.

El conjunto de ítems de **pLC**<sup>i</sup> se define mediante:

$$\mathcal{I}_{\text{pLC}^i} = \mathcal{I}_{\text{pLC}^i}^{(i)} \cup \mathcal{I}_{\text{pLC}^i}^{(ii)} \cup \mathcal{I}_{\text{pLC}^i}^{(iii)} \cup \mathcal{I}_{\text{pLC}^i}^{(iv)}$$

El conjunto de ítems predictivos es:

$$\mathcal{I}_{\text{pLC}^i}^{(i)} = \{\{M^\gamma, j\}\}$$

donde  $M^\gamma \in V_N$  y  $\gamma \in I \cup A$ . El índice  $0 \leq j$  establece la posición de la cadena de entrada donde se inicia el reconocimiento del nodo  $M^\gamma$ .

Y el conjunto de ítems *left corner* viene definido por los tres siguientes subconjuntos.

$$\mathcal{I}_{\text{pLC}^i}^{(ii)} = \{\{C^\gamma; M^\gamma \rightarrow P^\gamma \delta \bullet \nu, i, j, false\}\}$$

donde  $M^\gamma \rightarrow P^\gamma \delta \nu \in \mathcal{P}(\gamma)$  y  $\gamma \in I \cup A$ . Los índices  $0 \leq i \leq j$  establecen las posiciones de la cadena de entrada que delimitan el fragmento reconocido por  $P^\gamma \delta$ .  $C^\gamma$  es el ancla de una relación LC y domina al nodo  $M^\gamma$  ( $C^\gamma >_\ell^* M^\gamma$ ). Para evitar que el punto aparezca a la derecha de la producción, se exige que  $\nu \neq \varepsilon$ . El valor *false* indica que no se ha completado ninguna adjunción en el nodo  $M^\gamma$ .

$$\mathcal{I}_{\text{pLC}^i}^{(iii)} = \{\{C^\gamma; M^\gamma \rightarrow \nu \bullet, i, j, radj\}\}$$

donde  $M^\gamma \rightarrow \nu \in \mathcal{P}(\gamma)$  y  $\gamma \in I \cup A$ . Los índices  $0 \leq i \leq j$  establecen las posiciones de la cadena de entrada que delimitan el fragmento reconocido por  $\nu$ .  $C^\gamma$  es el ancla de una relación LC y domina al nodo  $M^\gamma$  ( $C^\gamma >_\ell^* M^\gamma$ ). Y *radj* = *false* si no se completó ninguna adjunción derecha sobre el nodo  $M^\gamma$  y *radj* = *true* si se completó una adjunción derecha sobre  $M^\gamma$ .

$$\mathcal{I}_{\text{pLC}^i}^{(iv)} = \{\{C^\gamma; M^\gamma \rightarrow \bullet P^\gamma \nu, j, j, false\}\}$$

donde  $M^\gamma \rightarrow P^\gamma \nu \in \mathcal{P}(\gamma)$  y  $\gamma \in I \cup A$ . El índice  $0 \leq j$  establece la posición en la cadena de entrada donde comienza el reconocimiento del subárbol que domina  $M^\gamma$ .  $C^\gamma$  es el ancla de una relación LC y domina al nodo  $M^\gamma$  ( $C^\gamma >_\ell^* M^\gamma$ ). El valor *false* indica que no se ha completado ninguna adjunción en el nodo  $M^\gamma$ . Además se debe cumplir que  $P^\gamma >_\ell \Delta$  o exista un  $\alpha$  tal que  $\alpha \in \text{subst}(P^\gamma)$ . Es decir, el punto sólo aparece al comienzo de la regla cuando el hijo izquierdo sea un nodo adjuntable por la izquierda o un nodo de sustitución, con objeto de lanzar el reconocimiento del árbol auxiliar izquierdo o del árbol inicial, respectivamente.

Con respecto al conjunto de pasos deductivos, definimos subconjuntos para reconocimiento y completación similares a los del esquema **Earley**<sup>i</sup> para TIG. La relación de esquina izquierda se aplicará a cinco casos de predicción: inicial, subárbol, pie, adjunción izquierda y adjunción derecha. Obsérvese que se introduce un filtro sobre la predicciones del pie,



aunque este tipo de predicción aparentemente no existe en el esquema **Earley**<sup>i</sup>. Pero si nos fijamos con atención, el paso de compleción de adjunción izquierda  $\mathcal{D}_{\text{Earley}^i}^{\text{LAdjComp}}$  hace una doble función: completa la adjunción izquierda e inicia el reconocimiento del subárbol escindido. Luego podemos aplicar un filtro sobre esta predicción de pie implícita.

Los pasos de esquina izquierda vienen en tres variedades, según el tipo de nodo en que finalice la relación: terminal, cadena vacía o nodo *bottom*, y no terminal. Los nodos etiquetados con  $\varepsilon$  y  $\perp$  se incluyen en el mismo caso porque los nodos  $\perp$  en las TIG se comportan igual que una cadena vacía, debido que no dominan nada. El último caso es necesario cuando la esquina izquierda es un nodo de adjunción izquierda o está marcado para sustitución.

Los pasos deductivos del esquema son:

$$\begin{aligned} \mathcal{D}_{\text{pLC}^i} = & \mathcal{D}_{\text{pLC}^i}^{\text{LI}_t} \cup \mathcal{D}_{\text{pLC}^i}^{\text{LI}_\varepsilon} \cup \mathcal{D}_{\text{pLC}^i}^{\text{LI}_{\text{pre}}} \cup \mathcal{D}_{\text{pLC}^i}^{\text{Scan}} \cup \mathcal{D}_{\text{pLC}^i}^\varepsilon \cup \mathcal{D}_{\text{pLC}^i}^{\text{LC}_t} \cup \\ & \mathcal{D}_{\text{pLC}^i}^{\text{LC}_\varepsilon} \cup \mathcal{D}_{\text{pLC}^i}^{\text{LC}_{\text{pre}}} \cup \mathcal{D}_{\text{pLC}^i}^{\text{LC}_n} \cup \mathcal{D}_{\text{pLC}^i}^{\text{Pre}} \cup \mathcal{D}_{\text{pLC}^i}^{\text{Comp}} \cup \mathcal{D}_{\text{pLC}^i}^{\text{LA}_t} \cup \\ & \mathcal{D}_{\text{pLC}^i}^{\text{LA}_\varepsilon} \cup \mathcal{D}_{\text{pLC}^i}^{\text{LA}_{\text{pre}}} \cup \mathcal{D}_{\text{pLC}^i}^{\text{LF}_t} \cup \mathcal{D}_{\text{pLC}^i}^{\text{LF}_\varepsilon} \cup \mathcal{D}_{\text{pLC}^i}^{\text{LF}_{\text{pre}}} \cup \mathcal{D}_{\text{pLC}^i}^{\text{RA}_\varepsilon} \cup \\ & \mathcal{D}_{\text{pLC}^i}^{\text{RA}_{\text{AdjComp}}} \cup \mathcal{D}_{\text{pLC}^i}^{\text{LS}_t} \cup \mathcal{D}_{\text{pLC}^i}^{\text{LS}_\varepsilon} \cup \mathcal{D}_{\text{pLC}^i}^{\text{LS}_{\text{pre}}} \cup \mathcal{D}_{\text{pLC}^i}^{\text{SubsComp}} \end{aligned}$$

### Filtrado de inicio

$$\mathcal{D}_{\text{pLC}^i}^{\text{LI}_t} = \frac{[a, 0, 1]}{[\top; O^\alpha \rightarrow P^\alpha \bullet \nu, 0, 1, \text{false}]} \quad \begin{array}{l} \alpha \in \mathbf{I} \\ \text{label}(\mathbf{R}^\alpha) = S \\ \text{label}(P^\alpha) = a \end{array}$$

$$\mathcal{D}_{\text{pLC}^i}^{\text{LI}_\varepsilon} = \frac{[a, 0, 0]}{[\top; O^\alpha \rightarrow P^\alpha \bullet \nu, 0, 0, \text{false}]} \quad \begin{array}{l} \alpha \in \mathbf{I} \\ \text{label}(\mathbf{R}^\alpha) = S \\ \text{label}(P^\alpha) = \varepsilon \end{array}$$

$$\mathcal{D}_{\text{pLC}^i}^{\text{LI}_{\text{pre}}} = \frac{[a, 0, 0]}{[\top; O^\alpha \rightarrow \bullet P^\alpha \nu, 0, 0, \text{false}]} \quad \begin{array}{l} \alpha \in \mathbf{I} \\ \text{label}(\mathbf{R}^\alpha) = S \\ P^\alpha >_\ell \Delta \circ \exists \alpha' \in \text{subst}(P^\alpha) \end{array}$$

En el paso  $\mathcal{D}_{\text{pLC}^i}^{\text{LI}_\varepsilon}$  no incluimos la condición  $\text{label}(P^\alpha) = \perp$  porque este paso se aplica exclusivamente a árboles iniciales.

### Reconocimiento

Los pasos de reconocimiento son similares a los del esquema **Earley**<sup>i</sup> pero sólo se aplican a nodos que no son hijos izquierdos.

$$\mathcal{D}_{\text{pLC}^i}^{\text{Scan}} = \frac{[a, j, j+1]}{[C^\gamma; N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \nu, i, j, \text{false}]} \quad \text{label}(M^\gamma) = a$$

$$\mathcal{D}_{\text{pLC}^i}^{\text{Scan}} = \frac{[a, j, j+1]}{[C^\gamma; N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, i, j+1, \text{false}]} \quad \text{label}(M^\gamma) = a$$

$$\mathcal{D}_{\text{pLC}^i}^\varepsilon = \frac{[C^\gamma; N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \nu, i, j, false]}{[C^\gamma; N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, i, j, false]} \quad \text{label}(M^\gamma) = \varepsilon$$

**Filtrado de predicción de subárbol**

$$\mathcal{D}_{\text{pLC}^i}^{\text{LC}_t} = \frac{\begin{array}{c} [M^\gamma, j] \\ [a, j, j+1] \end{array}}{[M^\gamma; O^\gamma \rightarrow P^\gamma \bullet \nu, j, j+1, false]} \quad \begin{array}{l} \text{nil} \in \text{ladj}(M^\gamma) \\ \text{label}(P^\gamma) = a \end{array}$$

$$\mathcal{D}_{\text{pLC}^i}^{\text{LC}_\varepsilon} = \frac{[M^\gamma, j]}{[M^\gamma; O^\gamma \rightarrow P^\gamma \bullet \nu, j, j, false]} \quad \begin{array}{l} \text{nil} \in \text{ladj}(M^\gamma) \\ \text{label}(P^\gamma) = \varepsilon \text{ o } \text{label}(P^\gamma) = \perp \end{array}$$

$$\mathcal{D}_{\text{pLC}^i}^{\text{LC}_{\text{pre}}} = \frac{[M^\gamma, j]}{[M^\gamma; O^\gamma \rightarrow \bullet P^\gamma \nu, j, j, false]} \quad \begin{array}{l} \text{nil} \in \text{ladj}(M^\gamma) \\ P^\gamma >_\ell \Delta \text{ o } \exists \alpha \in \text{subst}(P^\gamma) \end{array}$$

**Compleción en la esquina izquierda**

El paso  $\mathcal{D}_{\text{pLC}^i}^{\text{LC}_n}$  es el que lleva a cabo las completaciones en los nodos que han sido filtrados por una relación LC.

$$\mathcal{D}_{\text{pLC}^i}^{\text{LC}_n} = \frac{[M^\gamma; O^\gamma \rightarrow \nu \bullet, j, k, \text{radj}]}{[M^\gamma; Q^\gamma \rightarrow O^\gamma \bullet \omega, j, k, false]} \quad M^\gamma \neq O^\gamma$$

donde se debe cumplir que si  $\text{radj} = false$  entonces  $\text{nil} \in \text{radj}(O^\gamma)$ .

**Predicción**

Este paso realiza la predicción de los nodos que dominan relaciones LC, excepto los nodos raíces de los árboles elementales.

$$\mathcal{D}_{\text{pLC}^i}^{\text{Pre}} = \frac{[C^\gamma; N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[M^\gamma, j]} \quad \text{label}(M^\gamma) \in V_N$$

**Compleción de subárbol**

$$\mathcal{D}_{\text{pLC}^i}^{\text{Comp}} = \frac{\begin{array}{c} [M^\gamma; M^\gamma \rightarrow \nu \bullet, j, k, \text{radj}] \\ [C^\gamma; N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false] \end{array}}{[C^\gamma; N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k, false]}$$

donde se debe cumplir que si  $\text{radj} = false$  entonces  $\text{nil} \in \text{radj}(O^\gamma)$ .

**Filtrado de predicción de adjunción izquierda**

$$\mathcal{D}_{\text{pLC}^i}^{\text{LA}_t} = \frac{\begin{array}{c} [M^\gamma, j] \\ [a, j, j+1] \end{array}}{[\top; O^\beta \rightarrow P^\beta \bullet \nu, j, j+1, false]} \quad \begin{array}{l} \beta \in \text{ladj}(M^\gamma) \\ \text{label}(P^\beta) = a \end{array}$$

$$\mathcal{D}_{\text{pLC}^i}^{\text{LA}_\varepsilon} = \frac{[M^\gamma, j]}{[\top; O^\beta \rightarrow P^\beta \bullet \nu, j, j, false]} \quad \begin{array}{l} \beta \in \text{ladj}(M^\gamma) \\ \text{label}(P^\beta) = \varepsilon \end{array}$$

$$\mathcal{D}_{\text{pLC}^i}^{\text{LA}_{\text{pre}}} = \frac{[M^\gamma, j]}{[\top; O^\beta \rightarrow \bullet P^\beta \nu, j, j, false]} \quad \begin{array}{l} \beta \in \text{ladj}(M^\gamma) \\ P^\beta >_\ell \Delta \text{ o } \exists \alpha \in \text{subst}(P^\gamma) \end{array}$$

En el paso  $\mathcal{D}_{\text{pLC}^i}^{\text{LA}_\varepsilon}$  no incluimos la condición  $\text{label}(P^\beta) = \perp$  porque en un árbol auxiliar izquierdo no se puede dar el caso  $\top >_\ell^* \perp$ .

### Filtrado de predicción de pie

Este conjunto de pasos son el resultado de aplicar un filtro LC al paso  $\mathcal{D}_{\text{Earley}^i}^{\text{LAdjComp}}$ , el cual, como dijimos antes, cumple una función de predicción de pie.

$$\mathcal{D}_{\text{pLC}^i}^{\text{LF}_t} = \frac{\begin{array}{c} [a, k, k+1] \\ [\top; \top \rightarrow \mathbf{R}^\beta \bullet, j, k, false] \\ [M^\gamma, j] \end{array}}{[M^\gamma; O^\gamma \rightarrow P^\gamma \bullet \nu, k, k+1, false]} \quad \begin{array}{l} \beta \in \text{ladj}(M^\gamma) \\ \text{label}(P^\gamma) = a \end{array}$$

$$\mathcal{D}_{\text{pLC}^i}^{\text{LF}_\varepsilon} = \frac{\begin{array}{c} [\top; \top \rightarrow \mathbf{R}^\beta \bullet, j, k, false] \\ [M^\gamma, j] \end{array}}{[M^\gamma; O^\gamma \rightarrow P^\gamma \bullet \nu, k, k, false]} \quad \begin{array}{l} \beta \in \text{ladj}(M^\gamma) \\ \text{label}(P^\gamma) = \varepsilon \text{ o } \text{label}(P^\gamma) = \perp \end{array}$$

$$\mathcal{D}_{\text{pLC}^i}^{\text{LF}_{\text{pre}}} = \frac{\begin{array}{c} [\top; \top \rightarrow \mathbf{R}^\beta \bullet, j, k, false] \\ [M^\gamma, j] \end{array}}{[M^\gamma; O^\gamma \rightarrow P^\gamma \bullet \nu, k, k, false]} \quad \begin{array}{l} \beta \in \text{ladj}(M^\gamma) \\ P^\gamma >_\ell \Delta \text{ o } \exists \alpha \in \text{subst}(P^\gamma) \end{array}$$

### Filtrado de predicción de adjunción derecha

Teniendo en cuenta que los árboles auxiliares derechos se caracterizan porque a la izquierda de la espina no se pueden adjuntar árboles auxiliares izquierdos y en esa parte de su frontera sólo pueden aparecer nodos etiquetados con  $\varepsilon$ , ésto reduce la finalización del filtro LC sobre la raíces de los árboles auxiliares derechos a un solo caso:  $\varepsilon$  o  $\perp$ .

$$\mathcal{D}_{\text{pLC}^i}^{\text{RA}_\varepsilon} = \frac{[C^\gamma; M^\gamma \rightarrow \delta \bullet, k, l, false]}{[\top; O^\beta \rightarrow P^\beta \bullet \nu, l, l, false]} \quad \begin{array}{l} \beta \in \text{radj}(M^\gamma) \\ \text{label}(P^\beta) = \varepsilon \text{ o } \text{label}(P^\beta) = \perp \end{array}$$

### Completación de adjunción derecha

$$\mathcal{D}_{\text{pLC}^i}^{\text{RAdjComp}} = \frac{[\top; \top \rightarrow \mathbf{R}^\beta \bullet, j, k, false] \quad [C^\gamma; M^\gamma \rightarrow \nu \bullet, i, j, false]}{[C^\gamma; M^\gamma \rightarrow \nu \bullet, i, k, true]} \quad \beta \in \text{radj}(M^\gamma)$$

Filtrado de predicción de sustitución

$$\mathcal{D}_{\text{pLC}^i}^{\text{LS}_t} = \frac{[M^\gamma, j] \quad [a, j, j+1]}{[\top; O^\alpha \rightarrow P^\alpha \bullet \nu, j, j+1, false]} \quad \begin{array}{l} \alpha \in \text{subst}(M^\gamma) \\ \text{label}(P^\alpha) = a \end{array}$$

$$\mathcal{D}_{\text{pLC}^i}^{\text{LS}_\varepsilon} = \frac{[M^\gamma, j]}{[\top; O^\alpha \rightarrow P^\alpha \bullet \nu, j, j, false]} \quad \begin{array}{l} \alpha \in \text{subst}(M^\gamma) \\ \text{label}(P^\alpha) = \varepsilon \end{array}$$

$$\mathcal{D}_{\text{pLC}^i}^{\text{LS}_{\text{pre}}} = \frac{[M^\gamma, j]}{[\top; O^\alpha \rightarrow \bullet P^\alpha \nu, j, j, false]} \quad \begin{array}{l} \alpha \in \text{subst}(M^\gamma) \\ P^\alpha >_\ell \Delta \circ \exists \alpha' \in \text{subst}(P^\alpha) \end{array}$$

Compleción de sustitución

$$\mathcal{D}_{\text{pLC}^i}^{\text{SubsComp}} = \frac{[\top; \top \rightarrow \mathbf{R}^\alpha \bullet, j, k, false] \quad [C^\gamma; N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[C^\gamma; N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k, false]} \quad \alpha \in \text{subst}(M^\gamma)$$

El conjunto de ítems finales del esquema se define mediante:

$$\mathcal{F}_{\text{pLC}^i} = \{[\top; \top \rightarrow \mathbf{R}^\alpha \bullet, 0, n, false] \mid \alpha \in I, \text{label}(\mathbf{R}^\alpha) = S\}$$

## 9.4 Esquema tipo Left Corner con ítems simplificados

Vamos a derivar un nuevo esquema, al que denominaremos  $\text{sLC}^i$ , simplificando los ítems del esquema  $\text{pLC}^i$  de forma similar a la que lo hicimos para derivar el esquema  $\text{sLC}$  para TAG. Es decir, al conjunto de ítems *left corner* de  $\text{pLC}^i$  le eliminamos la parte predictiva y lo convertimos en ítems Earley en el esquema  $\text{sLC}^i$ . Evidentemente, esta modificación del dominio conlleva una adaptación del conjunto de pasos deductivos.

El dominio del esquema  $\text{sLC}^i$  presenta dos tipos de ítems: predictivos y Earley. Y se define de la siguiente manera:

$$\mathcal{I}_{\text{sLC}^i} = \mathcal{I}_{\text{sLC}^i}^{(i)} \cup \mathcal{I}_{\text{sLC}^i}^{(ii)} \cup \mathcal{I}_{\text{sLC}^i}^{(iii)} \cup \mathcal{I}_{\text{sLC}^i}^{(iv)}$$

El conjunto de ítems predictivos es igual al del esquema anterior:

$$\mathcal{I}_{\text{sLC}^i}^{(i)} = \mathcal{I}_{\text{pLC}^i}^{(i)}$$

El conjunto de ítems Earley es similar al del esquema **Earley**<sup>i</sup>, pero cierto tipo de ítems con el punto al comienzo de las reglas son filtrados del dominio. Dicho conjunto viene definido por los siguientes subconjuntos:

$$\mathcal{I}_{\text{sLC}^i}^{(\text{ii})} = \{[M^\gamma \rightarrow P^\gamma \delta \bullet \nu, i, j, \text{false}]\}$$

donde  $M^\gamma \rightarrow P^\gamma \delta \nu \in \mathcal{P}(\gamma)$  y  $\gamma \in I \cup A$ . Los índices  $0 \leq i \leq j$  establecen las posiciones de la cadena de entrada que delimitan el fragmento reconocido por  $P^\gamma \delta$ . Para evitar que el punto aparezca a la derecha de la producción, se exige que  $\nu \neq \varepsilon$ . El valor *false* indica que no se ha completado ninguna adjunción en el nodo  $M^\gamma$ .

$$\mathcal{I}_{\text{sLC}^i}^{(\text{iii})} = \{[M^\gamma \rightarrow \nu \bullet, i, j, \text{radj}]\}$$

donde  $M^\gamma \rightarrow \nu \in \mathcal{P}(\gamma)$  y  $\gamma \in I \cup A$ . Los índices  $0 \leq i \leq j$  establecen las posiciones de la cadena de entrada que delimitan el fragmento reconocido por  $\nu$ . Y *radj* = *false* si no se completó ninguna adjunción derecha sobre el nodo  $M^\gamma$  y *radj* = *true* si se completó una adjunción derecha sobre  $M^\gamma$ .

$$\mathcal{I}_{\text{sLC}^i}^{(\text{iv})} = \{[M^\gamma \rightarrow \bullet P^\gamma \nu, j, j, \text{false}]\}$$

donde  $M^\gamma \rightarrow P^\gamma \nu \in \mathcal{P}(\gamma)$  y  $\gamma \in I \cup A$ . El índice  $0 \leq j$  establece la posición en la cadena de entrada donde comienza el reconocimiento del subárbol que domina  $M^\gamma$ . El valor *false* indica que no se ha completado ninguna adjunción en el nodo  $M^\gamma$ . Además se debe cumplir que  $P^\gamma >_\ell \Delta$  o exista un  $\alpha$  tal que  $\alpha \in \text{subst}(P^\gamma)$ .

Los pasos deductivos del esquema son:

$$\begin{aligned} \mathcal{D}_{\text{sLC}^i} = & \mathcal{D}_{\text{sLC}^i}^{\text{LI}_t} \cup \mathcal{D}_{\text{sLC}^i}^{\text{LI}_\varepsilon} \cup \mathcal{D}_{\text{sLC}^i}^{\text{LI}_{\text{pre}}} \cup \mathcal{D}_{\text{sLC}^i}^{\text{Scan}} \cup \mathcal{D}_{\text{sLC}^i}^\varepsilon \cup \mathcal{D}_{\text{sLC}^i}^{\text{LC}_t} \cup \\ & \mathcal{D}_{\text{sLC}^i}^{\text{LC}_\varepsilon} \cup \mathcal{D}_{\text{sLC}^i}^{\text{LC}_{\text{pre}}} \cup \mathcal{D}_{\text{sLC}^i}^{\text{LC}_n} \cup \mathcal{D}_{\text{sLC}^i}^{\text{Pre}} \cup \mathcal{D}_{\text{sLC}^i}^{\text{Comp}} \cup \mathcal{D}_{\text{sLC}^i}^{\text{LA}_t} \cup \\ & \mathcal{D}_{\text{sLC}^i}^{\text{LA}_\varepsilon} \cup \mathcal{D}_{\text{sLC}^i}^{\text{LA}_{\text{pre}}} \cup \mathcal{D}_{\text{sLC}^i}^{\text{LF}_t} \cup \mathcal{D}_{\text{sLC}^i}^{\text{LF}_\varepsilon} \cup \mathcal{D}_{\text{sLC}^i}^{\text{LF}_{\text{pre}}} \cup \mathcal{D}_{\text{sLC}^i}^{\text{RA}_\varepsilon} \cup \\ & \mathcal{D}_{\text{sLC}^i}^{\text{RA}_{\text{Comp}}} \cup \mathcal{D}_{\text{sLC}^i}^{\text{LS}_t} \cup \mathcal{D}_{\text{sLC}^i}^{\text{LS}_\varepsilon} \cup \mathcal{D}_{\text{sLC}^i}^{\text{LS}_{\text{pre}}} \cup \mathcal{D}_{\text{sLC}^i}^{\text{SubsComp}} \end{aligned}$$

### Filtrado de inicio

$$\mathcal{D}_{\text{sLC}^i}^{\text{LI}_t} = \frac{[a, 0, 1]}{[O^\alpha \rightarrow P^\alpha \bullet \nu, 0, 1, \text{false}]} \quad \begin{array}{l} \top >_\ell^* O^\alpha \\ \alpha \in I \\ \text{label}(\mathbf{R}^\alpha) = S \\ \text{label}(P^\alpha) = a \end{array}$$

$$\mathcal{D}_{\text{sLC}^i}^{\text{LI}_\varepsilon} = \frac{[O^\alpha \rightarrow P^\alpha \bullet \nu, 0, 0, \text{false}]}{[O^\alpha \rightarrow P^\alpha \bullet \nu, 0, 0, \text{false}]} \quad \begin{array}{l} \top >_\ell^* O^\alpha \\ \alpha \in I \\ \text{label}(\mathbf{R}^\alpha) = S \\ \text{label}(P^\alpha) = \varepsilon \end{array}$$

$$\mathcal{D}_{\text{sLC}^i}^{\text{LIpre}} = \frac{[O^\alpha \rightarrow \bullet P^\alpha \nu, 0, 0, false]}{\quad} \begin{array}{l} \top >_\ell^* O^\alpha \\ \alpha \in \mathbf{I} \\ \text{label}(\mathbf{R}^\alpha) = S \\ P^\alpha >_\ell \Delta \text{ o } \exists \alpha' \in \text{subst}(P^\alpha) \end{array}$$

### Reconocimiento

$$\mathcal{D}_{\text{sLC}^i}^{\text{Scan}} = \frac{[a, j, j + 1] \quad [N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \nu, i, j, false]}{[N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, i, j + 1, false]} \quad \text{label}(M^\gamma) = a$$

$$\mathcal{D}_{\text{sLC}^i}^\varepsilon = \frac{[N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \nu, i, j, false]}{[N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, i, j, false]} \quad \text{label}(M^\gamma) = \varepsilon$$

### Filtrado de predicción de subárbol

$$\mathcal{D}_{\text{sLC}^i}^{\text{LCt}} = \frac{[M^\gamma, j] \quad [a, j, j + 1]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, j + 1, false]} \quad \begin{array}{l} M^\gamma >_\ell^* O^\gamma \\ \text{nil} \in \text{ladj}(M^\gamma) \\ \text{label}(P^\gamma) = a \end{array}$$

$$\mathcal{D}_{\text{sLC}^i}^{\text{LC}_\varepsilon} = \frac{[M^\gamma, j]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, j, false]} \quad \begin{array}{l} M^\gamma >_\ell^* O^\gamma \\ \text{nil} \in \text{ladj}(M^\gamma) \\ \text{label}(P^\gamma) = \varepsilon \text{ o } \text{label}(P^\gamma) = \perp \end{array}$$

$$\mathcal{D}_{\text{sLC}^i}^{\text{LCpre}} = \frac{[M^\gamma, j]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j, false]} \quad \begin{array}{l} M^\gamma >_\ell^* O^\gamma \\ \text{nil} \in \text{ladj}(M^\gamma) \\ P^\gamma >_\ell \Delta \text{ o } \exists \alpha \in \text{subst}(P^\gamma) \end{array}$$

### Compleción en la esquina izquierda

$$\mathcal{D}_{\text{sLC}^i}^{\text{LC}_n} = \frac{[O^\gamma \rightarrow \nu \bullet, j, k, \text{radj}]}{[Q^\gamma \rightarrow O^\gamma \bullet \omega, j, k, false]} \quad Q^\gamma >_\ell O^\gamma$$

donde se debe cumplir que si  $\text{radj} = false$  entonces  $\text{nil} \in \text{radj}(O^\gamma)$ .

### Predicción

$$\mathcal{D}_{\text{sLC}^i}^{\text{Pre}} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[M^\gamma, j]} \quad \text{label}(M^\gamma) \in V_N$$

### Compleción de subárbol



$$\mathcal{D}_{\text{sLC}^i}^{\text{Comp}} = \frac{\begin{array}{l} [M^\gamma \rightarrow \nu \bullet, j, k, \text{radj}] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, \text{false}] \end{array}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k, \text{false}]}$$

donde se debe cumplir que si  $\text{radj} = \text{false}$  entonces  $\text{nil} \in \text{radj}(O^\gamma)$ .

#### Filtrado de predicción de adjunción izquierda

$$\mathcal{D}_{\text{sLC}^i}^{\text{LA}_t} = \frac{\begin{array}{l} [M^\gamma, j] \\ [a, j, j+1] \end{array}}{[O^\beta \rightarrow P^\beta \bullet \nu, j, j+1, \text{false}]} \quad \begin{array}{l} \top >_\ell^* O^\beta \\ \beta \in \text{ladj}(M^\gamma) \\ \text{label}(P^\beta) = a \end{array}$$

$$\mathcal{D}_{\text{sLC}^i}^{\text{LA}_\varepsilon} = \frac{[M^\gamma, j]}{[O^\beta \rightarrow P^\beta \bullet \nu, j, j, \text{false}]} \quad \begin{array}{l} \top >_\ell^* O^\beta \\ \beta \in \text{ladj}(M^\gamma) \\ \text{label}(P^\beta) = \varepsilon \end{array}$$

$$\mathcal{D}_{\text{sLC}^i}^{\text{LA}_{\text{pre}}} = \frac{[M^\gamma, j]}{[O^\beta \rightarrow \bullet P^\beta \nu, j, j, \text{false}]} \quad \begin{array}{l} \top >_\ell^* O^\beta \\ \beta \in \text{ladj}(M^\gamma) \\ P^\beta >_\ell \Delta \text{ o } \exists \alpha \in \text{subst}(P^\gamma) \end{array}$$

#### Filtrado de predicción de pie

$$\mathcal{D}_{\text{sLC}^i}^{\text{LF}_t} = \frac{\begin{array}{l} [a, k, k+1] \\ [\top \rightarrow \mathbf{R}^\beta \bullet, j, k, \text{false}] \\ [M^\gamma, j] \end{array}}{[O^\gamma \rightarrow P^\gamma \bullet \nu, k, k+1, \text{false}]} \quad \begin{array}{l} M^\gamma >_\ell^* O^\gamma \\ \beta \in \text{ladj}(M^\gamma) \\ \text{label}(P^\gamma) = a \end{array}$$

$$\mathcal{D}_{\text{sLC}^i}^{\text{LF}_\varepsilon} = \frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\beta \bullet, j, k, \text{false}] \\ [M^\gamma, j] \end{array}}{[O^\gamma \rightarrow P^\gamma \bullet \nu, k, k, \text{false}]} \quad \begin{array}{l} M^\gamma >_\ell^* O^\gamma \\ \beta \in \text{ladj}(M^\gamma) \\ \text{label}(P^\gamma) = \varepsilon \text{ o } \text{label}(P^\gamma) = \perp \end{array}$$

$$\mathcal{D}_{\text{sLC}^i}^{\text{LF}_{\text{pre}}} = \frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\beta \bullet, j, k, \text{false}] \\ [M^\gamma, j] \end{array}}{[O^\gamma \rightarrow \bullet P^\gamma \nu, k, k, \text{false}]} \quad \begin{array}{l} M^\gamma >_\ell^* O^\gamma \\ \beta \in \text{ladj}(M^\gamma) \\ P^\gamma >_\ell \Delta \text{ o } \exists \alpha \in \text{subst}(P^\gamma) \end{array}$$

#### Filtrado de predicción de adjunción derecha

$$\mathcal{D}_{\text{sLC}^i}^{\text{RA}_\varepsilon} = \frac{[M^\gamma \rightarrow \delta \bullet, k, l, \text{false}]}{[O^\beta \rightarrow P^\beta \bullet \nu, l, l, \text{false}]} \quad \begin{array}{l} \top >_\ell^* O^\beta \\ \beta \in \text{radj}(M^\gamma) \\ \text{label}(P^\beta) = \varepsilon \text{ o } \text{label}(P^\beta) = \perp \end{array}$$

### Compleción de adjunción derecha

$$\mathcal{D}_{\text{sLC}^i}^{\text{RAAdjComp}} = \frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\beta \bullet, j, k, \text{false}] \\ [M^\gamma \rightarrow \nu \bullet, i, j, \text{false}] \end{array}}{[M^\gamma \rightarrow \nu \bullet, i, k, \text{true}]} \quad \beta \in \text{radj}(M^\gamma)$$

### Filtrado de predicción de sustitución

$$\mathcal{D}_{\text{sLC}^i}^{\text{LS}_\varepsilon} = \frac{\begin{array}{l} [M^\gamma, j] \\ [a, j, j+1] \end{array}}{[O^\alpha \rightarrow P^\alpha \bullet \nu, j, j+1, \text{false}]} \quad \begin{array}{l} \top >_\varepsilon^* O^\alpha \\ \alpha \in \text{subst}(M^\gamma) \\ \text{label}(P^\alpha) = a \end{array}$$

$$\mathcal{D}_{\text{sLC}^i}^{\text{LS}_\varepsilon} = \frac{[M^\gamma, j]}{[O^\alpha \rightarrow P^\alpha \bullet \nu, j, j, \text{false}]} \quad \begin{array}{l} \top >_\varepsilon^* O^\alpha \\ \alpha \in \text{subst}(M^\gamma) \\ \text{label}(P^\alpha) = \varepsilon \end{array}$$

$$\mathcal{D}_{\text{sLC}^i}^{\text{LS}_{\text{pre}}} = \frac{[M^\gamma, j]}{[O^\alpha \rightarrow \bullet P^\alpha \nu, j, j, \text{false}]} \quad \begin{array}{l} \top >_\varepsilon^* O^\alpha \\ \alpha \in \text{subst}(M^\gamma) \\ P^\alpha >_\varepsilon \Delta \circ \exists \alpha' \in \text{subst}(P^\alpha) \end{array}$$

### Compleción de sustitución

$$\mathcal{D}_{\text{sLC}^i}^{\text{SubsComp}} = \frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\alpha \bullet, j, k, \text{false}] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{false}] \end{array}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k, \text{false}]} \quad \alpha \in \text{subst}(M^\gamma)$$

El conjunto de ítems finales es igual al del esquema **Earley**<sup>i</sup>:

$$\mathcal{F}_{\text{sLC}^i} = \mathcal{F}_{\text{Earley}^i}$$

## 9.5 Esquema tipo Left Corner

En esta sección definiremos el esquema **LC**<sup>i</sup>, que introducimos en [28], y que se deriva del esquema **sLC**<sup>i</sup> mediante la eliminación del dominio de éste del tipo de ítem predictivo.

El conjunto de ítems del esquema **LC**<sup>i</sup> coincide con los ítems Earley del esquema **sLC**<sup>i</sup>:

$$\mathcal{I}_{\text{LC}^i} = \mathcal{I}_{\text{sLC}^i}^{(\text{ii})} \cup \mathcal{I}_{\text{sLC}^i}^{(\text{iii})} \cup \mathcal{I}_{\text{sLC}^i}^{(\text{iv})}$$

Para obtener el conjunto de pasos deductivos del esquema eliminamos el paso  $\mathcal{D}_{\text{sLC}^i}^{\text{Pre}}$ , que es el que genera los ítems predictivos, y adaptamos el resto al nuevo dominio. Para realizar esta adaptación basta con sustituir los ítems predictivos en cada paso donde aparezcan como antecedentes por el antecedente del paso  $\mathcal{D}_{\text{sLC}^i}^{\text{Pre}}$ .

$$\begin{aligned}
\mathcal{D}_{LC^i} &= \mathcal{D}_{LC^i}^{LI_t} \cup \mathcal{D}_{LC^i}^{LI_\varepsilon} \cup \mathcal{D}_{LC^i}^{LI_{pre}} \cup \mathcal{D}_{LC^i}^{Scan} \cup \mathcal{D}_{LC^i}^\varepsilon \cup \mathcal{D}_{LC^i}^{LC_t} \cup \\
&\mathcal{D}_{LC^i}^{LC_\varepsilon} \cup \mathcal{D}_{LC^i}^{LC_{pre}} \cup \mathcal{D}_{LC^i}^{LC_n} \cup \mathcal{D}_{LC^i}^{Comp} \cup \mathcal{D}_{LC^i}^{LA_t} \cup \mathcal{D}_{LC^i}^{LA_\varepsilon} \cup \\
&\mathcal{D}_{LC^i}^{LA_{pre}} \cup \mathcal{D}_{LC^i}^{LF_t} \cup \mathcal{D}_{LC^i}^{LF_\varepsilon} \cup \mathcal{D}_{LC^i}^{LF_{pre}} \cup \mathcal{D}_{LC^i}^{RA_\varepsilon} \cup \\
&\mathcal{D}_{LC^i}^{RA_{adjComp}} \cup \mathcal{D}_{LC^i}^{LS_t} \cup \mathcal{D}_{LC^i}^{LS_\varepsilon} \cup \mathcal{D}_{LC^i}^{LS_{pre}} \cup \mathcal{D}_{LC^i}^{SubsComp}
\end{aligned}$$

Todos los pasos deductivos son iguales a los del esquema  $sLC^i$ , excepto los de filtrado de predicción de subárbol, adjunción izquierda, pie y sustitución:

$$\begin{aligned}
\mathcal{D}_{LC^i}^{LI_t} &= \mathcal{D}_{sLC^i}^{LI_t} & \mathcal{D}_{LC^i}^{LI_\varepsilon} &= \mathcal{D}_{sLC^i}^{LI_\varepsilon} & \mathcal{D}_{LC^i}^{LI_{pre}} &= \mathcal{D}_{sLC^i}^{LI_{pre}} & \mathcal{D}_{LC^i}^{Scan} &= \mathcal{D}_{sLC^i}^{Scan} \\
\mathcal{D}_{LC^i}^\varepsilon &= \mathcal{D}_{sLC^i}^\varepsilon & \mathcal{D}_{LC^i}^{LC_n} &= \mathcal{D}_{sLC^i}^{LC_n} & \mathcal{D}_{LC^i}^{Comp} &= \mathcal{D}_{sLC^i}^{Comp} \\
\mathcal{D}_{LC^i}^{RA_\varepsilon} &= \mathcal{D}_{sLC^i}^{RA_\varepsilon} & \mathcal{D}_{LC^i}^{RA_{adjComp}} &= \mathcal{D}_{sLC^i}^{RA_{adjComp}} & \mathcal{D}_{LC^i}^{SubsComp} &= \mathcal{D}_{sLC^i}^{SubsComp}
\end{aligned}$$

#### Filtrado de predicción de subárbol

$$\mathcal{D}_{LC^i}^{LC_t} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false] \quad [a, j, j+1]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, j+1, false]} \quad \begin{array}{l} M^\gamma >_\ell^* O^\gamma \\ \mathbf{nil} \in \text{ladj}(M^\gamma) \\ \text{label}(P^\gamma) = a \end{array}$$

$$\mathcal{D}_{LC^i}^{LC_\varepsilon} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, j, false]} \quad \begin{array}{l} M^\gamma >_\ell^* O^\gamma \\ \mathbf{nil} \in \text{ladj}(M^\gamma) \\ \text{label}(P^\gamma) = \varepsilon \text{ o } \text{label}(P^\gamma) = \perp \end{array}$$

$$\mathcal{D}_{LC^i}^{LC_{pre}} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j, false]} \quad \begin{array}{l} M^\gamma >_\ell^* O^\gamma \\ \mathbf{nil} \in \text{ladj}(M^\gamma) \\ P^\gamma >_\ell \Delta \text{ o } \exists \alpha \in \text{subst}(P^\gamma) \end{array}$$

#### Filtrado de predicción de adjunción izquierda

$$\mathcal{D}_{LC^i}^{LA_t} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false] \quad [a, j, j+1]}{[O^\beta \rightarrow P^\beta \bullet \nu, j, j+1, false]} \quad \begin{array}{l} \top >_\ell^* O^\beta \\ \beta \in \text{ladj}(M^\gamma) \\ \text{label}(P^\beta) = a \end{array}$$

$$\mathcal{D}_{LC^i}^{LA_\varepsilon} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false]}{[O^\beta \rightarrow P^\beta \bullet \nu, j, j, false]} \quad \begin{array}{l} \top >_\ell^* O^\beta \\ \beta \in \text{ladj}(M^\gamma) \\ \text{label}(P^\beta) = \varepsilon \end{array}$$

$$\mathcal{D}_{LC^i}^{LA_{pre}} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false]}{[O^\beta \rightarrow \bullet P^\beta \nu, j, j, false]} \quad \begin{array}{l} \top >_\ell^* O^\beta \\ \beta \in \text{ladj}(M^\gamma) \\ P^\beta >_\ell \Delta \text{ o } \exists \alpha \in \text{subst}(P^\gamma) \end{array}$$

**Filtrado de predicción de pie**

$$\mathcal{D}_{LC^i}^{LF_t} = \frac{\begin{array}{l} [a, k, k + 1] \\ [\top \rightarrow \mathbf{R}^{\beta \bullet}, j, k, false] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false] \end{array}}{[O^\gamma \rightarrow P^\gamma \bullet \nu, k, k + 1, false]} \quad \begin{array}{l} M^\gamma >_{\ell}^* O^\gamma \\ \beta \in \text{ladj}(M^\gamma) \\ \text{label}(P^\gamma) = a \end{array}$$

$$\mathcal{D}_{LC^i}^{LF_\varepsilon} = \frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^{\beta \bullet}, j, k, false] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false] \end{array}}{[O^\gamma \rightarrow P^\gamma \bullet \nu, k, k, false]} \quad \begin{array}{l} M^\gamma >_{\ell}^* O^\gamma \\ \beta \in \text{ladj}(M^\gamma) \\ \text{label}(P^\gamma) = \varepsilon \text{ o } \text{label}(P^\gamma) = \perp \end{array}$$

$$\mathcal{D}_{LC^i}^{LF_{\text{pro}}} = \frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^{\beta \bullet}, j, k, false] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false] \end{array}}{[O^\gamma \rightarrow \bullet P^\gamma \nu, k, k, false]} \quad \begin{array}{l} M^\gamma >_{\ell}^* O^\gamma \\ \beta \in \text{ladj}(M^\gamma) \\ P^\gamma >_{\ell} \Delta \text{ o } \exists \alpha \in \text{subst}(P^\gamma) \end{array}$$

**Filtrado de predicción de sustitución**

$$\mathcal{D}_{LC^i}^{LS_t} = \frac{\begin{array}{l} [N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false] \\ [a, j, j + 1] \end{array}}{[O^\alpha \rightarrow P^\alpha \bullet \nu, j, j + 1, false]} \quad \begin{array}{l} \top >_{\ell}^* O^\alpha \\ \alpha \in \text{subst}(M^\gamma) \\ \text{label}(P^\alpha) = a \end{array}$$

$$\mathcal{D}_{LC^i}^{LS_\varepsilon} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false]}{[O^\alpha \rightarrow P^\alpha \bullet \nu, j, j, false]} \quad \begin{array}{l} \top >_{\ell}^* O^\alpha \\ \alpha \in \text{subst}(M^\gamma) \\ \text{label}(P^\alpha) = \varepsilon \end{array}$$

$$\mathcal{D}_{LC^i}^{LS_{\text{pre}}} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false]}{[O^\alpha \rightarrow \bullet P^\alpha \nu, j, j, false]} \quad \begin{array}{l} \top >_{\ell}^* O^\alpha \\ \alpha \in \text{subst}(M^\gamma) \\ P^\alpha >_{\ell} \Delta \text{ o } \exists \alpha' \in \text{subst}(P^\alpha) \end{array}$$

El conjunto de ítems finales es igual al del esquema  $\mathbf{sLC}^i$ :

$$\mathcal{F}_{LC^i} = \mathcal{F}_{\mathbf{sLC}^i}$$

**9.6 Ampliando la red de analizadores para TIG**

En esta sección vamos a establecer las relaciones formales entre los esquemas descritos en este capítulo para, a continuación, integrar los mismos en la red de analizadores que presentamos en el capítulo anterior.

### 9.6.1 Relaciones entre esquemas tipo Left Corner para TIG

**Teorema 9.1** *Relación entre los esquemas  $\mathbf{pLC}^i$ ,  $\mathbf{sLC}^i$  y  $\mathbf{LC}^i$*

*Se mantienen las siguientes relaciones de refinamiento de pasos e ítems:*

$$\mathbf{LC}^i \xrightarrow{\text{sr}} \mathbf{sLC}^i \xrightarrow{\text{ir}} \mathbf{pLC}^i$$

*Prueba*

Primero vamos a probar la relación  $\mathbf{sLC}^i \xrightarrow{\text{ir}} \mathbf{pLC}^i$ . Debemos demostrar que existe una función regular  $f : \mathcal{I}_{\mathbf{pLC}^i} \rightarrow \mathcal{I}_{\mathbf{sLC}^i}$  tal que:

1.  $\mathcal{I}_{\mathbf{sLC}^i} = f(\mathcal{I}_{\mathbf{pLC}^i})$
2.  $\Delta_{\mathbf{sLC}^i} = f(\Delta_{\mathbf{pLC}^i})$

Dada la función regular:

$$f([C^\gamma; N^\gamma \rightarrow \delta \bullet \nu, i, j \mid p, q]) = [N^\gamma \rightarrow \delta \bullet \nu, i, j \mid p, q]$$

$$f([C^\gamma, j]) = [C^\gamma, j]$$

se sigue inmediatamente que  $\mathcal{I}_{\mathbf{sLC}^i} = f(\mathcal{I}_{\mathbf{pLC}^i})$  y  $\Delta_{\mathbf{sLC}^i} = f(\Delta_{\mathbf{pLC}^i})$  por inducción en la longitud de las secuencias de derivación.

Ahora probaremos la relación  $\mathbf{LC}^i \xrightarrow{\text{sr}} \mathbf{sLC}^i$ . Para ello tenemos que demostrar que:

1.  $\mathcal{I}_{\mathbf{LC}^i} \subseteq \mathcal{I}_{\mathbf{sLC}^i}$
2.  $\vdash_{\mathbf{LC}^i}^* \subseteq \vdash_{\mathbf{sLC}^i}^*$

Como el dominio del esquema  $\mathbf{LC}^i$  está formado por los conjuntos de ítems de tipo Earley del esquema  $\mathbf{sLC}^i$ , se cumple que:

$$\mathcal{I}_{\mathbf{LC}^i} \subset \mathcal{I}_{\mathbf{sLC}^i}$$

Para demostrar que  $\vdash_{\mathbf{LC}^i}^* \subseteq \vdash_{\mathbf{sLC}^i}^*$  hay que probar que  $\mathcal{D}_{\mathbf{LC}^i} \subseteq \mathcal{D}_{\mathbf{sLC}^i}^*$ . Veamos cada paso:

- Un paso deductivo  $\mathcal{D}_{\mathbf{LC}^i}^{\text{LC}^i}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\mathbf{sLC}^i}^{\text{Pre}}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{false}]}{[M^\gamma, j]}$$

seguido de otro  $\mathcal{D}_{\mathbf{sLC}^i}^{\text{LC}^i}$

$$\frac{\frac{[M^\gamma, j]}{[a, j, j+1]}}{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, j+1, \text{false}]}$$

- Un paso deductivo  $\mathcal{D}_{\mathbf{LC}^i}^{\text{LC}^e}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\mathbf{sLC}^i}^{\text{Pre}}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{false}]}{[M^\gamma, j]}$$

seguido de otro  $\mathcal{D}_{\text{sLC}^i}^{\text{LC}_\epsilon}$

$$\frac{[M^\gamma, j]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, j, \text{false}]}$$

- Un paso deductivo  $\mathcal{D}_{\text{LC}^i}^{\text{LC}_{\text{pre}}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{sLC}^i}^{\text{Pre}}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{false}]}{[M^\gamma, j]}$$

seguido de otro  $\mathcal{D}_{\text{sLC}^i}^{\text{LC}_{\text{pre}}}$

$$\frac{[M^\gamma, j]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j, \text{false}]}$$

- Un paso deductivo  $\mathcal{D}_{\text{LC}^i}^{\text{LA}_t}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{sLC}^i}^{\text{Pre}}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{false}]}{[M^\gamma, j]}$$

seguido de otro  $\mathcal{D}_{\text{sLC}^i}^{\text{LA}_t}$

$$\frac{[M^\gamma, j] \\ [a, j, j + 1]}{[O^\beta \rightarrow P^\beta \bullet \nu, j, j + 1, \text{false}]}$$

- Un paso deductivo  $\mathcal{D}_{\text{LC}^i}^{\text{LA}_\epsilon}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{sLC}^i}^{\text{Pre}}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{false}]}{[M^\gamma, j]}$$

seguido de otro  $\mathcal{D}_{\text{sLC}^i}^{\text{LA}_\epsilon}$

$$\frac{[M^\gamma, j]}{[O^\beta \rightarrow P^\beta \bullet \nu, j, j, \text{false}]}$$

- Un paso deductivo  $\mathcal{D}_{\text{LC}^i}^{\text{LA}_{\text{pre}}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{sLC}^i}^{\text{Pre}}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{false}]}{[M^\gamma, j]}$$

seguido de otro  $\mathcal{D}_{\text{sLC}^i}^{\text{LA}_{\text{pre}}}$

$$\frac{[M^\gamma, j]}{[O^\beta \rightarrow \bullet P^\beta \nu, j, j, \text{false}]}$$

- Un paso deductivo  $\mathcal{D}_{\text{LC}^i}^{\text{LF}_t}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{sLC}^i}^{\text{Pre}}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{false}]}{[M^\gamma, j]}$$

seguido de otro  $\mathcal{D}_{\text{sLC}^i}^{\text{LF}_t}$

$$\frac{[a, k, k + 1] \\ [\top \rightarrow \mathbf{R}^\beta \bullet, j, k, \text{false}] \\ [M^\gamma, j]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, k, k + 1, \text{false}]}$$



- Un paso deductivo  $\mathcal{D}_{\text{sLC}^i}^{\text{LF}_\epsilon}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{sLC}^i}^{\text{Pre}}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{false}]}{[M^\gamma, j]}$$

seguido de otro  $\mathcal{D}_{\text{sLC}^i}^{\text{LF}_\epsilon}$

$$\frac{\frac{[\top \rightarrow \mathbf{R}^\beta \bullet, j, k, \text{false}]}{[M^\gamma, j]}}{[O^\gamma \rightarrow P^\gamma \bullet \nu, k, k, \text{false}]}$$

- Un paso deductivo  $\mathcal{D}_{\text{LC}^i}^{\text{LF}_{\text{pre}}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{sLC}^i}^{\text{Pre}}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{false}]}{[M^\gamma, j]}$$

seguido de otro  $\mathcal{D}_{\text{sLC}^i}^{\text{LF}_{\text{pre}}}$

$$\frac{\frac{[\top \rightarrow \mathbf{R}^\beta \bullet, j, k, \text{false}]}{[M^\gamma, j]}}{[O^\gamma \rightarrow \bullet P^\gamma \nu, k, k, \text{false}]}$$

- Un paso deductivo  $\mathcal{D}_{\text{LC}^i}^{\text{LS}_t}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{sLC}^i}^{\text{Pre}}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{false}]}{[M^\gamma, j]}$$

seguido de otro  $\mathcal{D}_{\text{sLC}^i}^{\text{LS}_t}$

$$\frac{\frac{[M^\gamma, j]}{[a, j, j + 1]}}{[O^\alpha \rightarrow P^\alpha \bullet \nu, j, j + 1, \text{false}]}$$

- Un paso deductivo  $\mathcal{D}_{\text{LC}^i}^{\text{LS}_\epsilon}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{sLC}^i}^{\text{Pre}}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{false}]}{[M^\gamma, j]}$$

seguido de otro  $\mathcal{D}_{\text{sLC}^i}^{\text{LS}_\epsilon}$

$$\frac{[M^\gamma, j]}{[O^\alpha \rightarrow P^\alpha \bullet \nu, j, j, \text{false}]}$$

- Un paso deductivo  $\mathcal{D}_{\text{LC}^i}^{\text{LS}_{\text{pre}}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{sLC}^i}^{\text{Pre}}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{false}]}{[M^\gamma, j]}$$

seguido de otro  $\mathcal{D}_{\text{sLC}^i}^{\text{LS}_{\text{pre}}}$

$$\frac{[M^\gamma, j]}{[O^\alpha \rightarrow \bullet P^\alpha \nu, j, j, \text{false}]}$$

- El resto de pasos homónimos en ambos esquemas son iguales.

**Teorema 9.2** *Relación entre los esquemas  $\text{buLC}^i$  y  $\text{LC}^i$* 

Se mantiene la siguientes relaciones de refinamiento de pasos y filtros dinámicos:

$$\text{buLC}^i \xrightarrow{\text{sr}} \text{buLC}_1^i \xrightarrow{\text{df}} \text{LC}_1^i \xleftarrow{\text{df}} \text{LC}^i$$

*Prueba*

El esquema  $\text{buLC}_1^i$  se obtiene mediante la ampliación del dominio de  $\text{buLC}^i$  con ítems con el punto al comienzo de la regla cuando el hijo izquierdo sea un nodo adjuntable o marcado para sustitución. Por tanto, el conjunto de ítems para este esquema es igual al del esquema  $\text{LC}^i$ :

$$\mathcal{I}_{\text{buLC}_1^i} = \mathcal{I}_{\text{LC}^i}$$

Hay que adaptar el conjunto de pasos deductivos al nuevo dominio:

$$\begin{aligned} \mathcal{D}_{\text{buLC}_1^i} = & \mathcal{D}_{\text{buLC}_1^i}^{\text{LC}_t} \cup \mathcal{D}_{\text{buLC}_1^i}^{\text{LC}_\epsilon} \cup \mathcal{D}_{\text{buLC}_1^i}^{\text{LC}_{\text{pre}}} \cup \mathcal{D}_{\text{buLC}_1^i}^{\text{LC}_n} \cup \mathcal{D}_{\text{buLC}_1^i}^{\text{LAdj}_t} \cup \mathcal{D}_{\text{buLC}_1^i}^{\text{LAdj}_\epsilon} \cup \mathcal{D}_{\text{buLC}_1^i}^{\text{LAdj}_{\text{pre}}} \cup \\ & \mathcal{D}_{\text{buLC}_1^i}^{\text{LAdj}_n} \cup \mathcal{D}_{\text{buLC}_1^i}^\epsilon \cup \mathcal{D}_{\text{buLC}_1^i}^{\text{Scan}} \cup \mathcal{D}_{\text{buLC}_1^i}^{\text{Comp}} \cup \mathcal{D}_{\text{buLC}_1^i}^{\text{RADj}} \cup \mathcal{D}_{\text{buLC}_1^i}^{\text{Subs}} \end{aligned}$$

donde todos los pasos son idénticos al los homónimos de  $\text{buLC}^i$ , excepto el paso  $\mathcal{D}_{\text{buLC}_1^i}^{\text{Subs}}$  y los nuevos pasos  $\mathcal{D}_{\text{buLC}_1^i}^{\text{LC}_{\text{pre}}}$  y  $\mathcal{D}_{\text{buLC}_1^i}^{\text{LAdj}_{\text{pre}}}$ .

$$\begin{aligned} \mathcal{D}_{\text{buLC}_1^i}^{\text{LC}_t} &= \mathcal{D}_{\text{buLC}^i}^{\text{LC}_t} & \mathcal{D}_{\text{buLC}_1^i}^{\text{LC}_\epsilon} &= \mathcal{D}_{\text{buLC}^i}^{\text{LC}_\epsilon} & \mathcal{D}_{\text{buLC}_1^i}^{\text{LC}_n} &= \mathcal{D}_{\text{buLC}^i}^{\text{LC}_n} & \mathcal{D}_{\text{buLC}_1^i}^{\text{LAdj}_t} &= \mathcal{D}_{\text{buLC}^i}^{\text{LAdj}_t} \\ \mathcal{D}_{\text{buLC}_1^i}^{\text{LAdj}_\epsilon} &= \mathcal{D}_{\text{buLC}^i}^{\text{LAdj}_\epsilon} & \mathcal{D}_{\text{buLC}_1^i}^{\text{LAdj}_n} &= \mathcal{D}_{\text{buLC}^i}^{\text{LAdj}_n} & \mathcal{D}_{\text{buLC}_1^i}^{\text{Scan}} &= \mathcal{D}_{\text{buLC}^i}^{\text{Scan}} \\ \mathcal{D}_{\text{buLC}_1^i}^\epsilon &= \mathcal{D}_{\text{buLC}^i}^\epsilon & \mathcal{D}_{\text{buLC}_1^i}^{\text{Comp}} &= \mathcal{D}_{\text{buLC}^i}^{\text{Comp}} & \mathcal{D}_{\text{buLC}_1^i}^{\text{RADj}} &= \mathcal{D}_{\text{buLC}^i}^{\text{RADj}} \end{aligned}$$

$$\mathcal{D}_{\text{buLC}_1^i}^{\text{LC}_{\text{pre}}} = \frac{[\text{O}^\gamma \rightarrow \bullet P^\gamma \nu, j, j, \text{false}]}{[\text{O}^\gamma \rightarrow \bullet P^\gamma \nu, j, j, \text{false}]} \quad \begin{array}{l} \text{nil} \in \text{ladj}(\text{O}^\gamma) \\ P^\gamma >_\ell \Delta \text{ o } \exists \alpha \in \text{subst}(P^\gamma) \end{array}$$

$$\mathcal{D}_{\text{buLC}_1^i}^{\text{LAdj}_{\text{pre}}} = \frac{[\text{T} \rightarrow \mathbf{R}^\beta \bullet, i, j, \text{false}]}{[\text{O}^\gamma \rightarrow \bullet P^\gamma \nu, j, j, \text{false}]} \quad \begin{array}{l} \beta \in \text{ladj}(\text{O}^\gamma) \\ P^\gamma >_\ell \Delta \text{ o } \exists \alpha \in \text{subst}(P^\gamma) \end{array}$$

$$\mathcal{D}_{\text{buLC}_1^i}^{\text{Subs}} = \frac{\begin{array}{l} [\text{T} \rightarrow \mathbf{R}^\alpha \bullet, j, k, \text{false}] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{false}] \end{array}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k, \text{false}]} \quad \alpha \in \text{subst}(M^\gamma)$$

El conjunto de ítems finales es igual al del esquema  $\text{buLC}^i$ :

$$\mathcal{F}_{\text{buLC}_1^i} = \mathcal{F}_{\text{buLC}^i}$$

Probemos ahora que  $\text{buLC}^i \xrightarrow{\text{sr}} \text{buLC}_1^i$ , para lo cual se tiene que cumplir que:

1.  $\mathcal{I}_{\text{buLC}^i} \subseteq \mathcal{I}_{\text{buLC}_1^i}$

$$2. \vdash_{\text{buLC}^i}^* \subseteq \vdash_{\text{buLC}_1^i}^*$$

Lo primero es cierto por definición, ya que el dominio del esquema  $\text{buLC}_1^i$  está formado por los ítems del esquema  $\text{buLC}^i$  ampliado con ítems con el punto a comienzo de la regla, luego se cumple:

$$\mathcal{I}_{\text{buLC}^i} \subset \mathcal{I}_{\text{buLC}_1^i}$$

Para demostrar que  $\vdash_{\text{buLC}^i}^* \subseteq \vdash_{\text{buLC}_1^i}^*$  nos basta con probar que  $\mathcal{D}_{\text{buLC}^i} \subseteq \mathcal{D}_{\text{buLC}_1^i}^*$ . Veamos únicamente los pasos en que difieren:

- Un paso deductivo  $\mathcal{D}_{\text{buLC}^i}^{\text{LC}_{\text{subs}}}$  es equivalente a una secuencia formada por una paso  $\mathcal{D}_{\text{buLC}_1^i}^{\text{LC}_{\text{pre}}}$

$$\frac{}{[O\gamma \rightarrow \bullet P\gamma\nu, j, j, \text{false}]}$$

seguido de otro  $\mathcal{D}_{\text{buLC}_1^i}^{\text{Subs}}$

$$\frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\alpha \bullet, j, k, \text{false}] \\ [O\gamma \rightarrow \bullet P\gamma\nu, j, j, \text{false}] \end{array}}{[O\gamma \rightarrow P\gamma \bullet \nu, j, k, \text{false}]}$$

- Un paso deductivo  $\mathcal{D}_{\text{buLC}^i}^{\text{LAdj}_{\text{subs}}}$  es equivalente a una secuencia formada por una paso  $\mathcal{D}_{\text{buLC}_1^i}^{\text{LAdj}_{\text{pre}}}$

$$\frac{[\top \rightarrow \mathbf{R}^\beta \bullet, i, j, \text{false}]}{[O\gamma \rightarrow \bullet P\gamma\nu, j, j, \text{false}]}$$

seguido de otro  $\mathcal{D}_{\text{buLC}_1^i}^{\text{Subs}}$

$$\frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\alpha \bullet, j, k, \text{false}] \\ [O\gamma \rightarrow \bullet P\gamma\nu, j, j, \text{false}] \end{array}}{[O\gamma \rightarrow P\gamma \bullet \nu, j, k, \text{false}]}$$

- Un paso deductivo  $\mathcal{D}_{\text{buLC}^i}^{\text{Subs}}$  está incluido en la inferencia del paso  $\mathcal{D}_{\text{buLC}_1^i}^{\text{Subs}}$ , ya que este último contempla la operación de sustitución en cualquier posición, mientras el primero sólo es válido para sustituciones en nodos que no sean hijos izquierdos. Por tanto:

$$\mathcal{D}_{\text{buLC}^i}^{\text{Subs}} \subset \mathcal{D}_{\text{buLC}_1^i}^{\text{Subs}}$$

El esquema  $\text{LC}_1^i$  se obtiene desdoblando el paso  $\mathcal{D}_{\text{LC}^i}^{\text{Comp}}$  del esquema  $\text{LC}^i$  en los pasos  $\mathcal{D}_{\text{LC}_1^i}^{\text{Comp}_1}$  y  $\mathcal{D}_{\text{LC}_1^i}^{\text{Comp}_2}$ , donde el primero se encarga de completar los subárboles en nodos que no son hijos izquierdos, mientras el segundo los completa en los hijos izquierdos.

El dominio de  $\text{LC}_1^i$  es igual al del esquema original:

$$\mathcal{I}_{\text{LC}_1^i} = \mathcal{I}_{\text{LC}^i}$$

El conjunto de pasos deductivos del nuevo esquema es:

$$\mathcal{D}_{\text{LC}_1^i} = \mathcal{D}_{\text{LC}_1^i}^{\text{LI}_t} \cup \mathcal{D}_{\text{LC}_1^i}^{\text{LI}_\epsilon} \cup \mathcal{D}_{\text{LC}_1^i}^{\text{LI}_{\text{pre}}} \cup \mathcal{D}_{\text{LC}_1^i}^{\text{Scan}} \cup \mathcal{D}_{\text{LC}_1^i}^\epsilon \cup \mathcal{D}_{\text{LC}_1^i}^{\text{LC}_t} \cup$$

$$\begin{aligned}
& \mathcal{D}_{LC_1^i}^{LC_\epsilon} \cup \mathcal{D}_{LC_1^i}^{LC_{pre}} \cup \mathcal{D}_{LC_1^i}^{LC_n} \cup \mathcal{D}_{LC_1^i}^{Comp_1} \cup \mathcal{D}_{LC_1^i}^{Comp_2} \cup \mathcal{D}_{LC_1^i}^{LA_t} \cup \\
& \mathcal{D}_{LC_1^i}^{LA_\epsilon} \cup \mathcal{D}_{LC_1^i}^{LA_{pre}} \cup \mathcal{D}_{LC_1^i}^{LF_t} \cup \mathcal{D}_{LC_1^i}^{LF_\epsilon} \cup \mathcal{D}_{LC_1^i}^{LF_{pre}} \cup \mathcal{D}_{LC_1^i}^{RA_\epsilon} \cup \\
& \mathcal{D}_{LC_1^i}^{RA_{adj}Comp} \cup \mathcal{D}_{LC_1^i}^{LS_t} \cup \mathcal{D}_{LC_1^i}^{LS_\epsilon} \cup \mathcal{D}_{LC_1^i}^{LS_{pre}} \cup \mathcal{D}_{LC_1^i}^{SubsComp}
\end{aligned}$$

donde todos los pasos deductivos son iguales a sus homónimos del esquema  $LC^i$ , excepto los dos nuevos obtenidos mediante el desdoble.

$$\begin{aligned}
\mathcal{D}_{LC_1^i}^{LI_t} &= \mathcal{D}_{LC^i}^{LI_t} & \mathcal{D}_{LC_1^i}^{LI_\epsilon} &= \mathcal{D}_{LC^i}^{LI_\epsilon} & \mathcal{D}_{LC_1^i}^{LI_{pre}} &= \mathcal{D}_{LC^i}^{LI_{pre}} \\
\mathcal{D}_{LC_1^i}^{Scan} &= \mathcal{D}_{LC^i}^{Scan} & \mathcal{D}_{LC_1^i}^\epsilon &= \mathcal{D}_{LC^i}^\epsilon & \mathcal{D}_{LC_1^i}^{LC_t} &= \mathcal{D}_{LC^i}^{LC_t} \\
\mathcal{D}_{LC_1^i}^{LC_\epsilon} &= \mathcal{D}_{LC^i}^{LC_\epsilon} & \mathcal{D}_{LC_1^i}^{LC_{pre}} &= \mathcal{D}_{LC^i}^{LC_{pre}} & \mathcal{D}_{LC_1^i}^{LC_n} &= \mathcal{D}_{LC^i}^{LC_n} \\
\mathcal{D}_{LC_1^i}^{LA_t} &= \mathcal{D}_{LC^i}^{LA_t} & \mathcal{D}_{LC_1^i}^{LA_\epsilon} &= \mathcal{D}_{LC^i}^{LA_\epsilon} & \mathcal{D}_{LC_1^i}^{LA_{pre}} &= \mathcal{D}_{LC^i}^{LA_{pre}} \\
\mathcal{D}_{LC_1^i}^{LF_t} &= \mathcal{D}_{LC^i}^{LF_t} & \mathcal{D}_{LC_1^i}^{LF_\epsilon} &= \mathcal{D}_{LC^i}^{LF_\epsilon} & \mathcal{D}_{LC_1^i}^{LF_{pre}} &= \mathcal{D}_{LC^i}^{LF_{pre}} \\
\mathcal{D}_{LC_1^i}^{RA_\epsilon} &= \mathcal{D}_{LC^i}^{RA_\epsilon} & \mathcal{D}_{LC_1^i}^{RA_{adj}Comp} &= \mathcal{D}_{LC^i}^{RA_{adj}Comp} & \mathcal{D}_{LC_1^i}^{LS_t} &= \mathcal{D}_{LC^i}^{LS_t} \\
\mathcal{D}_{LC_1^i}^{LS_\epsilon} &= \mathcal{D}_{LC^i}^{LS_\epsilon} & \mathcal{D}_{LC_1^i}^{LS_{pre}} &= \mathcal{D}_{LC^i}^{LS_{pre}} & \mathcal{D}_{LC_1^i}^{SubsComp} &= \mathcal{D}_{LC^i}^{SubsComp} \\
\mathcal{D}_{LC_1^i}^{Comp_1} &= \frac{[M^\gamma \rightarrow \nu \bullet, j, k, radj] [N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \omega, i, j, false]}{[N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, i, k, false]}
\end{aligned}$$

donde se debe cumplir que si  $radj = false$  entonces  $\mathbf{nil} \in \text{radj}(O^\gamma)$ .

$$\mathcal{D}_{LC_1^i}^{Comp_2} = \frac{[M^\gamma \rightarrow \nu \bullet, j, k, radj] [N^\gamma \rightarrow \bullet M^\gamma \omega, j, j, false]}{[N^\gamma \rightarrow M^\gamma \bullet \nu, j, k, false]}$$

donde se debe cumplir que si  $radj = false$  entonces  $\mathbf{nil} \in \text{radj}(O^\gamma)$ .

El conjunto de ítems finales también es igual al del esquema  $LC^i$ :

$$\mathcal{I}_{LC_1^i} = \mathcal{I}_{LC^i}$$

Para probar que  $LC^i \xrightarrow{df} LC_1^i$  tenemos que demostrar que:

1.  $\mathcal{I}_{LC_1^i} \subseteq \mathcal{I}_{LC^i}$
2.  $\vdash_{LC_1^i} \subseteq \vdash_{LC^i}$

Lo primero es cierto por definición, ya que los dominios de ambos esquemas son iguales:

$$\mathcal{I}_{LC_1^i} = \mathcal{I}_{LC^i}$$

Para probar que  $\vdash_{LC_1^i} \subseteq \vdash_{LC^i}$  debemos demostrar que  $\mathcal{D}_{LC_1^i} \subseteq \vdash_{LC^i}$ . Sólo vamos a considerar los dos pasos sobre los que se aplica el filtro dinámico, el resto son idénticos:

- Dado un paso

$$\frac{\begin{array}{l} [M^\gamma \rightarrow \nu \bullet, j, k, \text{radj}] \\ [N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \omega, i, j, \text{false}] \end{array}}{[N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, i, k, \text{false}]} \in \mathcal{D}_{\text{LC}_1^i}^{\text{Comp}_1}$$

existe un paso

$$\frac{\begin{array}{l} [M^\gamma \rightarrow \nu \bullet, j, k, \text{radj}] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, \text{false}] \end{array}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k, \text{false}]} \in \mathcal{D}_{\text{LC}_1^i}^{\text{Comp}}$$

y, por tanto, existe la inferencia

$$\frac{\begin{array}{l} [M^\gamma \rightarrow \nu \bullet, j, k, \text{radj}] \\ [N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \omega, i, j, \text{false}] \end{array}}{[N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, i, k, \text{false}]} \vdash_{\text{LC}_1^i}$$

- Dado un paso

$$\frac{\begin{array}{l} [M^\gamma \rightarrow \nu \bullet, j, k, \text{radj}] \\ [N^\gamma \rightarrow \bullet M^\gamma \omega, j, j, \text{false}] \end{array}}{[N^\gamma \rightarrow M^\gamma \bullet \nu, j, k, \text{false}]} \in \mathcal{D}_{\text{LC}_1^i}^{\text{Comp}_2}$$

existe un paso

$$\frac{\begin{array}{l} [M^\gamma \rightarrow \nu \bullet, j, k, \text{radj}] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, \text{false}] \end{array}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k, \text{false}]} \in \mathcal{D}_{\text{LC}_1^i}^{\text{Comp}}$$

y, por tanto, existe la inferencia

$$\frac{\begin{array}{l} [M^\gamma \rightarrow \nu \bullet, j, k, \text{radj}] \\ [N^\gamma \rightarrow \bullet M^\gamma \omega, j, j, \text{false}] \end{array}}{[N^\gamma \rightarrow M^\gamma \bullet \nu, j, k, \text{false}]} \vdash_{\text{LC}_1^i}$$

Nos queda probar que  $\text{buLC}_1^i \stackrel{\text{df}}{\iff} \text{LC}_1^i$ , para lo cual tenemos que demostrar que:

1.  $\mathcal{I}_{\text{LC}_1^i} \subseteq \mathcal{I}_{\text{buLC}_1^i}$
2.  $\vdash_{\text{LC}_1^i} \subseteq \vdash_{\text{buLC}_1^i}$

Lo primero es cierto por definición, ya que los dominios de ambos esquemas son iguales:

$$\mathcal{I}_{\text{LC}_1^i} = \mathcal{I}_{\text{LC}_1^i} = \mathcal{I}_{\text{buLC}_1^i}$$

Para probar que  $\vdash_{\text{LC}_1^i} \subseteq \vdash_{\text{buLC}_1^i}$  debemos demostrar que  $\mathcal{D}_{\text{LC}_1^i} \subseteq \mathcal{D}_{\text{buLC}_1^i}$ . Vamos a ver cada paso:

- Dado un paso

$$\frac{[a, 0, 1]}{[O^\alpha \rightarrow P^\alpha \bullet \nu, 0, 1, \text{false}]} \in \mathcal{D}_{\text{LC}_1^i}^{\text{Ll}_t}$$

existe un paso

$$\frac{[a, j, j+1]}{[O^\alpha \rightarrow P^\alpha \bullet \nu, j, j+1, \text{false}]} \in \mathcal{D}_{\text{buLC}_1^i}^{\text{LC}_t}$$

y, por tanto, existe la inferencia

$$[a, 0, 1] \vdash_{\text{buLC}_1^i} [O^\alpha \rightarrow P^\alpha \bullet \nu, 0, 1, \text{false}]$$

- Dado un paso

$$\frac{}{[O^\alpha \rightarrow P^\alpha \bullet \nu, 0, 0, false]} \in \mathcal{D}_{LC_1}^{LI_\epsilon}$$

existe un paso

$$\frac{}{[O^\alpha \rightarrow P^\alpha \bullet \nu, j, j, false]} \in \mathcal{D}_{buLC_1}^{LC_\epsilon}$$

y, por tanto, existe la inferencia

$$\vdash_{buLC_1} [O^\alpha \rightarrow P^\alpha \bullet \nu, 0, 0, false]$$

- Dado un paso

$$\frac{}{[O^\alpha \rightarrow \bullet P^\alpha \nu, 0, 0, false]} \in \mathcal{D}_{LC_1}^{LI_{pre}}$$

existe un paso

$$\frac{}{[O^\alpha \rightarrow \bullet P^\alpha \nu, j, j, false]} \in \mathcal{D}_{buLC_1}^{LC_{pre}}$$

y, por tanto, existe la inferencia

$$\vdash_{buLC_1} [O^\alpha \rightarrow \bullet P^\alpha \nu, 0, 0, false]$$

- Dado un paso

$$\frac{\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[a, j, j+1]}}{[O^\gamma \rightarrow P^\gamma \bullet \omega, j, j+1, false]} \in \mathcal{D}_{LC_1}^{LC_t}$$

existe un paso

$$\frac{[a, j, j+1]}{[O^\gamma \rightarrow P^\gamma \bullet \omega, j, j+1, false]} \in \mathcal{D}_{buLC_1}^{LC_t}$$

y, por tanto, existe la inferencia

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[a, j, j+1]} \vdash_{buLC_1} [O^\gamma \rightarrow P^\gamma \bullet \omega, j, j+1, false]$$

- Dado un paso

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[O^\gamma \rightarrow P^\gamma \bullet \omega, j, j, false]} \in \mathcal{D}_{LC_1}^{LC_\epsilon}$$

existe un paso

$$\frac{}{[O^\gamma \rightarrow P^\gamma \bullet \omega, j, j, false]} \in \mathcal{D}_{buLC_1}^{LC_\epsilon}$$

y, por tanto, existe la inferencia

$$[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false] \vdash_{buLC_1} [O^\gamma \rightarrow P^\gamma \bullet \omega, j, j, false]$$



- Dado un paso

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[O^\gamma \rightarrow \bullet P^\gamma \omega, j, j, false]} \in \mathcal{D}_{LC_1^i}^{LC_{pre}}$$

existe un paso

$$\overline{[O^\gamma \rightarrow \bullet P^\gamma \omega, j, j, false]} \in \mathcal{D}_{buLC_1^i}^{LC_{pre}}$$

y por tanto, existe la inferencia

$$[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false] \vdash_{buLC_1^i} [O^\gamma \rightarrow \bullet P^\gamma \omega, j, j, false]$$

- Dado un paso

$$\frac{\begin{array}{l} [O^\gamma \rightarrow \nu \bullet, j, k, radj] \\ [Q^\gamma \rightarrow \bullet O^\gamma \omega, j, j, false] \end{array}}{[Q^\gamma \rightarrow O^\gamma \bullet \omega, j, k, false]} \in \mathcal{D}_{LC_1^i}^{Comp_2}$$

existe un paso

$$\frac{[O^\gamma \rightarrow \nu \bullet, j, k, radj]}{[Q^\gamma \rightarrow O^\gamma \bullet \omega, j, k, false]} \in \mathcal{D}_{buLC_1^i}^{LC_n}$$

y, por tanto, existe la inferencia

$$\frac{[O^\gamma \rightarrow \nu \bullet, j, k, radj]}{[Q^\gamma \rightarrow \bullet O^\gamma \omega, j, j, false]} \vdash_{buLC_1^i} [Q^\gamma \rightarrow O^\gamma \bullet \omega, j, k, false]$$

- Dado un paso

$$\frac{\begin{array}{l} [N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false] \\ [a, j, j+1] \end{array}}{[O^\beta \rightarrow P^\beta \bullet \nu, j, j+1, false]} \in \mathcal{D}_{LC_1^i}^{LA_t}$$

existe un paso

$$\frac{[a, j, j+1]}{[O^\beta \rightarrow P^\beta \bullet \nu, j, j+1, false]} \in \mathcal{D}_{buLC_1^i}^{LC_t}$$

y, por tanto, existe la inferencia

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false]}{[a, j, j+1]} \vdash_{buLC_1^i} [O^\beta \rightarrow P^\beta \bullet \nu, j, j+1, false]$$

- Dado un paso

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false]}{[O^\beta \rightarrow P^\beta \bullet \nu, j, j, false]} \in \mathcal{D}_{LC_1^i}^{LA_\epsilon}$$

existe un paso

$$\overline{[O^\beta \rightarrow P^\beta \bullet \nu, j, j, false]} \in \mathcal{D}_{buLC_1^i}^{LC_\epsilon}$$

y, por tanto, existe la inferencia

$$[N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false] \vdash_{buLC_1^i} [O^\beta \rightarrow P^\beta \bullet \nu, j, j, false]$$

- Dado un paso

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false]}{[O^\beta \rightarrow \bullet P^\beta \nu, j, j, false]} \in \mathcal{D}_{LC_1^i}^{LA_{pre}}$$

existe un paso

$$\frac{}{[O^\beta \rightarrow \bullet P^\beta \nu, j, j, false]} \in \mathcal{D}_{buLC_1^i}^{LC_{pre}}$$

y, por tanto, existe la inferencia

$$[N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false] \vdash_{buLC_1^i} [O^\beta \rightarrow \bullet P^\beta \nu, j, j, false]$$

- Dado un paso

$$\frac{\begin{array}{l} [a, k, k+1] \\ [\top \rightarrow \mathbf{R}^\beta \bullet, j, k, false] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false] \end{array}}{[O^\gamma \rightarrow P^\gamma \bullet \nu, k, k+1, false]} \in \mathcal{D}_{LC_1^i}^{LF_t}$$

existe un paso

$$\frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\beta \bullet, j, k, false] \\ [a, k, k+1] \end{array}}{[O^\gamma \rightarrow P^\gamma \bullet \nu, k, k+1, false]} \in \mathcal{D}_{buLC_1^i}^{LAdj_t}$$

y, por tanto, existe la inferencia

$$\begin{array}{l} [a, k, k+1] \\ [\top \rightarrow \mathbf{R}^\beta \bullet, j, k, false] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false] \end{array} \vdash_{buLC_1^i} [O^\gamma \rightarrow P^\gamma \bullet \nu, k, k+1, false]$$

- Dado un paso

$$\frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\beta \bullet, j, k, false] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false] \end{array}}{[O^\gamma \rightarrow P^\gamma \bullet \nu, k, k, false]} \in \mathcal{D}_{LC_1^i}^{LF_e}$$

existe un paso

$$\frac{[\top \rightarrow \mathbf{R}^\beta \bullet, j, k, false]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, k, k, false]} \in \mathcal{D}_{buLC_1^i}^{LAdj_e}$$

y, por tanto, existe la inferencia

$$\begin{array}{l} [\top \rightarrow \mathbf{R}^\beta \bullet, j, k, false] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false] \end{array} \vdash_{buLC_1^i} [O^\gamma \rightarrow P^\gamma \bullet \nu, k, k, false]$$

- Dado un paso

$$\frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\beta \bullet, j, k, false] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false] \end{array}}{[O^\gamma \rightarrow \bullet P^\gamma \nu, k, k, false]} \in \mathcal{D}_{LC_1^i}^{LF_{pre}}$$

existe un paso

$$\frac{[\top \rightarrow \mathbf{R}^\beta \bullet, j, k, false]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, k, k, false]} \in \mathcal{D}_{buLC_1^i}^{LAdj_{pre}}$$

y, por tanto, existe la inferencia

$$\begin{array}{l} [\top \rightarrow \mathbf{R}^\beta \bullet, j, k, false] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \omega, i, j, false] \end{array} \vdash_{buLC_1^i} [O^\gamma \rightarrow \bullet P^\gamma \nu, k, k, false]$$

- Dado un paso

$$\frac{[M^\gamma \rightarrow \delta \bullet, k, l, false]}{[O^\beta \rightarrow P^\beta \bullet \nu, l, l, false]} \in \mathcal{D}_{LC_1^i}^{RA_\epsilon}$$

existe un paso

$$\overline{[O^\beta \rightarrow P^\beta \bullet \nu, l, l, false]} \in \mathcal{D}_{buLC_1^i}^{LC_\epsilon}$$

y, por tanto, existe la inferencia

$$[M^\gamma \rightarrow \delta \bullet, k, l, false] \vdash_{buLC_1^i} [O^\beta \rightarrow P^\beta \bullet \nu, l, l, false]$$

- Dado un paso

$$\frac{\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[a, j, j+1]}}{[O^\alpha \rightarrow P^\alpha \bullet \omega, j, j+1, false]} \in \mathcal{D}_{LC_1^i}^{LS_t}$$

existe un paso

$$\overline{[a, j, j+1]} \in \mathcal{D}_{buLC_1^i}^{LC_t}$$

y, por tanto, existe la inferencia

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[a, j, j+1]} \vdash_{buLC_1^i} [O^\alpha \rightarrow P^\alpha \bullet \omega, j, j+1, false]$$

- Dado un paso

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[O^\alpha \rightarrow P^\alpha \bullet \omega, j, j, false]} \in \mathcal{D}_{LC_1^i}^{LS_\epsilon}$$

existe un paso

$$\overline{[O^\alpha \rightarrow P^\alpha \bullet \omega, j, j, false]} \in \mathcal{D}_{buLC_1^i}^{LC_\epsilon}$$

y, por tanto, existe la inferencia

$$[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false] \vdash_{buLC_1^i} [O^\alpha \rightarrow P^\alpha \bullet \omega, j, j, false]$$

- Dado un paso

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[O^\alpha \rightarrow \bullet P^\alpha \omega, j, j, false]} \in \mathcal{D}_{LC_1^i}^{LS_{pre}}$$

existe un paso

$$\overline{[O^\alpha \rightarrow \bullet P^\alpha \omega, j, j, false]} \in \mathcal{D}_{buLC_1^i}^{LC_{pre}}$$

y, por tanto, existe la inferencia

$$[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false] \vdash_{buLC_1^i} [O^\alpha \rightarrow \bullet P^\alpha \omega, j, j, false]$$

- El resto de pasos deductivos cumplen las siguientes equivalencias:

$$\begin{aligned} \mathcal{D}_{LC^i}^{\text{Scan}} &= \mathcal{D}_{\text{buLC}^i}^{\text{Scan}} \\ \mathcal{D}_{LC^i}^e &= \mathcal{D}_{\text{buLC}^i}^e \\ \mathcal{D}_{LC^i}^{\text{LC}_n} &= \mathcal{D}_{\text{buLC}^i}^{\text{LC}_n} \\ \mathcal{D}_{LC^i}^{\text{Comp}_1} &= \mathcal{D}_{\text{buLC}^i}^{\text{Comp}} \\ \mathcal{D}_{LC^i}^{\text{RAjComp}} &= \mathcal{D}_{\text{buLC}^i}^{\text{RAj}} \\ \mathcal{D}_{LC^i}^{\text{SubsComp}} &= \mathcal{D}_{\text{buLC}^i}^{\text{Subs}} \end{aligned}$$

### 9.6.2 Integrando los esquemas tipo Left Corner en la red de analizadores para TIG

**Teorema 9.3** *Relación entre los esquemas  $\text{buE}^k$  y  $\text{buLC}^i$*

*Se mantiene la siguiente relación de contracción de secuencias deductivas:*

$$\text{buE}^k \xrightarrow{\text{sc}} \text{buLC}^i$$

*Prueba*

*Tenemos que demostrar que:*

1.  $\mathcal{I}_{\text{buLC}^i} \subseteq \mathcal{I}_{\text{buE}^k}$
2.  $\vdash_{\text{buLC}^i}^* \subseteq \vdash_{\text{buE}^k}^*$

*Lo primero es cierto, ya que el dominio del esquema  $\text{buLC}^i$  es igual al del esquema  $\text{buE}^k$  sin ítems con el punto al comienzo de la regla:*

$$\mathcal{I}_{\text{buLC}^i} \subset \mathcal{I}_{\text{buE}^k}$$

*Para probar que  $\vdash_{\text{buLC}^i}^* \subseteq \vdash_{\text{buE}^k}^*$  debemos demostrar que  $\mathcal{D}_{\text{buLC}^i} \subseteq \vdash_{\text{buE}^k}^*$ . Veamos cada uno de los pasos:*

- Un paso deductivo  $\mathcal{D}_{\text{buLC}^i}^{\text{LC}_t}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{buE}^k}^{\text{Ini}}$

$$\overline{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j, \text{false}]}$$

*seguido de otro  $\mathcal{D}_{\text{buE}^k}^{\text{Scan}}$*

$$\frac{\begin{array}{c} [a, j, j + 1] \\ [O^\gamma \rightarrow \bullet P^\gamma \nu, j, j, \text{false}] \end{array}}{[O^\gamma \rightarrow \delta P^\gamma \bullet \nu, j, j + 1, \text{false}]}$$

- Un paso deductivo  $\mathcal{D}_{\text{buLC}^i}^{\text{LC}_\epsilon}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{buE}^k}^{\text{Ini}}$

$$\overline{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j, \text{false}]}$$

seguido de otro  $\mathcal{D}_{\text{buE}^k}^\epsilon$

$$\frac{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j, \text{false}]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, j, \text{false}]}$$

- Un paso deductivo  $\mathcal{D}_{\text{buLC}^i}^{\text{LC}_{\text{subs}}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{buE}^k}^{\text{Ini}}$

$$\overline{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j, \text{false}]}$$

seguido de otro  $\mathcal{D}_{\text{buE}^k}^{\text{Subs}}$

$$\frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\alpha \bullet, j, k, \text{false}] \\ [O^\gamma \rightarrow \bullet P^\gamma \nu, j, j, \text{false}] \end{array}}{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, k, \text{false}]}$$

- Un paso deductivo  $\mathcal{D}_{\text{buLC}^i}^{\text{LC}_n}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{buE}^k}^{\text{Ini}}$

$$\overline{[Q^\gamma \rightarrow \bullet O^\gamma \omega, j, j, \text{false}]}$$

seguido de otro  $\mathcal{D}_{\text{buE}^k}^{\text{Comp}}$

$$\frac{\begin{array}{l} [O^\gamma \rightarrow \nu \bullet, j, k, \text{radj}] \\ [Q^\gamma \rightarrow \bullet O^\gamma \omega, j, j, \text{false}] \end{array}}{[Q^\gamma \rightarrow O^\gamma \bullet \omega, j, k, \text{false}]}$$

- Un paso deductivo  $\mathcal{D}_{\text{buLC}^i}^{\text{LAdj}_t}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{buE}^k}^{\text{LAdj}}$

$$\frac{[\top \rightarrow \mathbf{R}^\beta \bullet, i, j, \text{false}]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, i, j, \text{false}]}$$

seguido de otro  $\mathcal{D}_{\text{buE}^k}^{\text{Scan}}$

$$\frac{\begin{array}{l} [a, j, j + 1] \\ [O^\gamma \rightarrow \bullet P^\gamma \nu, i, j, \text{false}] \end{array}}{[O^\gamma \rightarrow \delta P^\gamma \bullet \nu, i, j + 1, \text{false}]}$$

- Un paso deductivo  $\mathcal{D}_{\text{buLC}^i}^{\text{LAdj}_\epsilon}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{buE}^k}^{\text{LAdj}}$

$$\frac{[\top \rightarrow \mathbf{R}^\beta \bullet, i, j, \text{false}]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, i, j, \text{false}]}$$

seguido de otro  $\mathcal{D}_{\text{buE}^k}^\epsilon$

$$\frac{[O^\gamma \rightarrow \bullet P^\gamma \nu, i, j, \text{false}]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, i, j, \text{false}]}$$

- Un paso deductivo  $\mathcal{D}_{\text{buLC}^i}^{\text{LAdjSubs}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{buE}^k}^{\text{LAdj}}$

$$\frac{[\top \rightarrow \mathbf{R}^\beta \bullet, i, j, \text{false}]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, i, j, \text{false}]}$$

seguido de otro  $\mathcal{D}_{\text{buE}^k}^{\text{Subs}}$

$$\frac{[\top \rightarrow \mathbf{R}^\alpha \bullet, j, k, \text{false}]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, i, j, \text{false}]} \\ \frac{[O^\gamma \rightarrow \bullet P^\gamma \nu, i, j, \text{false}]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, i, k, \text{false}]}$$

- Un paso deductivo  $\mathcal{D}_{\text{buLC}^i}^{\text{LAdjIn}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\text{buE}^k}^{\text{LAdj}}$

$$\frac{[\top \rightarrow \mathbf{R}^\beta \bullet, i, j, \text{false}]}{[Q^\gamma \rightarrow \bullet O^\gamma \nu, i, j, \text{false}]}$$

seguido de otro  $\mathcal{D}_{\text{buE}^k}^{\text{Comp}}$

$$\frac{[O^\gamma \rightarrow \nu \bullet, j, k, \text{radj}]}{[Q^\gamma \rightarrow \bullet O^\gamma \omega, i, j, \text{false}]} \\ \frac{[Q^\gamma \rightarrow \bullet O^\gamma \omega, i, j, \text{false}]}{[Q^\gamma \rightarrow O^\gamma \bullet \omega, i, k, \text{false}]}$$

- El resto de pasos deductivos del esquema  $\text{buLC}^i$  están incluidos en las inferencias de sus homónimos del esquema  $\text{buE}^k$ , ya que estos últimos contemplan las operaciones en cualquier posición, mientras que los pertenecientes a  $\text{buLC}^i$  sólo son aplicables en nodos que no sean hijos izquierdos. De manera que se cumple:

$$\mathcal{D}_{\text{buLC}^i}^{\text{Scan}} \subset \mathcal{D}_{\text{buE}^k}^{\text{Scan}}$$

$$\mathcal{D}_{\text{buLC}^i}^\epsilon \subset \mathcal{D}_{\text{buE}^k}^\epsilon$$

$$\mathcal{D}_{\text{buLC}^i}^{\text{Comp}} \subset \mathcal{D}_{\text{buE}^k}^{\text{Comp}}$$

$$\mathcal{D}_{\text{buLC}^i}^{\text{Subs}} \subset \mathcal{D}_{\text{buE}^k}^{\text{Subs}}$$

**Teorema 9.4** Relación entre los esquemas  $\text{Earley}^i$  y  $\text{LC}^i$

Se mantiene la siguiente relación de contracción de secuencias deductivas:

$$\text{Earley}^i \xrightarrow{\text{sc}} \text{LC}^i$$

*Prueba*

Tenemos que demostrar que:

$$1. \mathcal{I}_{\text{LC}^i} \subseteq \mathcal{I}_{\text{Earley}^i}$$

$$2. \vdash_{\text{LC}^i}^* \subseteq \vdash_{\text{Earley}^i}^*$$

Lo primero es cierto, ya que el dominio del esquema  $\mathbf{LC}^i$  es igual al del esquema  $\mathbf{Earley}^i$  sin el subconjunto de los ítems con el punto al comienzo de la regla:

$$\mathcal{I}_{\mathbf{LC}^i} \subset \mathcal{I}_{\mathbf{Earley}^i}$$

Para probar que  $\vdash_{\mathbf{LC}^i}^* \subseteq \vdash_{\mathbf{Earley}^i}^*$  debemos demostrar que  $\mathcal{D}_{\mathbf{LC}^i} \subseteq \mathcal{D}_{\mathbf{Earley}^i}^*$ . Veamos cada uno de los pasos:

- Un paso  $\mathcal{D}_{\mathbf{LC}^i}^{\text{LLt}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\mathbf{Earley}^i}^{\text{Ini}}$

$$\overline{[\top \rightarrow \bullet \mathbf{R}^\alpha, 0, 0, \text{false}]}$$

seguido de una secuencia de pasos  $\mathcal{D}_{\mathbf{Earley}^i}^{\text{Pred}}$

$$\frac{[\top \rightarrow \bullet \mathbf{R}^\alpha, 0, 0, \text{false}]}{\overline{[\mathbf{R}^\alpha \rightarrow \bullet M^\gamma \nu, 0, 0, \text{false}]}}$$

$$\frac{[\mathbf{R}^\alpha \rightarrow \bullet M^\gamma \nu, 0, 0, \text{false}]}{\overline{[M^\gamma \rightarrow \bullet O^\gamma \nu, 0, 0, \text{false}]}}$$

$$\frac{[M^\gamma \rightarrow \bullet O^\gamma \nu, 0, 0, \text{false}]}{\overline{[O^\gamma \rightarrow \bullet P^\gamma \delta, 0, 0, \text{false}]}}$$

y de un paso  $\mathcal{D}_{\mathbf{Earley}^i}^{\text{Scan}}$

$$\frac{[O^\gamma \rightarrow \bullet P^\gamma \nu, 0, 0, \text{false}]}{[a, 0, 1]}$$

$$\overline{[O^\gamma \rightarrow P^\gamma \bullet \nu, 0, 1, \text{false}]}$$

- Un paso  $\mathcal{D}_{\mathbf{LC}^i}^{\text{LLe}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\mathbf{Earley}^i}^{\text{Ini}}$

$$\overline{[\top \rightarrow \bullet \mathbf{R}^\alpha, 0, 0, \text{false}]}$$

seguido de una secuencia de pasos  $\mathcal{D}_{\mathbf{Earley}^i}^{\text{Pred}}$

$$\frac{[\top \rightarrow \bullet \mathbf{R}^\alpha, 0, 0, \text{false}]}{\overline{[\mathbf{R}^\alpha \rightarrow \bullet M^\gamma \nu, 0, 0, \text{false}]}}$$

$$\frac{[\mathbf{R}^\alpha \rightarrow \bullet M^\gamma \nu, 0, 0, \text{false}]}{\overline{[M^\gamma \rightarrow \bullet O^\gamma \nu, 0, 0, \text{false}]}}$$

$$\frac{[M^\gamma \rightarrow \bullet O^\gamma \nu, 0, 0, \text{false}]}{\overline{[O^\gamma \rightarrow \bullet P^\gamma \delta, 0, 0, \text{false}]}}$$

y de un paso  $\mathcal{D}_{\mathbf{Earley}^i}^\epsilon$

$$\frac{[O^\gamma \rightarrow \bullet P^\gamma \nu, 0, 0, \text{false}]}{\overline{[O^\gamma \rightarrow P^\gamma \bullet \nu, 0, 0, \text{false}]}}$$

- Un paso  $\mathcal{D}_{\mathbf{LC}^i}^{\text{LLpre}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{\mathbf{Earley}^i}^{\text{Ini}}$

$$\overline{[\top \rightarrow \bullet \mathbf{R}^\alpha, 0, 0, \text{false}]}$$

seguido de una secuencia de pasos  $\mathcal{D}_{\mathbf{Earley}^i}^{\text{Pred}}$

$$\frac{[\top \rightarrow \bullet \mathbf{R}^\alpha, 0, 0, \text{false}]}{\overline{[\mathbf{R}^\alpha \rightarrow \bullet M^\gamma \nu, 0, 0, \text{false}]}}$$

$$\frac{[\mathbf{R}^\alpha \rightarrow \bullet M^\gamma \nu, 0, 0, \text{false}]}{\overline{[M^\gamma \rightarrow \bullet O^\gamma \nu, 0, 0, \text{false}]}}$$

$$\frac{[M^\gamma \rightarrow \bullet O^\gamma \nu, 0, 0, \text{false}]}{\overline{[O^\gamma \rightarrow \bullet P^\gamma \delta, 0, 0, \text{false}]}}$$

- Un paso  $\mathcal{D}_{LC^i}^{LC^t}$  es equivalente a una secuencia de pasos  $\mathcal{D}_{Earley^i}^{Pred}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[M^\gamma \rightarrow \bullet O^\gamma \omega, j, j, false]}$$

$$\frac{[M^\gamma \rightarrow \bullet O^\gamma \omega, j, j, false]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j, false]}$$

seguida de un paso  $\mathcal{D}_{Earley^i}^{Scan}$

$$\frac{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j, false]}{[a, j, j + 1]}$$

$$\frac{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, j + 1, false]}{}$$

- Un paso  $\mathcal{D}_{LC^i}^{LC^e}$  es equivalente a una secuencia de pasos  $\mathcal{D}_{Earley^i}^{Pred}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[M^\gamma \rightarrow \bullet O^\gamma \omega, j, j, false]}$$

$$\frac{[M^\gamma \rightarrow \bullet O^\gamma \omega, j, j, false]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j, false]}$$

seguida de un paso  $\mathcal{D}_{Earley^i}^e$

$$\frac{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j, false]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, j, false]}$$

- Un paso  $\mathcal{D}_{LC^i}^{LC^{pre}}$  es equivalente a una secuencia de pasos  $\mathcal{D}_{Earley^i}^{Pred}$ :

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[M^\gamma \rightarrow \bullet O^\gamma \omega, j, j, false]}$$

$$\frac{[M^\gamma \rightarrow \bullet O^\gamma \omega, j, j, false]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j, false]}$$

- Un paso  $\mathcal{D}_{LC^i}^{LC^n}$  es equivalente a una secuencia de pasos  $\mathcal{D}_{Earley^i}^{Pred}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[M^\gamma \rightarrow \bullet O^\gamma \omega, j, j, false]}$$

$$\frac{[M^\gamma \rightarrow \bullet O^\gamma \omega, j, j, false]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j, false]}$$

seguida de un paso  $\mathcal{D}_{Earley^i}^{Comp}$

$$\frac{[P^\gamma \rightarrow \delta \bullet, j, k, radj]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, j, j, false]}$$

$$\frac{[O^\gamma \rightarrow P^\gamma \bullet \nu, j, k, false]}{}$$



- Un paso  $\mathcal{D}_{LC^i}^{LA_t}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{Earley^i}^{LAdjPred}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[\top \rightarrow \bullet R^\beta, j, j, false]}$$

una secuencia de pasos  $\mathcal{D}_{Earley^i}^{Pred}$

$$\frac{[\top \rightarrow \bullet R^\beta, j, j, false]}{[R^\beta \rightarrow \bullet M^\beta \omega, j, j, false]}$$

$$\frac{[M^\beta \rightarrow \bullet O^\beta \omega, j, j, false]}{[O^\beta \rightarrow \bullet P^\beta \nu, j, j, false]}$$

y un paso  $\mathcal{D}_{Earley^i}^{Scan}$

$$\frac{[O^\beta \rightarrow \bullet P^\beta \nu, j, j, false], \\ [a, j, j + 1]}{[O^\beta \rightarrow P^\beta \bullet \nu, j, j + 1, false]}$$

- Un paso  $\mathcal{D}_{LC^i}^{LA_\epsilon}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{Earley^i}^{LAdjPred}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[\top \rightarrow \bullet R^\beta, j, j, false]}$$

una secuencia de pasos  $\mathcal{D}_{Earley^i}^{Pred}$

$$\frac{[\top \rightarrow \bullet R^\beta, j, j, false]}{[R^\beta \rightarrow \bullet M^\beta \omega, j, j, false]}$$

$$\frac{[M^\beta \rightarrow \bullet O^\beta \omega, j, j, false]}{[O^\beta \rightarrow \bullet P^\beta \nu, j, j, false]}$$

y un paso  $\mathcal{D}_{Earley^i}^\epsilon$

$$\frac{[O^\beta \rightarrow \bullet P^\beta \nu, j, j, false]}{[O^\beta \rightarrow P^\beta \bullet \nu, j, j, false]}$$

- Un paso  $\mathcal{D}_{LC^i}^{LA_{pro}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{Earley^i}^{LAdjPred}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[\top \rightarrow \bullet R^\beta, j, j, false]}$$

seguido de una secuencia de pasos  $\mathcal{D}_{Earley^i}^{Pred}$

$$\frac{[\top \rightarrow \bullet R^\beta, j, j, false]}{[R^\beta \rightarrow \bullet M^\beta \omega, j, j, false]}$$

$$\frac{[M^\beta \rightarrow \bullet O^\beta \omega, j, j, false]}{[O^\beta \rightarrow \bullet P^\beta \nu, j, j, false]}$$

- Un paso  $\mathcal{D}_{LC^i}^{LF_t}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{Earley^i}^{LAdjComp}$

$$\frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\beta \bullet, j, k, false] \\ [O^\gamma \rightarrow \delta \bullet N^\gamma \nu, i, j, false] \end{array}}{[N^\gamma \rightarrow \bullet M^\gamma \nu, j, k, false]}$$

una secuencia de pasos  $\mathcal{D}_{Earley^i}^{Pred}$

$$\frac{[N^\gamma \rightarrow \bullet M^\gamma \nu, k, k, false]}{[M^\gamma \rightarrow \bullet O^\gamma \omega, k, k, false]}$$

$$\frac{[M^\gamma \rightarrow \bullet O^\gamma \omega, k, k, false]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, k, k, false]}$$

y un paso  $\mathcal{D}_{Earley^i}^{Scan}$

$$\frac{\begin{array}{l} [O^\gamma \rightarrow \bullet P^\gamma \nu, k, k, false] \\ [a, k, k + 1] \end{array}}{[O^\gamma \rightarrow P^\gamma \bullet \nu, k, k + 1, false]}$$

- Un paso  $\mathcal{D}_{LC^i}^{LF_\epsilon}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{Earley^i}^{LAdjComp}$

$$\frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\beta \bullet, j, k, false] \\ [O^\gamma \rightarrow \delta \bullet N^\gamma \nu, i, j, false] \end{array}}{[N^\gamma \rightarrow \bullet M^\gamma \nu, j, k, false]}$$

una secuencia de pasos  $\mathcal{D}_{Earley^i}^{Pred}$

$$\frac{[N^\gamma \rightarrow \bullet M^\gamma \nu, k, k, false]}{[M^\gamma \rightarrow \bullet O^\gamma \omega, k, k, false]}$$

$$\frac{[M^\gamma \rightarrow \bullet O^\gamma \omega, k, k, false]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, k, k, false]}$$

y un paso  $\mathcal{D}_{Earley^i}^\epsilon$

$$\frac{[O^\gamma \rightarrow \bullet P^\gamma \nu, k, k, false]}{[O^\gamma \rightarrow P^\gamma \bullet \nu, k, k, false]}$$

- Un paso  $\mathcal{D}_{LC^i}^{LF_{pre}}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{Earley^i}^{LAdjComp}$

$$\frac{\begin{array}{l} [\top \rightarrow \mathbf{R}^\beta \bullet, j, k, false] \\ [O^\gamma \rightarrow \delta \bullet N^\gamma \nu, i, j, false] \end{array}}{[N^\gamma \rightarrow \bullet M^\gamma \nu, j, k, false]}$$

seguido de una secuencia de pasos  $\mathcal{D}_{Earley^i}^{Pred}$

$$\frac{[N^\gamma \rightarrow \bullet M^\gamma \nu, k, k, false]}{[M^\gamma \rightarrow \bullet O^\gamma \omega, k, k, false]}$$

$$\frac{[M^\gamma \rightarrow \bullet O^\gamma \omega, k, k, false]}{[O^\gamma \rightarrow \bullet P^\gamma \nu, k, k, false]}$$



- Un paso  $\mathcal{D}_{LC^i}^{RA_\varepsilon}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{Earley^i}^{RAjPred}$

$$\frac{[N^\gamma \rightarrow \delta \bullet, i, j, false]}{[\top \rightarrow \bullet R^\beta, j, j, false]}$$

una secuencia de pasos  $\mathcal{D}_{Earley^i}^{Pred}$

$$\frac{[\top \rightarrow \bullet R^\beta, j, j, false]}{[R^\beta \rightarrow \bullet M^{\beta\omega}, j, j, false]}$$

$$\frac{[M^\beta \rightarrow \bullet O^{\beta\omega}, j, j, false]}{[O^\beta \rightarrow \bullet P^{\beta\nu}, j, j, false]}$$

y un paso  $\mathcal{D}_{Earley^i}^\varepsilon$

$$\frac{[O^\beta \rightarrow \bullet P^{\beta\nu}, j, j, false]}{[O^\beta \rightarrow P^\beta \bullet \nu, j, j, false]}$$

- Un paso  $\mathcal{D}_{LC^i}^{LS_t}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{Earley^i}^{SubsPred}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[\top \rightarrow \bullet R^\alpha, j, j, false]}$$

una secuencia de pasos  $\mathcal{D}_{Earley^i}^{Pred}$

$$\frac{[\top \rightarrow \bullet R^\alpha, j, j, false]}{[R^\alpha \rightarrow \bullet M^{\alpha\omega}, j, j, false]}$$

$$\frac{[M^\alpha \rightarrow \bullet O^{\alpha\omega}, j, j, false]}{[O^\alpha \rightarrow \bullet P^{\alpha\nu}, j, j, false]}$$

y un paso  $\mathcal{D}_{Earley^i}^{Scan}$

$$\frac{[O^\alpha \rightarrow \bullet P^{\alpha\nu}, j, j, false], [a, j, j + 1]}{[O^\alpha \rightarrow P^\alpha \bullet \nu, j, j + 1, false]}$$

- Un paso  $\mathcal{D}_{LC^i}^{LS_\varepsilon}$  es equivalente a la secuencia formada por un paso  $\mathcal{D}_{Earley^i}^{SubsPred}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[\top \rightarrow \bullet R^\alpha, j, j, false]}$$

una secuencia de pasos  $\mathcal{D}_{Earley^i}^{Pred}$

$$\frac{[\top \rightarrow \bullet R^\alpha, j, j, false]}{[R^\alpha \rightarrow \bullet M^{\alpha\omega}, j, j, false]}$$

$$\frac{[M^\alpha \rightarrow \bullet O^{\alpha\omega}, j, j, false]}{[O^\alpha \rightarrow \bullet P^{\alpha\nu}, j, j, false]}$$

y un paso  $\mathcal{D}_{Earley^i}^\varepsilon$

$$\frac{[O^\alpha \rightarrow \bullet P^{\alpha\nu}, j, j, false]}{[O^\alpha \rightarrow P^\alpha \bullet \nu, j, j, false]}$$

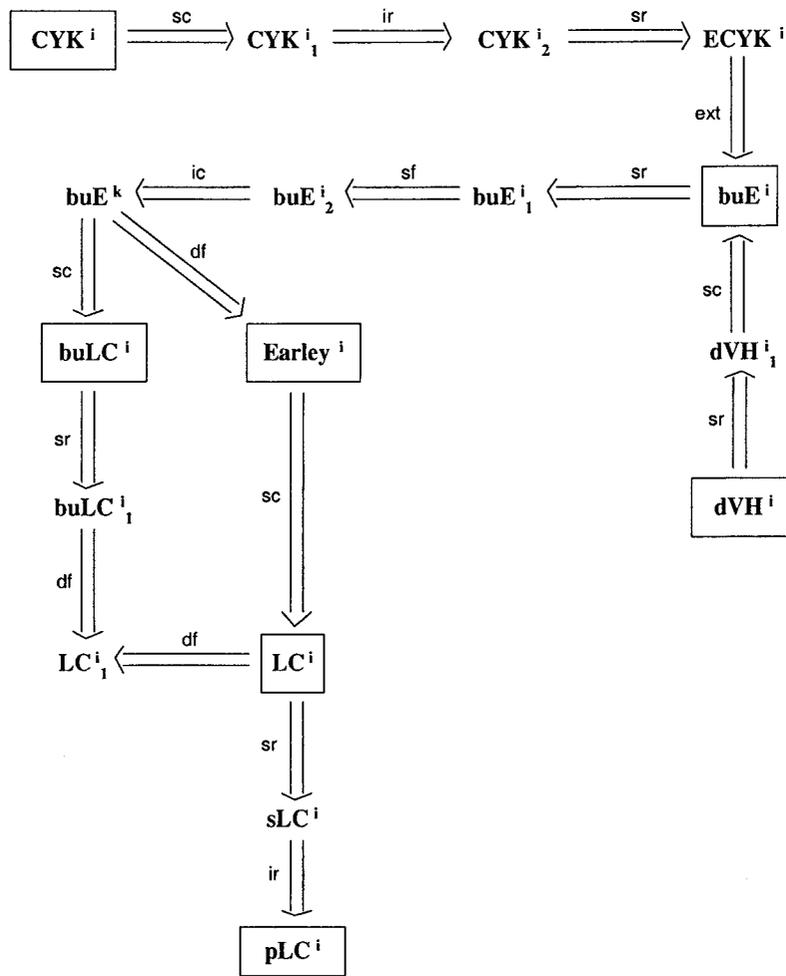


Figura 9.1: Red ampliada de analizadores para TIG

- Un paso  $D_{LC^i}^{LS_{pre}}$  es equivalente a la secuencia formada por un paso  $D_{Earley^i}^{SubsPred}$

$$\frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[\top \rightarrow \bullet R^\alpha, j, j, false]}$$

seguido de una secuencia de pasos  $D_{Earley^i}^{Pred}$

$$\frac{[\top \rightarrow \bullet R^\alpha, j, j, false]}{[R^\alpha \rightarrow \bullet M^\alpha \omega, j, j, false]}$$

$$\frac{[R^\alpha \rightarrow \bullet M^\alpha \omega, j, j, false]}{[M^\alpha \rightarrow \bullet O^\alpha \omega, j, j, false]}$$

$$\frac{[M^\alpha \rightarrow \bullet O^\alpha \omega, j, j, false]}{[O^\alpha \rightarrow \bullet P^\alpha \nu, j, j, false]}$$

- Los pasos deductivos de reconocimiento del esquema  $LC^i$  están incluidos en las inferencias de sus homónimos del esquema  $Earley^i$ , ya que estos últimos contemplan

las operaciones en cualquier posición, mientras que los pertenecientes a  $LC^i$  sólo son aplicables en nodos que no sean hijos izquierdos. Luego se cumple:

$$\mathcal{D}_{LC^i}^{\text{Scan}} \subset \mathcal{D}_{\text{Earley}^i}^{\text{Scan}}$$

$$\mathcal{D}_{LC^i}^{\varepsilon} \subset \mathcal{D}_{\text{Earley}^i}^{\varepsilon}$$

- El resto de pasos son idénticos en ambos esquemas:

$$\mathcal{D}_{LC^i}^{\text{Comp}} = \mathcal{D}_{\text{Earley}^i}^{\text{Comp}}$$

$$\mathcal{D}_{LC^i}^{\text{RAdjComp}} = \mathcal{D}_{\text{Earley}^i}^{\text{RAdjComp}}$$

$$\mathcal{D}_{LC^i}^{\text{SubsComp}} = \mathcal{D}_{\text{Earley}^i}^{\text{SubsComp}}$$

En la figura 9.1 se muestra la red de analizadores para TIG del capítulo anterior ampliada con los esquemas y las relaciones presentados en este capítulo.

## Parte IV

# Análisis sintáctico de TAG dirigido por el núcleo TIG

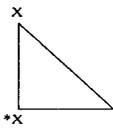
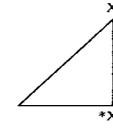
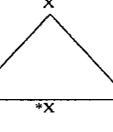
## Capítulo 10

# Integrando analizadores sintácticos para TIG y TAG

La importancia fundamental del formalismo TIG se basa en el hecho de que la mayor parte de las estructuras de los árboles definidos para gramáticas de amplia cobertura se corresponden con estructuras de árboles TIG. Un claro ejemplo lo tenemos en la gramática del inglés XTAG [144], donde, según se indica en [114], el 99% de los árboles y adjunciones permitidas existentes se encuentran dentro del formalismo TIG. En la figura 10.1 mostramos un estudio detallado de la distribución de los 650 árboles auxiliares que existen en la última versión de esta gramática de amplia cobertura.

En este capítulo proponemos un conjunto de analizadores para TAG cuyo objetivo es sacar provecho de esta circunstancia, de forma que durante su ejecución se comporten de forma dinámica, trabajando como analizadores para TIG cuando analizan árboles TAG que cumplen las restricciones impuestas en el formalismo TIG o como analizadores para TAG cuando analizan otro tipo de árboles. Es evidente, al tratarse de analizadores para TAG, que sus cotas de complejidad teórica permanecen en las estándar para las mismas, es decir,  $\mathcal{O}(n^6)$  la temporal y  $\mathcal{O}(n^4)$  la espacial. Sin embargo, veremos como la complejidad temporal se ve significativamente reducida cuando los algoritmos analizan árboles con estructuras TIG.

El capítulo se organiza de la siguiente forma. En la primera sección se adoptan las convenciones que se emplearan para las definiciones de los analizadores, y en las siguientes secciones se presentan un conjunto de analizadores sin VPP combinados para TAG y TIG similar a los presentados anteriormente para cada uno de estos formalismos, es decir, tres con estrategia ascendente (CYK, tipo Earley y bidireccional de tipo dVH) y uno ascendente con predicciones tipo Earley sin VPP. Para el algoritmo de tipo Earley también incluimos una comparativa de prestaciones respecto al mismo para gramáticas TAG. Mostraremos, en general, que estos esquemas son contracciones de ítems de los definidos para TAG, para lo que probaremos la relación existente entre los dos algoritmos basados en CYK.

Forma del árbol	% del total	Distribución	
		Nodo raíz	%
	81 %	Nodo raíz	%
		NP	87
		VP, S	4
		Resto	9*
	15 %	Nodo raíz	%
		S	43
		VP	21
		Resto	36*
	4 %	Estructura	%
		Pasivas	10
		Conjunciones	53
		Puntuación	34
		Adjetivos Comparativos	3

\* Distribuidos a una media del 5% por cada tipo de nodo raíz

Figura 10.1: Distribución de los árboles auxiliares en XTAG

## 10.1 Convenciones y notación

Nuestro objetivo es integrar la funcionalidad de los esquemas presentados para TIG con sus homónimos para TAG. Recordemos las diferencias fundamentales entre TIG y TAG para tomar las decisiones oportunas de cara a realizar la mencionada integración.

Desde el punto de vista de los árboles elementales, los árboles TIG son un subconjunto de los árboles TAG que cumplen una serie de restricciones. Cuando definimos el formalismo TIG dividimos los árboles elementales en tres grupos: iniciales (*I*), auxiliares izquierdos (*A<sub>L</sub>*) y auxiliares derechos (*A<sub>R</sub>*), puesto que la definición no permitía otro tipo de estructura distinta a la de éstos. Sin embargo, el formalismo TAG también permite árboles auxiliares wrapping, incluyendo en este grupo tanto los árboles auxiliares con nodos frontera no vacíos a la derecha e izquierda del nodo pie, como aquellos árboles que no cumplen las restricciones de adjunción impuestas en la definición de TIG, aunque presenten estructura de árbol auxiliar izquierdo o derecho.

Puesto que nuestro objetivo es que el análisis varíe de forma dinámica según se aplique sobre árboles TIG o TAG, debemos identificar dentro del conjunto de árboles elementales de cualquier gramática TAG sobre la que se va aplicar este tipo de analizadores cual es el núcleo de árboles que responden a la definición de TIG. Para ello dividimos el conjunto de árboles elementales en cuatro grupos disjuntos: iniciales (*I*), *auxiliares fuertemente izquierdos* (*A<sub>SL</sub>*), *auxiliares fuertemente derechos* (*A<sub>SR</sub>*) y auxiliares wrapping (*A<sub>W</sub>*). Un árbol auxiliar es fuertemente izquierdo (resp. derecho) si todos sus nodos frontera no vacíos se

encuentran a la izquierda (resp. derecha) de su nodo pie y prohíbe: (1) la adjunción en su nodo pie, (2) la adjunción en los nodos situados a la derecha (resp. izquierda) de la espina, y (3) la adjunción de árboles fuertemente derechos (resp. izquierdos) y árboles auxiliares wrapping en su espina. Los árboles auxiliares de wrapping son aquellos que no se encuentran en el conjunto  $A_{SL} \cup A_{SR}$ .

Estos cambios nos obligan a incluir alguna notación nueva en las TAG. El conjunto de árboles auxiliares de una gramática TAG será  $A = A_{SL} \cup A_{SR} \cup A_W$ . Si un árbol auxiliar  $\beta$  es un árbol auxiliar fuertemente izquierdo, entonces  $\beta \in A_{SL}$ . Si un árbol auxiliar  $\beta$  es un árbol auxiliar fuertemente derecho, entonces  $\beta \in A_{SR}$ . Si un árbol auxiliar fuertemente izquierdo  $\beta \in A_{SL}$  es adjuntable en un nodo  $M^\gamma$  lo denotamos como  $\beta \in \text{ladj}(M^\gamma)$ . Si un árbol auxiliar fuertemente derecho  $\beta \in A_{SR}$  es adjuntable en un nodo  $M^\gamma$  lo denotamos como  $\beta \in \text{radj}(M^\gamma)$ . Si un árbol auxiliar wrapping  $\beta \in A_W$  es adjuntable en un nodo  $M^\gamma$  lo denotamos como  $\beta \in \text{wadj}(M^\gamma)$ . Denominamos *adjunción izquierda* (resp. *derecha*) a la adjunción de un árbol auxiliar fuertemente izquierdo (resp. derecho) en un nodo adjuntable. Denominamos *adjunción wrapping* a la adjunción de un árbol auxiliar wrapping en un nodo adjuntable. Cuando la adjunción izquierda (resp. derecha) no es obligatoria en un nodo  $M^\gamma$  lo denotamos como  $\text{nil} \in \text{ladj}(M^\gamma)$  (resp.  $\text{nil} \in \text{radj}(M^\gamma)$ ). Cuando la adjunción wrapping no es obligatoria en un nodo  $M^\gamma$  lo denotamos como  $\text{nil} \in \text{wadj}(M^\gamma)$ . Cuando un nodo  $M^\gamma$  no presenta adjunciones izquierda, derecha y wrapping obligatorias ( $\text{nil} \in \text{ladj}(M^\gamma), \text{nil} \in \text{radj}(M^\gamma), \text{nil} \in \text{wadj}(M^\gamma)$ ) lo denotamos como  $\text{nil} \in \text{adj}(M^\gamma)$ .

Respecto a la operación de composición básica, mientras las TAG sólo permiten una adjunción sobre un nodo, la versión de TIG que hemos empleado a lo largo de esta memoria permite la adjunción de un árbol auxiliar izquierdo y otro derecho sobre un nodo. En este punto se plantean dos alternativas: limitamos a una adjunción en un nodo como obliga el formalismo TAG, o permitimos varias adjunciones como propone el formalismo TIG. Hemos elegido la primera opción fundamentalmente porque no podemos olvidar que nuestro objetivo en este capítulo es diseñar analizadores eficientes para TAG y, en la medida de lo posible, debemos adoptar las restricciones de las mismas. Obviamente esta decisión va a provocar ciertas limitaciones a la hora de definir ciertas formaciones lingüísticas con árboles elementales de tipo TIG. Por ejemplo, árboles derivados como los que se muestran en la figura 10.2 no se podrían obtener si un árbol auxiliar fuera fuertemente izquierdo y otro fuertemente derecho, debido a que sólo se puede adjuntar uno de ellos, el cual precisamente prohíbe la adjunción del otro en su espina. Esto no quiere decir que estos tipos de árboles derivados no sean factibles, simplemente permitiendo la adjunción del árbol fuertemente derecho (resp. izquierdo) en la raíz del fuertemente izquierdo (resp. derecho) ya se podrían derivar. Lo que ocurre es que el árbol fuertemente izquierdo (resp. derecho) dejaría de serlo para convertirse en un árbol auxiliar de wrapping y, por consiguiente, su coste computacional de análisis sería superior, como veremos más adelante.

En las definiciones de analizadores para TIG incluimos la operación de sustitución. Sin embargo, por claridad, obviamos dicha operación en los algoritmos para TAG. Dado que

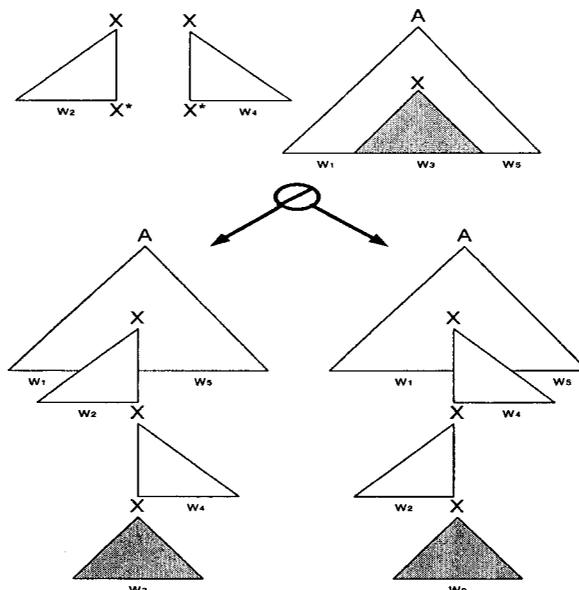


Figura 10.2: *Adjunción de dos árboles auxiliares en un nodo*

la sustitución no modifica la potencia de ninguno de ambos formalismos y puesto que ya la hemos introducido en los esquemas para TIG, vamos a incluirla en los analizadores que presentamos en este capítulo. Así cuando un nodo frontera  $M^\gamma$  se puede sustituir por un árbol inicial  $\alpha$ , lo denotaremos como  $\alpha \in \text{subst}(M^\gamma)$ .

Las convenciones y notación adoptadas van a influir de cara a la fusión de los esquemas homónimos para TAG y TIG, especialmente en estos últimos. De manera que por una mayor claridad sería conveniente redefinir los analizadores para TIG introducidos en los capítulos previos para que recojan, por un lado, la adjunción simple en un nodo y, por otro, los cambios de notación provocados por los nuevos grupos de árboles auxiliares. Teniendo en cuenta que el interés fundamental de este tema no son estas adaptaciones, hemos optado por situar en el apéndice B las redefiniciones de estos esquemas.

## 10.2 Esquema tipo CYK

El esquema  $\text{CYK}^{\text{Mix}}$  es el resultado de la fusión de los esquemas  $\text{CYK}^j$  y  $\text{CYK}$ . Para conseguir un analizador más eficiente que el ya definido para TAG, observemos el paso de adjunción del algoritmo para TAG  $\mathcal{D}_{\text{CYK}}^{\text{Adj}}$ . Este paso presenta una complejidad temporal de  $\mathcal{O}(n^6)$ , debido a que en sus antecedentes se incorpora la información referente al fragmento de la cadena de entrada reconocido por el subárbol que pende del nodo pie del árbol auxiliar que se va a adjuntar, concretamente los índices  $k$  y  $l$ . Sin embargo, cuando se trata de árboles auxiliares fuertemente izquierdos o derechos no es necesario transmitir dicha información.

Teniendo ésto en cuenta, el esquema  $\text{CYK}^{\text{Mix}}$  toma como base el algoritmo para TAG, pero el paso de adjunción lo divide en tres, uno para cada tipo de árbol auxiliar. De esta forma disminuye la complejidad teórica cuando realiza adjunciones izquierdas o derechas y, por tanto, reduce la complejidad teórica media.

El conjunto de ítems del esquema  $\mathcal{I}_{\text{CYK}^{\text{Mix}}}$  se define mediante:

$$\mathcal{I}_{\text{CYK}^{\text{Mix}}} = [N^\gamma, i, j \mid p, q \mid \text{adj}]$$

donde  $\text{label}(N^\gamma) \in V_N$  y  $\gamma \in I \cup A$ . Los índices  $0 \leq i \leq j$  establecen las posiciones dentro de la cadena de entrada que delimitan el fragmento reconocido por  $N^\gamma$ . Si  $p$  y  $q$  presentan valores definidos, entonces  $N^\gamma$  pertenece a la espina de un árbol auxiliar wrapping ( $\gamma \in \mathbf{A}_W$ ) y se cumple  $(p, q) \leq (i, j)$ . El componente booleano  $\text{adj} \in \{\text{true}, \text{false}\}$  permite al analizador determinar si se ha llevado a cabo una adjunción (izquierda, derecha o wrapping) sobre el nodo  $N^\gamma$ .

Las reglas deductivas son las siguientes:

$$\mathcal{D}_{\text{CYK}^{\text{Mix}}} = \mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{Scan}} \cup \mathcal{D}_{\text{CYK}^{\text{Mix}}}^\varepsilon \cup \mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{Foot}} \cup \mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{CompUna}} \cup \mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{CompBin}} \cup \mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{LAdj}} \cup \mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{RAAdj}} \cup \mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{WAdj}} \cup \mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{Subs}}$$

$$\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{Scan}} = \frac{[a, j, j+1]}{[N^\gamma, j, j+1 \mid -, - \mid \text{false}]} \quad N^\gamma \rightarrow a \in \mathcal{P}(\gamma)$$

$$\mathcal{D}_{\text{CYK}^{\text{Mix}}}^\varepsilon = \frac{N^\gamma \rightarrow \varepsilon \in \mathcal{P}(\gamma)}{[N^\gamma, j, j \mid -, - \mid \text{false}]} \quad \text{o } (\text{label}(N^\gamma) = \perp \text{ y } \gamma \in \mathbf{A}_{SL} \cup \mathbf{A}_{SR})$$

$$\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{Foot}} = \frac{\text{label}(P^\beta) = \perp}{[P^\beta, k, l \mid k, l \mid \text{false}]} \quad \beta \in \mathbf{A}_W$$

$$\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{CompUna}} = \frac{[O^\gamma, i, j \mid p, q \mid \text{adj}]}{[M^\gamma, i, j \mid p, q \mid \text{false}]} \quad M^\gamma \rightarrow O^\gamma \in \mathcal{P}(\gamma)$$

donde se debe cumplir que  $(\text{nil} \in \text{adj}(O^\gamma) \text{ y } \text{adj} = \text{false}) \text{ o } (\exists \beta \in \text{adj}(O^\gamma) \text{ y } \text{adj} = \text{true})$ .

$$\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{CompBin}} = \frac{\begin{array}{l} [O_1^\gamma, i, j \mid p, q \mid \text{adj}] \\ [O_2^\gamma, j, k \mid p', q' \mid \text{adj}'] \end{array}}{[M^\gamma, i, k \mid p \cup p' \mid q \cup q' \mid \text{false}]} \quad M^\gamma \rightarrow O_1^\gamma O_2^\gamma \in \mathcal{P}(\gamma)$$

donde se debe cumplir:

- (i)  $(\text{nil} \in \text{adj}(O_1^\gamma) \text{ y } \text{adj} = \text{false}) \text{ o } (\exists \beta \in \text{adj}(O_1^\gamma) \text{ y } \text{adj} = \text{true})$ ,
- (ii)  $(\text{nil} \in \text{adj}(O_2^\gamma) \text{ y } \text{adj} = \text{false}) \text{ o } (\exists \beta \in \text{adj}(O_2^\gamma) \text{ y } \text{adj} = \text{true})$ .

$$\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{LAdj}} = \frac{\begin{array}{l} [P^\beta, i, j \mid -, - \mid \text{false}] \\ [M^\gamma, j, k \mid p, q \mid \text{false}] \end{array}}{[M^\gamma, i, k \mid p, q \mid \text{true}]} \quad \begin{array}{l} \text{label}(P^\beta) = \top \\ \beta \in \text{ladj}(M^\gamma) \end{array}$$

$$\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{RAAdj}} = \frac{\begin{array}{l} [P^\beta, j, k \mid -, - \mid \text{false}] \\ [M^\gamma, i, j \mid p, q \mid \text{false}] \end{array}}{[M^\gamma, i, k \mid p, q \mid \text{true}]} \quad \begin{array}{l} \text{label}(P^\beta) = \top \\ \beta \in \text{ladj}(M^\gamma) \end{array}$$



$$\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{WAdj}} = \frac{\begin{array}{l} [P^\beta, j, m \mid k, l \mid \text{false}] \\ [M^\gamma, k, l \mid p, q \mid \text{false}] \\ [M^\gamma, j, m \mid p, q \mid \text{true}] \end{array}}{\quad} \quad \begin{array}{l} \text{label}(P^\beta) = \top \\ \beta \in \text{wadj}(M^\gamma) \end{array}$$

$$\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{Subs}} = \frac{\begin{array}{l} [P^\alpha, i, j \mid -, - \mid \text{false}] \\ [M^\gamma, i, j \mid -, - \mid \text{false}] \end{array}}{\quad} \quad \begin{array}{l} \text{label}(P^\alpha) = \top \\ \alpha \in \text{subst}(M^\gamma) \end{array}$$

El reconocimiento ascendente se inicia con los pasos deductivos  $\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{Scan}}$  y  $\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{e}}$ .  $\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{Scan}}$  introduce en el análisis los subárboles que dominan directamente nodos etiquetados con símbolos terminales que coincida con algún símbolo de la cadena de entrada. Y  $\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{e}}$  introduce aquellos subárboles que dominan directamente nodos etiquetados con la palabra vacía. Los dos pasos son semejantes a sus homónimos del esquema **CYK**, aunque también se incluye en el paso  $\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{e}}$  los nodos pies de los árboles auxiliares fuertemente izquierdos y derechos, debido a que en estos tipos de árboles auxiliares no es necesario transmitir la información del subárbol escindido.

El paso deductivo  $\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{Foot}}$  establece todos los posibles fragmentos de la cadena de entrada que puede cubrir el nodo pie. Este paso es análogo al del mismo nombre del esquema **CYK**, pero sólo se aplica a los árboles auxiliares wrapping, ya que los nodos pies del resto de árboles auxiliares se encuentran recogidos en el paso  $\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{e}}$ .

Una vez que el reconocimiento de un subárbol dominado por un nodo se ha completado, se debe continuar con el reconocimiento del superárbol respecto a dicho nodo. El paso deductivo  $\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{CompUna}}$  se aplica a nodos que dominan directamente a un solo nodo y el paso  $\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{CompBin}}$  se aplica a los que dominan directamente a dos nodos. Si dicho subárbol domina el nodo pie de un árbol en  $\mathbf{A}_W$ , entonces se transmite la información del segmento que cuelga del mismo en los índices del consecuente. La condición que acompaña a ambos pasos de compleción comprueba que el nodo que domina el subárbol que se va a completar no presente adjunción (izquierda, derecha o wrapping) obligatoria y, en el caso de presentarla, que se haya efectuado. Estos dos pasos son totalmente análogos a sus homónimos en el esquema **CYK**.

Como dijimos anteriormente, la principal diferencia de este esquema con el conocido para TAG consiste en el desglose del paso deductivo de adjunción  $\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{Adj}}$  en tres casos en función del tipo de árbol auxiliar que se va adjuntar:  $\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{LAdj}}$  si es un árbol auxiliar fuertemente izquierdo,  $\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{RAAdj}}$  si es un árbol auxiliar fuertemente derecho y  $\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{WAdj}}$  si es un árbol auxiliar de wrapping. Los dos primeros pasos se caracterizan porque los ítems de los antecedentes que representan el reconocimiento de los árboles auxiliares tienen indefinidos los índices que llevan la información del pie. Estos pasos son análogos a sus homónimos del esquema **CYK<sup>j</sup>**, pero en los antecedentes que representan el reconocimiento del subárbol escindido por la adjunción se incluyen los índices que llevan la información del pie para el caso de adjunciones (izquierdas o derechas) en nodos de la espina de árboles auxiliares wrapping. Como era de suponer, el paso  $\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{WAdj}}$  es igual al paso  $\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{Adj}}$ , ya que lleva a cabo la adjunción de árboles wrapping. Para impedir más de una adjunción sobre el mismo nodo, en los antecedentes que representan el subárbol que cuelga del nodo pie se requiere

que el componente *adj* sea *false* y además el componente *adj* del consecuente se fija a *true*.

El paso de sustitución  $\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{Subs}}$  es análogo al del mismo nombre del esquema  $\text{CYK}^{\text{j}}$ .

El conjunto de ítems finales se define mediante:

$$\mathcal{F}_{\text{CYK}^{\text{Mix}}} = \{[P^\alpha, 0, n \mid -, - \mid \text{false}] \mid \alpha \in I, \text{label}(P^\alpha) = \top, \text{label}(\mathbf{R}^\alpha) = S\}$$

En cuanto a la complejidad teórica, tanto temporal como espacial, en el caso peor permanece en la cota estándar para TAG. Para determinar las mejoras obtenidas mediante esta integración de esquemas se requiere un estudio más exhaustivo. Los únicos árboles elementales que se analizan con un coste espacial de  $\mathcal{O}(n^4)$  en el caso peor son los árboles auxiliares wrapping, mientras la complejidad espacial en el análisis de los árboles elementales del conjunto  $I \cup \mathbf{A}_{SL} \cup \mathbf{A}_{SR}$  se reduce a  $\mathcal{O}(n^2)$ . En cuanto a la complejidad temporal en el caso peor:

- Debido a los pasos  $\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{LAdj}}$  y  $\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{RAAdj}}$ , las adjunciones de árboles auxiliares fuertemente izquierdos o derechos en:
  - un nodo de un árbol en  $I \cup \mathbf{A}_{SL} \cup \mathbf{A}_{SR}$  o en un nodo que no se encuentre en la espina de un árbol en  $\mathbf{A}_W$  presentan una complejidad  $\mathcal{O}(n^3)$ .
  - un nodo de la espina de un árbol auxiliar en  $\mathbf{A}_W$  presentan una complejidad  $\mathcal{O}(n^5)$ .
- Debido al paso  $\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{WAdj}}$ , las adjunciones de árboles auxiliares wrapping en:
  - un nodo de un árbol en  $I \cup \mathbf{A}_{SL} \cup \mathbf{A}_{SR}$  o en un nodo que no se encuentre en la espina de un árbol en  $\mathbf{A}_W$  presentan una complejidad  $\mathcal{O}(n^4)$ .
  - un nodo de la espina de un árbol auxiliar en  $\mathbf{A}_W$  presentan la complejidad máxima del algoritmo de  $\mathcal{O}(n^6)$ .
- Las operaciones de sustitución se realizan en tiempo  $\mathcal{O}(n^2)$ .
- Las operaciones de reconocimiento y compleción permanecen en las mismas cotas de complejidad del esquema  $\text{CYK}$ . Aunque las operaciones de compleción unaria y binaria en nodos pertenecientes a las espinas de árboles auxiliares en  $\mathbf{A}_{SL} \cup \mathbf{A}_{SR}$  se reducen a  $\mathcal{O}(n^2)$  y  $\mathcal{O}(n^3)$ , respectivamente.

### 10.2.1 Relación entre los esquemas tipo CYK para TAG

En esta sección mostraremos que el esquema presentado en esta sección no es más que el resultado de aplicar una contracción sobre los ítems del esquema  $\text{CYK}$ . Aunque inmediatamente presentaremos la demostración formal, de forma intuitiva podemos decir que esta relación se deduce de que podemos considerar que todos los ítems que representan subárboles

dominados por nodos de la espina de árboles del conjunto  $A_{SL} \cup A_{SR}$  en el esquema **CYK** y que llevan información del nodo pie, es decir, que tienen la forma

$$[N^\beta, i, j \mid k, l \mid adj],$$

son representados en el esquema **CYK<sup>Mix</sup>** por ítems de la forma

$$[N^\beta, i, j \mid -, - \mid adj].$$

Para poder establecer dicha relación, vamos a suponer que en la definición original de **CYK** incluimos la operación de sustitución. Es decir, redefinimos el esquema **CYK** añadiéndole al conjunto de pasos deductivos  $\mathcal{D}_{\text{CYK}}$  el paso de sustitución:

$$\mathcal{D}_{\text{CYK}}^{\text{Subs}} = \mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{Subs}}$$

**Teorema 10.1** *Relación entre los esquemas **CYK** y **CYK<sup>Mix</sup>***

*Se mantiene la siguiente relación de contracción de ítems entre esquemas:*

$$\text{CYK} \xrightarrow{\text{ic}} \text{CYK}^{\text{Mix}}$$

*Prueba*

*Tenemos que demostrar que existe una función regular  $f : \mathcal{I}_{\text{CYK}} \rightarrow \mathcal{I}_{\text{CYK}^{\text{Mix}}}$  tal que:*

1.  $\mathcal{I}_{\text{CYK}^{\text{Mix}}} = f(\mathcal{I}_{\text{CYK}})$
2.  $\Delta_{\text{CYK}^{\text{Mix}}} = f(\Delta_{\text{CYK}})$

*Una función regular de contracción de ítems que cumple estas condiciones es la siguiente:*

$$f([N^\beta, i, k \mid j, k \mid adj]) = [N^\beta, i, j \mid -, - \mid adj] \text{ si } \beta \in A_{SL}$$

$$f([N^\beta, i, k \mid i, j \mid adj]) = [N^\beta, j, k \mid -, - \mid adj] \text{ si } \beta \in A_{SR}$$

$$f([N^\gamma, i, k \mid p, q \mid adj]) = [N^\gamma, i, k \mid p, q \mid adj] \text{ para el resto de casos}$$

*De  $f$  se sigue inmediatamente que  $\mathcal{I}_{\text{CYK}^{\text{Mix}}} = f(\mathcal{I}_{\text{CYK}})$  y que  $\Delta_{\text{CYK}^{\text{Mix}}} = f(\Delta_{\text{CYK}})$  por inducción en la longitud de las secuencias derivación. Veamos la función para cada uno de los pasos deductivos:*

- *Dado un paso*

$$\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{e}} = \overline{[N^\gamma, j, j \mid -, - \mid false]} \text{ label}(N^\gamma) = \perp \text{ y } \gamma \in A_{SL} \cup A_{SR}$$

*Si  $\text{label}(N^\gamma) = \perp$  y  $\gamma \in A_{SL}$ , entonces por el caso 1 de la definición de  $f$  se cumple:*

$$[N^\gamma, j, j \mid -, - \mid false] = f([N^\gamma, j, k \mid j, k \mid false])$$

Si  $\text{label}(N^\gamma) = \perp$  y  $\gamma \in \mathbf{A}_{SR}$ , entonces por el caso 2 de la definición de  $f$  se cumple:

$$[N^\gamma, j, j \mid -, - \mid \text{false}] = f([N^\gamma, i, j \mid i, j \mid \text{false}])$$

Para la parte del paso  $\mathcal{D}_{\text{CYK}^{\text{Mix}}}^\varepsilon$  que se ocupa del reconocimiento del nodo pie se cumple:

$$\mathcal{D}_{\text{CYK}^{\text{Mix}}}^\varepsilon = f(\mathcal{D}_{\text{CYK}}^{\text{Foot}})$$

- Dado un paso

$$\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{Foot}} = \frac{\text{label}(P^\beta) = \perp}{[P^\beta, k, l \mid k, l \mid \text{false}]} \quad \beta \in \mathbf{A}_W$$

Por el caso 3 de la definición de  $f$  se cumple:

$$[P^\beta, k, l \mid k, l \mid \text{false}] = f([P^\beta, k, l \mid k, l \mid \text{false}])$$

Por consiguiente:

$$\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{Foot}} = f(\mathcal{D}_{\text{CYK}}^{\text{Foot}})$$

- Dado un paso

$$\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{CompUna}} = \frac{[O^\gamma, i, j \mid p, q \mid \text{adj}]}{[M^\gamma, i, j \mid p, q \mid \text{false}]} \quad M^\gamma \rightarrow O^\gamma \in \mathcal{P}(\gamma)$$

Si  $O^\gamma$  pertenece a la espina de un árbol en  $\mathbf{A}_{SL} \cup \mathbf{A}_{SR}$ , entonces  $M^\gamma$  también pertenecerá y los índices  $p$  y  $q$  estarán indefinidos. En esta situación se pueden dar dos casos:

- si  $\gamma \in \mathbf{A}_{SL}$ , por el caso 1 de la definición de  $f$  se cumple:

$$[O^\gamma, i, j \mid -, - \mid \text{adj}] = f([O^\gamma, i, k \mid j, k \mid \text{adj}])$$

$$[M^\gamma, i, j \mid -, - \mid \text{false}] = f([M^\gamma, i, k \mid j, k \mid \text{false}])$$

- si  $\gamma \in \mathbf{A}_{SR}$ , por el caso 2 de la definición de  $f$  se cumple:

$$[O^\gamma, i, j \mid -, - \mid \text{adj}] = f([O^\gamma, k, j \mid k, i \mid \text{adj}])$$

$$[M^\gamma, i, j \mid -, - \mid \text{false}] = f([M^\gamma, k, j \mid k, i \mid \text{false}])$$

El resto de situaciones son iguales en ambos esquemas y se contemplan en el caso 3 de la definición de  $f$ :

$$[O^\gamma, i, j \mid p, q \mid \text{adj}] = f([O^\gamma, i, j \mid p, q \mid \text{adj}])$$

$$[M^\gamma, i, j \mid p, q \mid \text{false}] = f([M^\gamma, i, j \mid p, q \mid \text{false}])$$

Por consiguiente:

$$\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{CompUna}} = f(\mathcal{D}_{\text{CYK}}^{\text{CompUna}})$$

- Dado un paso

$$\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{CompBin}} = \frac{[O_1^\gamma, i, j \mid p, q \mid \text{adj}] [O_2^\gamma, j, k \mid p', q' \mid \text{adj}']}{[M^\gamma, i, k \mid p \cup p' \mid q \cup q' \mid \text{false}]} \quad M^\gamma \rightarrow O_1^\gamma O_2^\gamma \in \mathcal{P}(\gamma)$$

Si  $M^\gamma$  pertenece a la espina de un árbol en  $\mathbf{A}_{\text{SL}} \cup \mathbf{A}_{\text{SR}}$ , entonces  $O_1^\gamma$  u  $O_2^\gamma$  también pertenecerá y sus índices  $(p, q)$  o  $(p', q')$  estarán indefinidos. Se pueden distinguir tres casos:

- si  $O_2^\gamma$  pertenece a la espina de  $\gamma \in \mathbf{A}_{\text{SL}}$ , por el caso 1 de la definición de  $f$  se cumple:

$$\begin{aligned} [O_2^\gamma, j, k \mid -, - \mid \text{adj}'] &= f([O_2^\gamma, j, l \mid k, l \mid \text{adj}']) \\ [M^\gamma, i, k \mid -, - \mid \text{false}] &= f([M^\gamma, i, l \mid k, l \mid \text{false}]) \end{aligned}$$

y por el caso 3 de la definición de  $f$ :

$$[O_1^\gamma, i, j \mid -, - \mid \text{adj}] = f([O_1^\gamma, i, j \mid -, - \mid \text{adj}])$$

- si  $O_1^\gamma$  pertenece a la espina de  $\gamma \in \mathbf{A}_{\text{SR}}$ , por el caso 2 de la definición de  $f$  se cumple:

$$\begin{aligned} [O_1^\gamma, i, j \mid -, - \mid \text{adj}] &= f([O_1^\gamma, l, j \mid l, i \mid \text{adj}]) \\ [M^\gamma, i, k \mid -, - \mid \text{false}] &= f([M^\gamma, l, k \mid l, i \mid \text{false}]) \end{aligned}$$

y por el caso 3 de la definición de  $f$ :

$$[O_2^\gamma, j, k \mid -, - \mid \text{adj}'] = f([O_2^\gamma, j, k \mid -, - \mid \text{adj}'])$$

- el resto de situaciones son iguales en ambos esquemas y se contemplan en el caso 3 de la definición de  $f$ .

Se concluye que:

$$\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{CompBin}} = f(\mathcal{D}_{\text{CYK}}^{\text{CompBin}})$$

- Dado un paso

$$\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{LAdj}} = \frac{[P^\beta, i, j \mid -, - \mid \text{false}] [M^\gamma, j, k \mid p, q \mid \text{false}]}{[M^\gamma, i, k \mid p, q \mid \text{true}]} \quad \begin{array}{l} \text{label}(P^\beta) = \top \\ \beta \in \text{ladj}(M^\gamma) \end{array}$$

Por el caso 1 de la definición de  $f$  para el primer antecedente se cumple:

$$[P^\beta, i, j \mid -, - \mid \text{false}] = f([P^\beta, i, l \mid j, l \mid \text{false}])$$

Teniendo en cuenta que un árbol auxiliar fuertemente izquierdo no se puede adjuntar en la espina de otro fuertemente derecho, nos encontramos con la siguiente casuística respecto al segundo antecedente y consecuente:

- si  $M^\gamma$  pertenece a la espina de un árbol en  $\mathbf{ASL}$ , por el caso 1 de la definición de  $f$  se cumple:

$$[M^\gamma, j, k \mid -, - \mid false] = f([M^\gamma, j, l \mid k, l \mid false])$$

$$[M^\gamma, i, k \mid -, - \mid true] = f([M^\gamma, i, l \mid k, l \mid true])$$

- El resto de situaciones son iguales en ambos esquemas y se recogen en el caso 3 de la definición de  $f$ :

$$[M^\gamma, j, k \mid p, q \mid false] = f([M^\gamma, i, j \mid p, q \mid false])$$

$$[M^\gamma, i, k \mid p, q \mid true] = f([M^\gamma, i, j \mid p, q \mid true])$$

De todo esto podemos concluir que:

$$\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{LAdj}} = f(\mathcal{D}_{\text{CYK}}^{\text{Adj}})$$

- Dado un paso

$$\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{RAAdj}} = \frac{\begin{array}{l} [P^\beta, j, k \mid -, - \mid false] \\ [M^\gamma, i, j \mid p, q \mid false] \end{array}}{[M^\gamma, i, k \mid p, q \mid true]} \quad \begin{array}{l} \text{label}(P^\beta) = \top \\ \beta \in \text{radj}(M^\gamma) \end{array}$$

Por el caso 2 de la definición de  $f$  para el primer antecedente se cumple:

$$[P^\beta, j, k \mid -, - \mid false] = f([P^\beta, l, k \mid l, j \mid false])$$

Teniendo en cuenta que un árbol auxiliar fuertemente derecho no se puede adjuntar en la espina de otro fuertemente izquierdo, nos encontramos con la siguiente casuística respecto al segundo antecedente y consecuente:

- si  $M^\gamma$  pertenece a la espina de un árbol en  $\mathbf{ASR}$ , por el caso 2 de la definición de  $f$  se cumple:

$$[M^\gamma, i, j \mid -, - \mid false] = f([M^\gamma, l, j \mid l, i \mid false])$$

$$[M^\gamma, i, k \mid -, - \mid true] = f([M^\gamma, l, k \mid l, i \mid true])$$

- El resto de situaciones son iguales en ambos esquemas y se recogen en el caso 3 de la definición de  $f$ :

$$[M^\gamma, i, j \mid p, q \mid false] = f([M^\gamma, i, j \mid p, q \mid false])$$

$$[M^\gamma, i, k \mid p, q \mid true] = f([M^\gamma, i, k \mid p, q \mid true])$$

Por consiguiente:

$$\mathcal{D}_{\text{CYK}^{\text{Mix}}}^{\text{RAAdj}} = f(\mathcal{D}_{\text{CYK}}^{\text{Adj}})$$

- Los pasos de reconocimiento y sustitución son iguales en **CYK** y **CYK<sup>Mix</sup>**.

**Teorema 10.2** Corrección del analizador **CYK<sup>Mix</sup>**

Para cualquier gramática  $G \in TAG$  y una cadena de entrada  $a_1 \dots a_n$  se cumple que:

$$[P^\alpha, 0, n \mid -, - \mid false] \in \mathcal{V}_{\text{CYK}^{\text{Mix}}} \text{ si y solo si } \alpha \in I, \text{label}(P^\alpha) = \top,$$

$$\text{label}(\mathbf{R}^\alpha) = S, \top \Rightarrow \mathbf{R}^\alpha \xrightarrow{*} a_1 \dots a_n$$

*Prueba*

Se deduce de la relación que mantiene con el esquema **CYK**, y que es la siguiente:

$$\text{CYK}^{\text{Mix}} \xrightarrow{\text{ir}} \text{CYK}$$

Por el corolario 2.1, la corrección de **CYK** implica la corrección de **CYK<sup>Mix</sup>**.

### 10.3 Esquema tipo Earley ascendente

Ahora definiremos una versión ascendente del analizador tipo Earley generado a partir de la fusión de los esquemas **buE** y **buE<sup>j</sup>**.

El dominio del nuevo esquema **buE<sup>Mix</sup>** también es el resultado de la integración de los dominios de ambos esquemas. Se obtiene tomando como base el dominio del **buE<sup>j</sup>** y añadiéndole los dos índices correspondientes a la información del pie para árboles auxiliares wrapping. Al introducir la adjunción wrapping, el parámetro *adj* en los ítems también se redefine para que incluya dicho tipo de adjunción. El dominio del esquema **buE<sup>Mix</sup>** es:

$$\mathcal{I}_{\text{buE}^{\text{Mix}}} = \mathcal{I}_{\text{buE}^{\text{Mix}}}^{(i)} \cup \mathcal{I}_{\text{buE}^{\text{Mix}}}^{(ii)}$$

$$\mathcal{I}_{\text{buE}^{\text{Mix}}}^{(i)} = \{[M^\gamma \rightarrow \delta \bullet \nu, i, j \mid p, q \mid false]\}$$

donde  $N^\gamma \rightarrow \delta \nu \in \mathcal{P}(\gamma)$ , siendo  $\gamma \in I \cup A$  y  $\nu \neq \varepsilon$ . Los índices  $0 \leq i \leq j$  establecen las posiciones dentro de la cadena de entrada que delimitan el fragmento reconocido por  $\delta$ . Si  $p$  y  $q$  presentan un valor definido, entonces  $M^\gamma$  es un nodo de la espina de un árbol auxiliar wrapping ( $\gamma \in A_W$ ) y se cumple  $(p, q) \leq (i, j)$ . El último parámetro *false* indica que no se ha completado ninguna adjunción sobre  $M^\gamma$ , ya que aquí hemos optado porque las adjunciones sobre un nodo se completen únicamente cuando se hayan reconocido en su totalidad tanto el árbol auxiliar como el subárbol que domina dicho nodo.

$$\mathcal{I}_{\text{buE}^{\text{Mix}}}^{(ii)} = \{[M^\gamma \rightarrow \nu \bullet, i, j \mid p, q \mid adj]\}$$

donde  $N^\gamma \rightarrow \nu \in \mathcal{P}(\gamma)$ , siendo  $\gamma \in I \cup A$ . Los índices  $0 \leq i \leq j$  establecen las posiciones dentro de la cadena de entrada que delimitan el fragmento reconocido por  $\nu$ . Si  $p$  y  $q$  presentan un valor definido, entonces  $M^\gamma$  es un nodo de la espina de un árbol auxiliar

wrapping ( $\gamma \in \mathbf{A}_W$ ) y se cumple  $(p, q) \leq (i, j)$ . El parámetro booleano  $adj \in \{true, false\}$  indica si se ha completado alguna adjunción (izquierda, derecha o wrapping) sobre  $M^\gamma$ .

El conjunto de pasos deductivos del esquema se define como:

$$\begin{aligned} \mathcal{D}_{\text{buE}} &= \mathcal{D}_{\text{buEMix}}^{\text{Ini}} \cup \mathcal{D}_{\text{buEMix}}^{\text{Foot}} \cup \mathcal{D}_{\text{buEMix}}^{\text{Scan}} \cup \mathcal{D}_{\text{buEMix}}^\varepsilon \cup \mathcal{D}_{\text{buEMix}}^{\text{Comp}} \cup \\ &\quad \mathcal{D}_{\text{buEMix}}^{\text{LAdj}} \cup \mathcal{D}_{\text{buEMix}}^{\text{RAAdj}} \cup \mathcal{D}_{\text{buEMix}}^{\text{WAdj}} \cup \mathcal{D}_{\text{buEMix}}^{\text{Subs}} \\ \mathcal{D}_{\text{buEMix}}^{\text{Ini}} &= \overline{[N^\gamma \rightarrow \bullet \delta, i, i \mid -, - \mid false]} \\ \mathcal{D}_{\text{buEMix}}^{\text{Foot}} &= \overline{[\mathbf{F}^\beta \rightarrow \perp \bullet, k, l \mid k, l \mid false]} \quad \beta \in \mathbf{A}_W \\ &\quad [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q \mid false], \\ &\quad [a, j, j + 1] \\ \mathcal{D}_{\text{buEMix}}^{\text{Scan}} &= \overline{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, j + 1 \mid p, q \mid false]} \quad \text{label}(M^\gamma) = a \\ \mathcal{D}_{\text{buEMix}}^\varepsilon &= \overline{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q \mid false]} \quad \text{label}(M^\gamma) = \varepsilon \\ &\quad \circ (\text{label}(M^\gamma) = \perp \text{ y } \gamma \in \mathbf{A}_{SL} \cup \mathbf{A}_{SR}) \\ \mathcal{D}_{\text{buEMix}}^{\text{Comp}} &= \overline{[M^\gamma \rightarrow \delta \bullet, j, k \mid p, q \mid adj] \\ &\quad [N^\gamma \rightarrow \nu \bullet M^\gamma \omega, i, j \mid p', q' \mid false]} \\ &\quad [N^\gamma \rightarrow \nu M^\gamma \bullet \omega, i, k \mid p \cup p', q \cup q' \mid false]} \end{aligned}$$

donde se debe cumplir que  $(\text{nil} \in \text{adj}(M^\gamma) \text{ y } adj = false)$  o  $(\exists \beta \in \text{adj}(M^\gamma) \text{ y } adj \neq false)$ .

$$\begin{aligned} \mathcal{D}_{\text{buEMix}}^{\text{LAdj}} &= \frac{[\top \rightarrow \mathbf{R}^\beta \bullet, i, j \mid -, - \mid false] \\ [M^\gamma \rightarrow \nu \bullet, j, k \mid p, q \mid false]}{[M^\gamma \rightarrow \nu \bullet, i, k \mid p, q \mid true]} \quad \beta \in \text{ladj}(M^\gamma) \\ \mathcal{D}_{\text{buEMix}}^{\text{RAAdj}} &= \frac{[\top \rightarrow \mathbf{R}^\beta \bullet, j, k \mid -, - \mid false] \\ [M^\gamma \rightarrow \nu \bullet, i, j \mid p, q \mid false]}{[M^\gamma \rightarrow \nu \bullet, i, k \mid p, q \mid true]} \quad \beta \in \text{radj}(M^\gamma) \\ \mathcal{D}_{\text{buEMix}}^{\text{WAdj}} &= \frac{[\top \rightarrow \mathbf{R}^\beta \bullet, j, m \mid k, l \mid false] \\ [M^\gamma \rightarrow \nu \bullet, k, l \mid p, q \mid false]}{[M^\gamma \rightarrow \nu \bullet, j, m \mid p, q \mid true]} \quad \beta \in \text{wadj}(M^\gamma) \\ \mathcal{D}_{\text{buEMix}}^{\text{Subs}} &= \frac{[\top \rightarrow \mathbf{R}^\alpha \bullet, j, k \mid -, - \mid false] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q \mid false]}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k \mid p, q \mid false]} \quad \alpha \in \text{subst}(M^\gamma) \end{aligned}$$

Los pasos deductivos que inician el reconocimiento ascendente son  $\mathcal{D}_{\text{buEMix}}^{\text{Ini}}$  y  $\mathcal{D}_{\text{buEMix}}^{\text{Foot}}$ . El paso  $\mathcal{D}_{\text{buEMix}}^{\text{Ini}}$  establece de partida la predicción de todos los subárboles participantes en los árboles elementales y el paso  $\mathcal{D}_{\text{buEMix}}^{\text{Foot}}$  completa el reconocimiento de los subárboles que dominan los nodos pie de los árboles auxiliares en  $\mathbf{A}_W$  entre todas las posibles posiciones de la cadena de entrada. Estos dos pasos son análogos a sus homónimos en el esquema **buE**, pero el último solo se aplica a los árboles auxiliares en  $\mathbf{A}_W$ .



Los pasos de reconocimiento  $\mathcal{D}_{\text{buEMix}}^{\text{Scan}}$  y  $\mathcal{D}_{\text{buEMix}}^{\varepsilon}$  se aplican cuando el reconocimiento alcanza un nodo etiquetado con un símbolo terminal, la cadena vacía o el nodo pie de los árboles auxiliares en  $\mathbf{A}_{SL} \cup \mathbf{A}_{SR}$ . Estos dos pasos son análogos a sus homónimos en el esquema **buE** pero, por los mismos motivos que en el esquema **CYK**, en el paso  $\mathcal{D}_{\text{buEMix}}^{\varepsilon}$  también se incluyen los nodos pie de los árboles auxiliares en  $\mathbf{A}_{SL} \cup \mathbf{A}_{SR}$ .

El paso deductivo  $\mathcal{D}_{\text{buEMix}}^{\text{Comp}}$  continúa el reconocimiento del superárbol respecto a un nodo una vez que el subárbol dominado por él ha sido completamente reconocido. Si dicho subárbol domina el nodo pie de un árbol en  $\mathbf{A}_W$ , entonces el antecedente que representa el reconocimiento del subárbol contiene en los índices  $p$  y  $q$  la información del segmento que pende del pie. La condición que acompaña a este paso garantiza que el nodo que domina el subárbol que se va a completar no presente adjunción (izquierda, derecha o wrapping) obligatoria y, en el caso de presentarla, que se haya efectuado.

Como es característico en todos los analizadores de este capítulo, la operación de adjunción distingue entre los tres tipos de árboles auxiliares, empleándose:  $\mathcal{D}_{\text{buEMix}}^{\text{LAdj}}$  para adjunciones de árboles auxiliares izquierdos,  $\mathcal{D}_{\text{buEMix}}^{\text{RAdj}}$  para adjunciones de árboles auxiliares derechos y  $\mathcal{D}_{\text{buEMix}}^{\text{WAdj}}$  para adjunciones de árboles auxiliares wrapping. Para evitar la posibilidad de múltiples operaciones de adjunción sobre un nodo, el componente *adj* se activa a *true* en el consecuente en todos los pasos de adjunción y se requiere que este componente esté a *false* en los antecedentes. A diferencia del esquema **buE**, estos pasos no completan la operación de adjunción, lo que hacen es introducir en el análisis ítems con la información del árbol adjuntado y subárbol escindido. Una posterior operación de completación  $\mathcal{D}_{\text{buEMix}}^{\text{Comp}}$  finalizará la adjunción.

El conjunto de ítems finales viene dado por:

$$\mathcal{F}_{\text{buEMix}} = \{[\top \rightarrow \mathbf{R}^\alpha \bullet, 0, n \mid -, - \mid \text{false}] \mid \alpha \in I, \text{label}(\mathbf{R}^\alpha) = S\}$$

Estudiamos ahora su complejidad teórica. Los árboles en  $\mathbf{A}_W$  se analizan con un coste espacial de  $\mathcal{O}(n^4)$  en el caso peor. Y en el resto de árboles elementales, es decir, los pertenecientes a  $I \cup \mathbf{A}_{SL} \cup \mathbf{A}_{SR}$ , se reduce a  $\mathcal{O}(n^2)$ . La complejidad temporal en el caso peor de:

- Las adjunciones de árboles auxiliares fuertemente izquierdos o derechos en:
  - un nodo de un árbol en  $I \cup \mathbf{A}_{SL} \cup \mathbf{A}_{SR}$  presentan una complejidad  $\mathcal{O}(n^3)$ , debido a los pasos  $\mathcal{D}_{\text{buEMix}}^{\text{LAdj}}$ ,  $\mathcal{D}_{\text{buEMix}}^{\text{RAdj}}$  y  $\mathcal{D}_{\text{buEMix}}^{\text{Comp}}$ .
  - un nodo que se encuentre en la espina de un árbol en  $\mathbf{A}_W$  presentan una complejidad  $\mathcal{O}(n^5)$ , debido a los pasos  $\mathcal{D}_{\text{buEMix}}^{\text{LAdj}}$ ,  $\mathcal{D}_{\text{buEMix}}^{\text{RAdj}}$  y  $\mathcal{D}_{\text{buEMix}}^{\text{Comp}}$ .
  - un nodo dominado directamente por un nodo de la espina de un árbol en  $\mathbf{A}_W$  y que se encuentra en el contexto derecho de la espina de dicho árbol presentan una complejidad  $\mathcal{O}(n^5)$ , debido al paso  $\mathcal{D}_{\text{buEMix}}^{\text{Comp}}$ .

- un nodo que encuentre en el contexto izquierdo de la espina de un árbol en  $\mathbf{A}_W$  o en el contexto derecho de un árbol en  $\mathbf{A}_W$  de la espina pero no está dominado directamente por ningún nodo perteneciente a la espina presentan una complejidad  $\mathcal{O}(n^3)$ , debido a los pasos  $\mathcal{D}_{\text{buEMix}}^{\text{LAdj}}$ ,  $\mathcal{D}_{\text{buEMix}}^{\text{RAAdj}}$  y  $\mathcal{D}_{\text{buEMix}}^{\text{Comp}}$ .
- Las adjunciones de árboles auxiliares wrapping en:
  - un nodo de un árbol en  $I \cup \mathbf{A}_{SL} \cup \mathbf{A}_{SR}$  presentan una complejidad  $\mathcal{O}(n^4)$ , debido al paso  $\mathcal{D}_{\text{buEMix}}^{\text{WAdj}}$ .
  - un nodo de la espina de un árbol auxiliar en  $\mathbf{A}_W$  presentan la complejidad máxima del algoritmo de  $\mathcal{O}(n^6)$ , debido al paso  $\mathcal{D}_{\text{buEMix}}^{\text{WAdj}}$ .
  - un nodo dominado directamente por un nodo de la espina de un árbol en  $\mathbf{A}_W$  y que se encuentra en el contexto derecho de la espina de dicho árbol presentan una complejidad  $\mathcal{O}(n^5)$ , debido al paso  $\mathcal{D}_{\text{buEMix}}^{\text{Comp}}$ .
  - un nodo que encuentre en el contexto izquierdo de la espina de un árbol en  $\mathbf{A}_W$  o en el contexto derecho de la espina de un árbol en  $\mathbf{A}_W$  pero no está dominado directamente por ningún nodo perteneciente a la espina presentan una complejidad  $\mathcal{O}(n^4)$ , debido al paso  $\mathcal{D}_{\text{buEMix}}^{\text{WAdj}}$ .
- Las operaciones de sustitución se realizan en tiempo  $\mathcal{O}(n^3)$ , salvo las sustituciones en los nodos dominados directamente por un nodo de la espina de un árbol en  $\mathbf{A}_W$  y que se encuentran en el contexto derecho de la espina de dicho árbol, las cuales presentan una complejidad  $\mathcal{O}(n^5)$ .
- Las operaciones de reconocimiento y compleción permanecen en las mismas cotas de complejidad del esquema **buE**. Aunque las operaciones de compleción en nodos pertenecientes a las espinas de árboles auxiliares en  $\mathbf{A}_{SL} \cup \mathbf{A}_{SR}$  se reducen a  $\mathcal{O}(n^3)$ .

Este analizador mejora la complejidad teórica media del analizador que vimos para TAG y, por tanto, su comportamiento práctico en muchas situaciones. Sin embargo, adolece de los mismos problemas que sus esquemas de origen **buE** y **buE<sup>j</sup>**, y que en su mayor parte son derivados del hecho de que introduce muchos ítems innecesarios en el proceso de análisis, debido fundamentalmente a que lanza el reconocimiento ascendente desde todos los subárboles en los árboles elementales. Así, del mismo modo que hicimos en los analizadores de este tipo para TIG y TAG, también podríamos aplicar al esquema **buE<sup>Mix</sup>** un filtro basado en la esquina izquierda de las producciones y obtener un esquema que mejore sus prestaciones.

## 10.4 Esquema tipo de Vreught y Honig

En esta sección definiremos un analizador para TAG con estrategia ascendente y recorrido bidireccional de la cadena de entrada, obtenido mediante la integración del esquema **dVH**

para TAG con el esquema  $\mathbf{dVH}^j$  para TIG.

Si examinamos el esquema  $\mathbf{dVH}^j$  del apéndice B podremos comprobar que se obtiene aplicándole las convenciones y notación introducidas en este capítulo al esquema  $\mathbf{dVH}^i$  que presentamos en el tema 8. Esto implica que, al igual que en los dos esquemas ascendentes anteriores, la adjunción de un árbol auxiliar sobre un nodo no se puede llevar a cabo hasta que no se complete el reconocimiento del subárbol que pende del nodo pie. Esta consideración hace que el dominio del nuevo esquema se pueda considerar como el resultado de la adición de los dos índices correspondientes a la información del pie y la redefinición del parámetro  $adj$  del esquema  $\mathbf{dVH}^j$  para adaptarlo a las posibles adjunciones wrapping. El conjunto de ítems del nuevo esquema  $\mathbf{dVH}^{\text{Mix}}$  es:

$$\mathcal{I}_{\mathbf{dVH}^{\text{Mix}}} = \mathcal{I}_{\mathbf{dVH}^{\text{Mix}}}^{(i)} \cup \mathcal{I}_{\mathbf{dVH}^{\text{Mix}}}^{(ii)}$$

$$\mathcal{I}_{\mathbf{dVH}^{\text{Mix}}}^{(i)} = \{[M^\gamma \rightarrow \nu \bullet \delta \bullet \omega, i, j \mid p, q \mid false]\}$$

donde  $M^\gamma \rightarrow \nu \delta \omega \in \mathcal{P}(\gamma)$ ,  $\gamma \in I \cup A$ ,  $\nu \neq \varepsilon$  o  $\omega \neq \varepsilon$ . Los índices  $0 \leq i \leq j$  establecen las posiciones dentro de la cadena de entrada que delimitan el fragmento reconocido por  $\delta$ . Si  $p$  y  $q$  presentan un valor definido, entonces  $M^\gamma$  es un nodo de la espina de un árbol auxiliar wrapping ( $\gamma \in A_W$ ) y se cumple  $(p, q) \leq (i, j)$ . El último parámetro  $false$  indica que no se ha completado ninguna adjunción sobre  $M^\gamma$ .

$$\mathcal{I}_{\mathbf{dVH}^{\text{Mix}}}^{(ii)} = \{[M^\gamma \rightarrow \bullet \delta \bullet, i, j \mid p, q \mid adj]\}$$

donde  $M^\gamma \rightarrow \delta \in \mathcal{P}(\gamma)$ , siendo  $\gamma \in I \cup A$ . Los índices  $0 \leq i \leq j$  establecen las posiciones dentro de la cadena de entrada que delimitan el fragmento reconocido por  $\delta$ . Si  $p$  y  $q$  presentan un valor definido, entonces  $M^\gamma$  es un nodo de la espina de un árbol auxiliar wrapping ( $\gamma \in A_W$ ) y se cumple  $(p, q) \leq (i, j)$ . El parámetro booleano  $adj \in \{true, false\}$  indica si se ha completado alguna adjunción (izquierda, derecha o wrapping) sobre  $M^\gamma$ .

El conjunto de pasos deductivos viene dado por:

$$\mathcal{D}_{\mathbf{dVH}^{\text{Mix}}} = \mathcal{D}_{\mathbf{dVH}^{\text{Mix}}}^{\text{Scan}} \cup \mathcal{D}_{\mathbf{dVH}^{\text{Mix}}}^\varepsilon \cup \mathcal{D}_{\mathbf{dVH}^{\text{Mix}}}^{\text{Comp}} \cup \mathcal{D}_{\mathbf{dVH}^{\text{Mix}}}^{\text{Con}} \cup \mathcal{D}_{\mathbf{dVH}^{\text{Mix}}}^{\text{Foot}} \cup \mathcal{D}_{\mathbf{dVH}^{\text{Mix}}}^{\text{LAdj}} \cup \mathcal{D}_{\mathbf{dVH}^{\text{Mix}}}^{\text{RAAdj}} \cup \mathcal{D}_{\mathbf{dVH}^{\text{Mix}}}^{\text{WAdj}} \cup \mathcal{D}_{\mathbf{dVH}^{\text{Mix}}}^{\text{Subs}}$$

$$\mathcal{D}_{\mathbf{dVH}^{\text{Mix}}}^{\text{Scan}} = \frac{[a, j, j+1]}{[N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, j, j+1 \mid -, - \mid false]} \quad label(M^\gamma) = a$$

$$\mathcal{D}_{\mathbf{dVH}^{\text{Mix}}}^\varepsilon = \frac{}{[N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, j, j \mid -, - \mid false]} \quad \begin{array}{l} label(M^\gamma) = \varepsilon \\ \text{ó } (label(M^\gamma) = \perp \text{ y } \gamma \in A_{SL} \cup A_{SR}) \end{array}$$

$$\mathcal{D}_{\mathbf{dVH}^{\text{Mix}}}^{\text{Comp}} = \frac{[M^\gamma \rightarrow \bullet \delta \bullet, i, j \mid p, q \mid adj]}{[N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, i, j \mid p, q \mid false]}$$

donde se debe cumplir que  $(\mathbf{nil} \in \text{adj}(M^\gamma)$  y  $\text{adj} = \text{false}$ ) o  $(\exists \beta \in \text{adj}(M^\gamma)$  y  $\text{adj} \neq \text{false}$ ).

$$\mathcal{D}_{\text{dVHMix}}^{\text{Con}} = \frac{\begin{array}{l} [N^\gamma \rightarrow \nu \bullet \delta \bullet \delta' \omega, i, j' \mid p, q \mid \text{false}] \\ [N^\gamma \rightarrow \nu \delta \bullet \delta' \bullet \omega, j', j \mid p', q' \mid \text{false}] \end{array}}{[N^\gamma \rightarrow \nu \bullet \delta \delta' \bullet \omega, i, j \mid p \cup p', q \cup q' \mid \text{false}]}$$

$$\mathcal{D}_{\text{dVHMix}}^{\text{Foot}} = \frac{[\mathbf{F}^\beta \rightarrow \bullet \perp \bullet, k, l \mid k, l \mid \text{false}]}{\beta \in \mathbf{A}_W}$$

$$\mathcal{D}_{\text{dVHMix}}^{\text{LAdj}} = \frac{\begin{array}{l} [\top \rightarrow \bullet \mathbf{R}^\beta \bullet, i, j \mid -, - \mid \text{false}] \\ [M^\gamma \rightarrow \bullet \delta \bullet, j, k \mid p, q \mid \text{false}] \end{array}}{[M^\gamma \rightarrow \bullet \delta \bullet, i, k \mid p, q \mid \text{true}]} \quad \beta \in \text{ladj}(M^\gamma)$$

$$\mathcal{D}_{\text{dVHMix}}^{\text{RAAdj}} = \frac{\begin{array}{l} [\top \rightarrow \bullet \mathbf{R}^\beta \bullet, j, k \mid -, - \mid \text{false}] \\ [M^\gamma \rightarrow \bullet \delta \bullet, i, j \mid p, q \mid \text{false}] \end{array}}{[M^\gamma \rightarrow \bullet \delta \bullet, i, k \mid p, q \mid \text{true}]} \quad \beta \in \text{radj}(M^\gamma)$$

$$\mathcal{D}_{\text{dVHMix}}^{\text{WAdj}} = \frac{\begin{array}{l} [\top \rightarrow \bullet \mathbf{R}^\beta \bullet, j, m \mid k, l \mid \text{false}] \\ [M^\gamma \rightarrow \bullet \delta \bullet, k, l \mid p, q \mid \text{false}] \end{array}}{[M^\gamma \rightarrow \bullet \delta \bullet, j, m \mid p, q \mid \text{true}]} \quad \beta \in \text{wadj}(M^\gamma)$$

$$\mathcal{D}_{\text{dVHMix}}^{\text{Subs}} = \frac{[\top \rightarrow \bullet \mathbf{R}^\alpha \bullet, j, k \mid -, - \mid \text{false}]}{[N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, j, k \mid -, - \mid \text{false}]} \quad \alpha \in \text{subst}(M^\gamma)$$

Los pasos deductivos que inician el reconocimiento ascendente son  $\mathcal{D}_{\text{dVHMix}}^{\text{Scan}}$ ,  $\mathcal{D}_{\text{dVHMix}}^\varepsilon$  y  $\mathcal{D}_{\text{dVHMix}}^{\text{Foot}}$ . Los pasos  $\mathcal{D}_{\text{dVHMix}}^{\text{Scan}}$  y  $\mathcal{D}_{\text{dVHMix}}^\varepsilon$  inician el reconocimiento desde los nodos etiquetados con símbolos terminales, la palabra vacía o el nodo pie de los árboles auxiliares en  $\mathbf{A}_{SL} \cup \mathbf{A}_{SR}$ . El paso  $\mathcal{D}_{\text{dVHMix}}^{\text{Foot}}$  completa el reconocimiento de los subárboles que dominan los nodos pie de los árboles auxiliares en  $\mathbf{A}_W$  entre todas las posibles posiciones de la cadena de entrada.

Una vez completado el reconocimiento de un subárbol el paso deductivo  $\mathcal{D}_{\text{dVHMix}}^{\text{Comp}}$  continúa el reconocimiento del superárbol. Si dicho subárbol domina el pie de un árbol en  $\mathbf{A}_W$ , entonces se transmite la información del segmento que cuelga del mismo en los índices  $p$  y  $q$ . Al igual que en los anteriores analizadores, se acompaña una condición lateral a este paso para garantizar que el nodo que domina el subárbol no presente adjunción (izquierda, derecha o wrapping) obligatoria y, en el caso de presentarla, que se haya efectuado.

Cuando el reconocimiento completa dos fragmentos adyacentes en una regla que dominan segmentos colindantes en la cadena de entrada, el analizador los concatena mediante el paso  $\mathcal{D}_{\text{dVHMix}}^{\text{Con}}$ . Si alguno de los dos fragmentos domina el nodo pie de un árbol auxiliar en  $\mathbf{A}_W$ , entonces el antecedente que representa el reconocimiento de dicho fragmento tendrá en sus índices la información del segmento de la cadena de entrada que cuelga del nodo pie.

De nuevo los pasos de adjunción se agrupan en tres tipos:  $\mathcal{D}_{dVH^{Mix}}^{LAdj}$  para adjunciones de árboles auxiliares izquierdos,  $\mathcal{D}_{dVH^{Mix}}^{RAdj}$  para adjunciones de árboles auxiliares derechos y  $\mathcal{D}_{dVH^{Mix}}^{WAdj}$  para adjunciones de árboles auxiliares wrapping. Los cuales funcionan igual que sus homónimos en el esquema  $\mathbf{buE}^{Mix}$ .

El conjunto de ítems finales se define mediante:

$$\mathcal{F}_{dVH^{Mix}} = \{[\top \rightarrow \bullet \mathbf{R}^\alpha \bullet, 0, n \mid -, - \mid false] \mid \alpha \in I, label(\mathbf{R}^\alpha) = S\}$$

Veamos ahora las mejoras que aporta este analizador respecto complejidad teórica media. Como en los anteriores esquemas, los árboles en  $\mathbf{AW}$  se analizan con un coste espacial de  $\mathcal{O}(n^4)$  en el caso peor y se reduce a  $\mathcal{O}(n^2)$  en el resto. La complejidad temporal en el caso peor de:

- Las adjunciones de árboles auxiliares fuertemente izquierdos o derechos en:
  - un nodo de un árbol en  $I \cup \mathbf{ASL} \cup \mathbf{ASR}$  o un nodo que no está en la espina de un árbol en  $\mathbf{AW}$  presentan una complejidad  $\mathcal{O}(n^3)$ , debido a los pasos  $\mathcal{D}_{dVH^{Mix}}^{LAdj}$  y  $\mathcal{D}_{dVH^{Mix}}^{RAdj}$ .
  - un nodo que se encuentre en la espina de un árbol en  $\mathbf{AW}$  presentan una complejidad  $\mathcal{O}(n^5)$ , debido a los pasos  $\mathcal{D}_{dVH^{Mix}}^{LAdj}$  y  $\mathcal{D}_{dVH^{Mix}}^{RAdj}$ .
- Las adjunciones de árboles auxiliares wrapping en:
  - un nodo de un árbol en  $I \cup \mathbf{ASL} \cup \mathbf{ASR}$  o un nodo que no se encuentre en la espina de un árbol auxiliar en  $\mathbf{AW}$  presentan una complejidad  $\mathcal{O}(n^4)$ , debido al paso  $\mathcal{D}_{dVH^{Mix}}^{WAdj}$ .
  - un nodo de la espina de un árbol auxiliar en  $\mathbf{AW}$  presentan la complejidad máxima del algoritmo de  $\mathcal{O}(n^6)$ , debido al paso  $\mathcal{D}_{dVH^{Mix}}^{WAdj}$ .
- Las operaciones de sustitución se realizan en tiempo  $\mathcal{O}(n^2)$ .
- Las operaciones de reconocimiento, completión y concatenación permanecen en las mismas cotas de complejidad del esquema  $\mathbf{dVH}$ . Aunque las operaciones de completión en nodos pertenecientes a las espinas de árboles auxiliares en  $\mathbf{ASL} \cup \mathbf{ASR}$  se reducen a  $\mathcal{O}(n^2)$  y las operaciones de concatenación en árboles auxiliares en  $\mathbf{ASL} \cup \mathbf{ASR}$  permanecen en todos los casos a casos en  $\mathcal{O}(n^3)$ .

## 10.5 Esquema tipo Earley

Para finalizar definimos un esquema ascendente con información predictiva que se obtiene mediante la fusión de los esquemas  $\mathbf{E}$  y  $\mathbf{E}^j$ . Al nuevo esquema lo denominamos  $\mathbf{E}^{Mix}$  y fue presentado en [9, 10].

El conjunto de ítems del esquema es:

$$\mathcal{I}_{E^{Mix}} = \mathcal{I}_{E^{Mix}}^{(i)} \cup \mathcal{I}_{E^{Mix}}^{(ii)}$$

$$\mathcal{I}_{E^{Mix}}^{(ii)} = \mathcal{I}_{buE^{Mix}}^{(ii)}$$

$$\mathcal{I}_{E^{Mix}}^{(i)} = \{[M^\gamma \rightarrow \delta \bullet \nu, i, j \mid p, q \mid adj]\}$$

donde  $N^\gamma \rightarrow \delta \nu \in \mathcal{P}(\gamma)$ , siendo  $\gamma \in I \cup A$  y  $\nu \neq \varepsilon$ . Los índices  $0 \leq i \leq j$  establecen las posiciones dentro de la cadena de entrada que delimitan el fragmento reconocido por  $\delta$ . Si  $p$  y  $q$  presentan un valor definido, entonces  $M^\gamma$  es un nodo de la espina de un árbol auxiliar wrapping ( $\gamma \in A_W$ ) y se cumple  $(p, q) \leq (i, j)$ . El parámetro  $adj \in \{true, false\}$  indica si se ha completado una adjunción izquierda en  $M^\gamma$ .

El conjunto de pasos deductivos es:

$$\begin{aligned} \mathcal{D}_{E^{Mix}} = & \mathcal{D}_{E^{Mix}}^{Ini} \cup \mathcal{D}_{E^{Mix}}^{Scan} \cup \mathcal{D}_{E^{Mix}}^\varepsilon \cup \mathcal{D}_{E^{Mix}}^{Pred} \cup \mathcal{D}_{E^{Mix}}^{Comp} \cup \\ & \mathcal{D}_{E^{Mix}}^{SubsPred} \cup \mathcal{D}_{E^{Mix}}^{SubsComp} \cup \mathcal{D}_{E^{Mix}}^{WAdjPred} \cup \mathcal{D}_{E^{Mix}}^{FootPred} \cup \mathcal{D}_{E^{Mix}}^{FootComp} \cup \\ & \mathcal{D}_{E^{Mix}}^{WAdjComp} \cup \mathcal{D}_{E^{Mix}}^{LAdjPred} \cup \mathcal{D}_{E^{Mix}}^{LAdjComp} \cup \mathcal{D}_{E^{Mix}}^{RAdjPred} \cup \mathcal{D}_{E^{Mix}}^{RAdjComp} \end{aligned}$$

$$\mathcal{D}_{E^{Mix}}^{WAdjComp} = \mathcal{D}_{buE^{Mix}}^{WAdj}$$

$$\mathcal{D}_{E^{Mix}}^{RAdjComp} = \mathcal{D}_{buE^{Mix}}^{RAdj}$$

$$\mathcal{D}_{E^{Mix}}^{Ini} = \frac{}{[\top \rightarrow \bullet R^\alpha, 0, 0 \mid -, - \mid false]} \quad \begin{array}{l} \alpha \in I \\ label(R^\alpha) = S \end{array}$$

$$\mathcal{D}_{E^{Mix}}^{Scan} = \frac{\begin{array}{l} [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q \mid adj], \\ [a, j, j+1] \end{array}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, j+1 \mid p, q \mid adj]} \quad label(M^\gamma) = S$$

$$\mathcal{D}_{E^{Mix}}^\varepsilon = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q \mid adj]}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, j \mid p, q \mid adj]} \quad \begin{array}{l} label(M^\gamma) = \varepsilon \\ \text{o } (label(M^\gamma) = \perp \text{ y } \gamma \in A_{SL} \cup A_{SR}) \end{array}$$

$$\mathcal{D}_{E^{Mix}}^{Pred} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q \mid adj]}{[M^\gamma \rightarrow \bullet \nu, j, j \mid -, - \mid false]}$$

donde se debe cumplir que  $nil \in adj(M^\gamma)$  o  $(\exists \beta \in A_{SL} \cup A_{SR} \text{ con } \beta \in adj(M^\gamma))$

$$\mathcal{D}_{E^{Mix}}^{Comp} = \frac{\begin{array}{l} [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q \mid adj], \\ [M^\gamma \rightarrow \nu \bullet, j, k \mid p', q' \mid adj'] \end{array}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k \mid p \cup p', q \cup q' \mid adj]}$$



donde se debe cumplir que ( $\text{nil} \in \text{adj}(M^\gamma)$  y  $\text{adj}' = \text{false}$ ) o ( $\exists \beta \in \text{adj}(M^\gamma)$  y  $\text{adj}' = \text{true}$ )

$$\mathcal{D}_{\text{EMix}}^{\text{SubsPred}} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q \mid \text{adj}]}{[\top \rightarrow \bullet \mathbf{R}^\alpha, j, j \mid -, - \mid \text{false}]} \quad \alpha \in \text{subst}(M^\gamma)$$

$$\mathcal{D}_{\text{EMix}}^{\text{SubsComp}} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q \mid \text{adj}], [\top \rightarrow \mathbf{R}^\alpha \bullet, j, k \mid -, - \mid \text{false}]}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k \mid p, q \mid \text{adj}]} \quad \alpha \in \text{subst}(M^\gamma)$$

$$\mathcal{D}_{\text{EMix}}^{\text{WAdjPred}} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q \mid \text{adj}]}{[\top \rightarrow \bullet \mathbf{R}^\beta, j, j \mid -, - \mid \text{false}]} \quad \beta \in \text{wadj}(M^\gamma)$$

$$\mathcal{D}_{\text{EMix}}^{\text{FootPred}} = \frac{[\mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, - \mid \text{false}]}{[M^\gamma \rightarrow \bullet \nu, k, k \mid -, - \mid \text{false}]} \quad \beta \in \text{wadj}(M^\gamma)$$

$$\mathcal{D}_{\text{EMix}}^{\text{FootComp}} = \frac{[\mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, - \mid \text{false}], [M^\gamma \rightarrow \nu \bullet, k, l \mid p', q' \mid \text{false}]}{[\mathbf{F}^\beta \rightarrow \perp \bullet, k, l \mid k, l \mid \text{false}]} \quad \beta \in \text{wadj}(M^\gamma)$$

$$\mathcal{D}_{\text{EMix}}^{\text{LAdjPred}} = \frac{[M^\gamma \rightarrow \bullet \nu, i, i \mid -, - \mid \text{false}]}{[\top \rightarrow \bullet \mathbf{R}^\beta, i, i \mid -, - \mid \text{false}]} \quad \beta \in \text{ladj}(M^\gamma)$$

$$\mathcal{D}_{\text{EMix}}^{\text{LAdjComp}} = \frac{[M^\gamma \rightarrow \bullet \nu, i, i \mid -, - \mid \text{false}], [\top \rightarrow \mathbf{R}^\beta \bullet, i, j \mid -, - \mid \text{false}]}{[M^\gamma \rightarrow \bullet \nu, i, j \mid -, - \mid \text{true}]} \quad \beta \in \text{ladj}(M^\gamma)$$

$$\mathcal{D}_{\text{EMix}}^{\text{RAAdjPred}} = \frac{[M^\gamma \rightarrow \nu \bullet, i, j \mid p, q \mid \text{false}]}{[\top \rightarrow \bullet \mathbf{R}^\beta, j, j \mid -, - \mid \text{false}]} \quad \beta \in \text{radj}(M^\gamma)$$

El conjunto de ítems finales es igual al del esquema  $\text{buE}^{\text{Mix}}$ :

$$\mathcal{F}_{\text{EMix}} = \mathcal{F}_{\text{buE}^{\text{Mix}}}$$

Mediante el paso  $\mathcal{D}_{\text{EMix}}^{\text{Ini}}$ , el analizador comienza lanzando la predicción de todos los árboles iniciales cuyas raíces estén etiquetadas con el axioma.

Los pasos deductivos que establecen la estrategia ascendente predictiva son los correspondientes a predicciones y compleciones. Para este algoritmo vamos a distinguir seis tipos de predicciones con sus pasos de compleción asociados: subárbol, adjunción izquierda, adjunción derecha, adjunción wrapping, sustitución y pie de los árboles auxiliares wrapping.

Los pasos de reconocimiento y compleción son similares a los mismos del esquema  $\text{buE}^{\text{Mix}}$ , salvo  $\mathcal{D}_{\text{EMix}}^{\text{FootComp}}$  y  $\mathcal{D}_{\text{EMix}}^{\text{LAdjComp}}$ . El paso  $\mathcal{D}_{\text{EMix}}^{\text{FootComp}}$  funciona igual que su homónimo en el esquema  $\mathbf{E}$ . La estrategia predictiva, a diferencia del esquema  $\text{buE}^{\text{Mix}}$ , obliga a que

las adjunciones izquierdas se efectúen antes de empezar a reconocer el subárbol. Esta consideración es recogida en el paso  $\mathcal{D}_{EMix}^{LAdjComp}$  y provoca que en el dominio aparezcan ítems  $(\mathcal{I}_{EMix}^{(i)})$  sin el punto al final de la regla que informan que un nodo que ha sido adjuntado por la izquierda. Muchos de los pasos del esquema deben adaptarse para que contemplen el caso que representa este tipo de ítems.

La complejidad espacial del algoritmo se mantiene en las mismas cotas que en el esquema  $buE^{Mix}$ . La complejidad temporal en el caso peor de:

- Las adjunciones de árboles auxiliares fuertemente izquierdos o derechos en:
  - un nodo de un árbol en  $I \cup A_{SL} \cup A_{SR}$  presentan una complejidad  $\mathcal{O}(n^3)$ , debido a los pasos  $\mathcal{D}_{EMix}^{RAdjComp}$  y  $\mathcal{D}_{EMix}^{Comp}$ .
  - un nodo que se encuentre en la espina de un árbol en  $A_W$  o un nodo dominado directamente por un nodo de la espina de un árbol en  $A_W$  y que esté en el contexto derecho de la espina de dicho árbol presentan una complejidad  $\mathcal{O}(n^5)$ , debido al paso  $\mathcal{D}_{EMix}^{Comp}$ .
  - un nodo que encuentre en el contexto izquierdo de la espina de un árbol en  $A_W$  o en el contexto derecho de un árbol en  $A_W$  de la espina pero no está dominado directamente por ningún nodo perteneciente a la espina presentan una complejidad  $\mathcal{O}(n^3)$ , debido a los pasos  $\mathcal{D}_{EMix}^{RAdjComp}$  y  $\mathcal{D}_{EMix}^{Comp}$ .
- Las adjunciones de árboles auxiliares wrapping en:
  - un nodo de un árbol en  $I \cup A_{SL} \cup A_{SR}$  presentan una complejidad  $\mathcal{O}(n^4)$ , debido al paso  $\mathcal{D}_{EMix}^{WAdjComp}$ .
  - un nodo de la espina de un árbol auxiliar en  $A_W$  presentan la complejidad máxima del algoritmo de  $\mathcal{O}(n^6)$ , debido al paso  $\mathcal{D}_{EMix}^{WAdjComp}$ .
  - un nodo dominado directamente por un nodo de la espina de un árbol en  $A_W$  y que se encuentra en el contexto derecho de la espina de dicho árbol presentan una complejidad  $\mathcal{O}(n^5)$ , debido al paso  $\mathcal{D}_{EMix}^{Comp}$ .
  - un nodo que encuentre en el contexto izquierdo de la espina de un árbol en  $A_W$  o en el contexto derecho de la espina de un árbol en  $A_W$  pero no está dominado directamente por ningún nodo perteneciente a la espina presentan una complejidad  $\mathcal{O}(n^4)$ , debido al paso  $\mathcal{D}_{EMix}^{WAdjComp}$ .
- Las operaciones de sustitución se realizan en tiempo  $\mathcal{O}(n^3)$ , salvo las sustituciones en los nodos dominados directamente por un nodo de la espina de un árbol en  $A_W$  y que se encuentran en el contexto derecho de la espina de dicho árbol, las cuales presentan una complejidad  $\mathcal{O}(n^5)$ .
- Las operaciones de reconocimiento y compleción permanecen en las mismas cotas de complejidad del esquema **E**. Aunque las operaciones de compleción en nodos pertenecientes a las espinas de árboles auxiliares en  $A_{SL} \cup A_{SR}$  se reducen a  $\mathcal{O}(n^3)$ .

Frase	E	E <sup>Mix</sup>	Mejora
Srini bought a book	0.61	0.49	19.67%
Srini bought Beth a book	0.77	0.71	7.79%
Srini bought a book at the bookstore	0.94	0.93	1.06%
he put the book on the table	0.83	0.71	14.46%
the sun melted the ice	0.71	0.66	7.04%
Elmo borrowed a book	0.55	0.49	10.91%
he hopes Muriel wins	0.93	0.77	17.20%
he hopes that Muriel wins	1.26	1.16	7.94%
the man who Muriel likes bought a book	2.14	1.48	30.84%
the man that Muriel likes bought a book	1.21	1.04	14.05%
Clove caught a frisbee	0.55	0.49	10.91%
who caught a frisbee	0.55	0.44	20.00%
the emu thinks that the aardvark smells terrible	1.48	1.32	10.81%
who does the emu think smells terrible	0.99	0.77	22.22%
who did the elephant think the panda heard the emu said smells terrible	3.13	2.36	24.60%
Herbert is angry and furious	0.55	0.55	0.00%
Herbert is more livid than angry	0.50	0.44	12.00%

Figura 10.3: *Comparativa entre analizadores tipo Earley para TAG*

### 10.5.1 Resultados experimentales

E y E<sup>Mix</sup> han sido comparados en la implementación de la máquina deductiva que hemos usado en los experimentos descritos en esta memoria. Hemos realizado varios experimentos. En primer lugar, hemos comparado el rendimiento de ambos algoritmos cuando trabajan sobre gramáticas TIG. Para ello se han diseñado dos gramáticas artificiales  $G_l$ , formada por los árboles  $\alpha_1$  y  $\beta_1$  del conjunto mostrado en la figura 5.3, y  $G_r$ , formada por los árboles  $\alpha_1$  y  $\beta_0$  del conjunto mostrado en la figura 5.3. Esperábamos que el analizador sintáctico combinado fuese considerablemente más rápido que el otro. En efecto, para  $G_l$  hemos observado que dicho analizador es hasta 18 veces más rápido que el analizador para TAG. Sin embargo, en el caso de  $G_r$  las diferencias entre ambos analizadores son mínimas.

Estos resultados fueron corroborados por un segundo experimento realizado sobre el mismo conjunto de TAG artificiales ( $G_1$ - $G_7$ ) empleado en la sección 5.1.1. Observamos que el rendimiento del analizador combinado mejora ostensiblemente cuando el análisis de la cadena de entrada involucra la realización de adjunciones de árboles fuertemente izquierdos.

Para el tercer experimento hemos tomado el subconjunto de 27 árboles de la gramática XTAG descrito en [35] con simulación de rasgos mediante restricciones locales. En la figura 10.3 se muestran los resultados obtenidos para algunas frases representativas. La aplicación del analizador combinado conlleva una reducción del tiempo de análisis del 31% en algunos casos, aunque dicha mejora depende del tipo de árboles auxiliares que se utilice en cada frase.

# Capítulo 11

## Conclusión

Partimos de un conjunto de analizadores tabulares para TAG definidos mediante esquemas de análisis sintáctico y derivados a partir de algoritmos conocidos para CFG con diversas estrategias. Concretamente, los esquemas basados en los algoritmos con estrategia ascendente **CYK**, **buE** y **dVH**, y los de estrategia ascendente predictiva **E** y **Nederhof**. Estos esquemas de análisis se organizan como una red de analizadores en base a las relaciones formales existentes entre ellos, la cual establece cómo se pueden derivar unos esquemas a partir de otros. En cuanto al formalismo TIG, sólo se disponía de un par de analizadores, uno basado en el método CYK, descrito en forma algorítmica, y otro basado en el algoritmo Earley, descrito mediante reglas deductivas.

Nuestro objetivo principal era definir nuevos analizadores para TAG que mejorasen el comportamiento práctico de los analizadores tabulares clásicos. Para ello decidimos emplear dos métodos: (1) extender el concepto de Left Corner de CFG a TAG y usarlo para definir esquemas más eficientes para TAG; (2) la construcción de analizadores que trabajasen de forma dinámica, comportándose como algoritmos de análisis para TIG cuando analizan árboles TIG y como algoritmos de análisis clásicos para TAG en el resto de situaciones.

Respecto al primer método, las principales aportaciones recogidas en la memoria son las siguientes:

- Un nuevo analizador ascendente, al que denominamos **buLC**, obtenido mediante la aplicación de un filtro al esquema **buE**.
- Se define la relación de Left Corner en TAG. Se introducen dos analizadores ascendentes con información predictiva, a los que denominamos **pLC** y **LC**, obtenidos mediante la aplicación de un filtro LC al esquema **E**. Además hemos incluido una versión que verifica la propiedad del prefijo válido para el esquema **pLC**.
- Un nuevo analizador, al que denominamos **LC'**, obtenido mediante un filtro sobre el esquema **LC** que elimina predicciones en el subárbol que domina el nodo pie durante una operación de adjunción.



- Aunque las mejoras dependerán en gran medida de las gramáticas y cadenas de entradas analizadas, hemos realizado experimentos con un grupo de gramáticas artificiales y un subconjunto de la gramática XTAG que ponen de manifiesto que los analizadores con filtros LC, tanto en la versión ascendente como predictiva, presentan comportamientos significativamente mejores.
- Demostramos las relaciones formales entre todos los esquemas tipo LC definidos. También probamos las relaciones entre éstos y los esquemas **buE** y **E**, ampliando así la red de analizadores para TAG ya establecida.

Para implementar al segundo método era necesario disponer de un conjunto de analizadores para TIG que empleasen estrategias similares a las de los existentes para TAG. Ante la inexistencia de estos esquemas, nos planteamos como subobjetivo la definición de los mismos, obteniendo los resultados siguientes:

- Un conjunto de nuevos analizadores con diversas estrategias: tres ascendentes (**CYK<sup>i</sup>**, **buE<sup>i</sup>** y **dVH<sup>i</sup>**) y uno predictivo (**Earley<sup>i</sup>**).
- Se define la relación de Left Corner en TIG. Se introducen dos analizadores ascendentes con información predictiva, a los que denominamos **pLC<sup>i</sup>** y **LC<sup>i</sup>**, obtenidos mediante la aplicación de un filtro LC al esquema **Earley<sup>i</sup>**.
- Demostramos las relaciones formales entre todos los esquemas definidos, creando de esta forma una red de analizadores para TIG similar a la existente para TAG.

De la integración de los esquemas para TAG y TIG con las mismas estrategias obtuvimos los siguientes resultados:

- Un conjunto de nuevos analizadores: tres ascendentes (**CYK<sup>Mix</sup>**, **buE<sup>Mix</sup>** y **dVH<sup>Mix</sup>**) y uno predictivo que no verifica la VPP (**E<sup>Mix</sup>**).
- Hacemos un estudio detallado de la complejidad computacional de todos estos algoritmos con objeto de mostrar en qué casos se producen reducciones de la complejidad.
- Hemos realizado experimentos con un grupo de gramáticas artificiales y un subconjunto de la gramática XTAG para comparar los comportamientos prácticos de los esquemas **E<sup>Mix</sup>** y **E**. Comprobándose que el analizador combinado, en general, presenta mejoras apreciables.

A raíz de este trabajo se abren numerosas líneas de investigación, entre las que podemos citar: el establecimiento de la relación de los esquemas LC con los dirigidos por núcleos, la ampliación de la red de analizadores definida para TIG, la construcción de esquemas que combinen las dos técnicas descritas en esta memoria, la creación de bancos de pruebas más extensos para llevar a cabo los experimentos o la definición de esquemas combinados que permitan adjunciones simultáneas sobre un nodo.

Parte V  
Apéndices

## Apéndice A

# Pruebas de corrección de esquemas para TAG

En este apéndice presentamos las pruebas de corrección de dos esquemas tipo Left Corner para TAG, concretamente el ascendente **buLC** y el predictivo **pLC**, con objeto de mostrar como la estrategia general propuesta por Sikkel en [126] se emplea para demostrar la corrección de esquemas de análisis para TAG. Las demostraciones de la corrección del resto de analizadores introducidos en la memoria se realizan de forma similar a éstas. Incluimos también la prueba de corrección del esquema **sLC** para poner de manifiesto como las relaciones formales entre analizadores ayudan en esta labor.

### A.1 Prueba de corrección del esquema buLC

El analizador es correcto si, para una gramática  $G \in TAG$  y una cadena de entrada  $a_1 \dots a_n$ , se cumple que:

$$[\top \rightarrow \mathbf{R}^\alpha \bullet, 0, n \mid -, -] \in \mathcal{V}_{\text{buLC}} \Leftrightarrow \top \Rightarrow \mathbf{R}^\alpha \xrightarrow{*} a_1 \dots a_n \wedge \alpha \in \mathbf{I}$$

Para demostrar la corrección vamos a proceder según el método descrito en [126]. De forma que para determinar el conjunto de ítems válidos  $\mathcal{V}_{\text{buLC}}$ , primero estableceremos un conjunto de *ítems viables*,  $\mathcal{W}_{\text{buLC}} \subseteq \mathcal{I}_{\text{buLC}}$ , que deben ser reconocidos por el sistema  $\mathbb{P}_{\text{buLC}}$  y a continuación probaremos que  $\mathcal{W}_{\text{buLC}} = \mathcal{V}_{\text{buLC}}$ .

#### Definición A.1 *ítems buLC viables*

Dada una gramática  $G \in TAG$  y una cadena de entrada  $a_1 \dots a_n$ :

- Un ítem  $\xi = [N^\gamma \rightarrow \delta \bullet \nu, i, m \mid p, q]$  es *buLC viable*, y los denotaremos como  $\xi \in \mathcal{W}_{\text{buLC}}$ , si cumple:
  - (i)  $\gamma \in \mathbf{I} \cup \mathbf{A}$
  - (ii) Para esta condición vamos a distinguir dos casos:

- **Caso 1:**  $(p, q) = (-, -)$ . Este ocurre cuando el nodo  $N^\gamma$  no está en la espina de un árbol auxiliar ( $\gamma \in I \cup A$ ) o  $N^\gamma$  pertenece a la espina ( $\gamma \in A$ ) pero no se ha completado el reconocimiento del nodo hijo que pertenece a la espina. Para este caso se debe cumplir que:

$$\delta\nu \xrightarrow{*} a_{i+1} \dots a_m \nu.$$

- **Caso 2:**  $(p, q) \neq (-, -)$ . Este ocurre cuando el nodo  $N^\gamma$  pertenece a la espina ( $\gamma \in A$ ) y lleva la información del subárbol que domina el nodo pie. Para este caso se debe cumplir que:

$$\delta\nu \xrightarrow{*} a_{i+1} \dots a_p \perp a_{q+1} \dots a_m \nu \Rightarrow_f a_{i+1} \dots a_p \omega a_{q+1} \dots a_m \nu \xrightarrow{*} a_{i+1} \dots a_m \nu,$$

donde  $\gamma \in \text{adj}(N^\varphi)$  y  $N^\varphi \rightarrow \omega$ .

- Las hipótesis  $[a, j - 1, j]$  son viables si  $a = a_j$ .

### Corolario A.1

Un ítem  $[\top \rightarrow \mathbf{R}^\alpha \bullet, 0, n \mid -, -] \in \mathcal{W}_{\text{buLC}}$  para una gramática  $G$  y una cadena de entrada  $a_1 \dots a_n$  si y solo si  $\top \Rightarrow \mathbf{R}^\alpha \xrightarrow{*} a_1 \dots a_n$  y  $\alpha \in I$ .

Ahora queda probar la equivalencia entre  $\mathcal{V}_{\text{buLC}}$  y  $\mathcal{W}_{\text{buLC}}$ .

### Lema A.1

$\mathcal{V}_{\text{buLC}} \subseteq \mathcal{W}_{\text{buLC}}$ , por tanto, el sistema  $\mathbb{P}_{\text{buLC}}$  es consistente.

#### Prueba

Por inducción sobre  $\vdash$ :

- todas las hipótesis son viables;
- para cada paso deductivo de  $\mathcal{D}_{\text{buLC}}$ , se cumple que de la viabilidad de los antecedentes se implica viabilidad del consecuente.

### Lema A.2

$\mathcal{W}_{\text{buLC}} \subseteq \mathcal{V}_{\text{buLC}}$ , por tanto, el sistema  $\mathbb{P}_{\text{buLC}}$  es completo.

#### Prueba

Para esta prueba construiremos una función de longitud de derivación (dlf)  $d$  sobre  $\mathcal{H}_{\text{buLC}} \cup \mathcal{W}_{\text{buLC}}$ , que nos permita probar la completitud por inducción en  $d(\xi)$  para cualquier  $\xi \in \mathcal{H}_{\text{buLC}} \cup \mathcal{W}_{\text{buLC}}$ . Nuestra dlf asignará a cada ítem el número de arcos atravesados en el recorrido ascendente del árbol de análisis que se necesita para derivarlo. Puesto que pueden existir varios árboles de análisis para el mismo ítem, tomaremos el valor mínimo de todos los recorridos.

La función parcial  $d: \mathcal{H}_{\text{buLC}} \cup \mathcal{W}_{\text{buLC}} \rightarrow \mathbb{N}$  se define mediante:

$$\bullet \quad d([N^\gamma \rightarrow \delta \bullet \nu, i, j \mid p, q]) = \min\{\mu \mid \delta \xrightarrow{\mu} a_{i+1} \dots a_j\}$$

- $d([a, j, j + 1]) = 0$

El motivo de esta definición es que cada paso en la derivación  $\delta \xrightarrow{\mu} a_{i+1} \dots a_j$  requiere una operación de completión o reconocimiento. Por tanto,  $\mu$  cuenta el número de completiones y reconocimientos necesarios para obtener  $[N^\gamma \rightarrow \delta \bullet \nu, i, j \mid p, q]$ .

Por simplicidad, en las adjunciones consideramos que existe un arco que baja desde el nodo adjuntado hasta el nodo top del árbol auxiliar. Sin embargo, no existirá ningún arco desde el nodo bottom hasta el nodo raíz del subárbol escindido, sino que este par de nodos se identifican según establece la operación de adjunción.

Es trivial demostrar que la función  $d$  está definida para cada  $\xi \in \mathcal{W}_{\text{buLC}}$ .

Para probar que  $d$  es una dlf sólo nos queda establecer que para cada  $\xi \in \mathcal{W}_{\text{buLC}}$  hay algún  $\eta_1, \dots, \eta_k \vdash \xi \in \mathcal{D}_{\text{buLC}}$  tal que  $\{\eta_1, \dots, \eta_k\} \subseteq \mathcal{W}_{\text{buLC}}$  y  $d(\eta_i) < d(\xi)$  para  $1 \leq i \leq k$ . Para ello vamos distinguir los siguientes casos:

- $\xi = [O^\gamma \rightarrow P^\gamma \bullet \nu, j, j + 1 \mid -, -]$  con  $\text{label}(P^\gamma) = a$ .

De la viabilidad de  $\xi$  tenemos que:

- (i)  $P^\gamma \Rightarrow a_{j+1}$ .

De (i) y de la condición de partida  $\text{label}(P^\gamma) = a$  obtenemos que:

$$\zeta = [a, j, j + 1] \in \mathcal{H}_{\text{buLC}}.$$

El paso deductivo  $\zeta \vdash \xi$  existe y es  $\mathcal{D}_{\text{buLC}}^{\text{LC}_t}$ .

Por definición de  $d$ :

$$\begin{aligned} d(\xi) &= 1, \\ d(\zeta) &= 0. \end{aligned}$$

Por lo que se cumple  $d(\xi) > d(\zeta)$ .

- $\xi = [N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, i, j + 1 \mid p, q]$  con  $\text{label}(M^\gamma) = a$ .

De la viabilidad de  $\xi$  tenemos que:

- (i.i) si  $(p, q) = (-, -)$

$$P^\gamma \delta M^\gamma \xrightarrow{\mu} a_{i+1} \dots a_j M^\gamma \Rightarrow a_{i+1} \dots a_{j+1};$$

- (i.ii) si  $(p, q) \neq (-, -)$

$$P^\gamma \delta M^\gamma \xrightarrow{\mu'} a_{i+1} \dots a_p \perp a_{q+1} \dots a_j M^\gamma \xrightarrow{\theta} a_{i+1} \dots a_j M^\gamma \Rightarrow a_{i+1} \dots a_{j+1},$$

donde  $\mu = \mu' + \theta$ .

De (i) y la condición de partida  $\text{label}(M^\gamma) = a$  tenemos que:

$$\zeta = [a, j, j + 1] \in \mathcal{H}_{\text{buLC}}.$$

De (i) se obtiene que:

$$\eta = [N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \nu, i, j \mid p, q] \in \mathcal{W}_{\text{buLC}}.$$

El paso deductivo  $\{\eta, \zeta\} \vdash \xi$  existe y es  $\mathcal{D}_{\text{buLC}}^{\text{Scan}}$ .

Por definición de  $d$ :

$$\begin{aligned} d(\xi) &= \min\{\mu + 1\} \\ d(\eta) &= \min\{\mu\} \\ d(\zeta) &= 0. \end{aligned}$$

Por lo que se cumple  $d(\xi) > d(\eta)$  y  $d(\xi) > d(\zeta)$ .

- $\xi = [N^\gamma \rightarrow P^\gamma \delta M^\gamma \bullet \nu, i, j \mid p, q]$  con  $\text{label}(M^\gamma) = \varepsilon$ .

De la viabilidad de  $\xi$  tenemos que:

(i.i) si  $(p, q) = (-, -)$

$$P^\gamma \delta M^\gamma \xrightarrow{\mu} a_{i+1} \dots a_j M^\gamma \Rightarrow a_{i+1} \dots a_j;$$

(i.ii) si  $(p, q) \neq (-, -)$

$$P^\gamma \delta M^\gamma \xrightarrow{\mu'} a_{i+1} \dots a_p \perp a_{q+1} \dots a_j M^\gamma \xrightarrow{\theta} a_{i+1} \dots a_j M^\gamma \Rightarrow a_{i+1} \dots a_j,$$

donde  $\mu = \mu' + \theta$ .

De (i) se obtiene que:

$$\eta = [N^\gamma \rightarrow P^\gamma \delta \bullet M^\gamma \nu, i, j \mid p, q] \in \mathcal{W}_{\text{buLC}}.$$

El paso deductivo  $\eta \vdash \xi$  existe y es  $\mathcal{D}_{\text{buLC}}^\varepsilon$ .

Por definición de  $d$ :

$$d(\xi) = \min\{\mu + 1\},$$

$$d(\eta) = \min\{\mu\}.$$

Por lo que se cumple  $d(\xi) > d(\eta)$ .

- $\xi = [O^\gamma \rightarrow P^\gamma \bullet \nu, j, k \mid p, q]$  con  $P^\gamma \in V_N$  y  $\text{label}(P^\gamma) \neq \perp$ .

De la viabilidad de  $\xi$  tenemos que:

(i.i) si  $(p, q) = (-, -)$

$$P^\gamma \Rightarrow \delta \xrightarrow{\mu} a_{j+1} \dots a_k;$$

(i.ii) si  $(p, q) \neq (-, -)$

$$P^\gamma \xrightarrow{\mu'} a_{j+1} \dots a_p \perp a_{q+1} \dots a_k \xrightarrow{\theta} a_{j+1} \dots a_k,$$

donde  $\mu = \mu' + \theta$ .

De (i) se obtiene que:

$$\eta = [P^\gamma \rightarrow \delta \bullet, j, k \mid p, q] \in \mathcal{W}_{\text{buLC}}.$$

El paso deductivo  $\eta \vdash \xi$  existe y es  $\mathcal{D}_{\text{buLC}}^{\text{LC}_n}$ .

Por definición de  $d$ :

$$d(\xi) = \min\{\mu + 1\},$$

$$d(\eta) = \min\{\mu\}.$$

Por lo que se cumple  $d(\xi) > d(\eta)$ .

- $\xi = [N^\gamma \rightarrow P^\gamma \nu M^\gamma \bullet \omega, i, k \mid p, q]$  con  $M^\gamma \in V_N$  y  $\text{nil} \in \text{adj}(M^\gamma)$ .

De la viabilidad de  $\xi$  tenemos que:

(i.i) si  $(p, q) = (-, -)$

$$P^\gamma \nu M^\gamma \xrightarrow{\mu} a_{i+1} \dots a_j M^\gamma \Rightarrow a_{i+1} \dots a_j \delta \xrightarrow{\theta} a_{i+1} \dots a_k,$$

donde  $M^\gamma \rightarrow \delta \in \mathcal{P}(\gamma)$ ;

(i.ii.i) si  $(p, q) \neq (-, -)$  y  $M^\gamma$  no pertenece a la espina de  $\gamma$

$$P^\gamma \nu M^\gamma \xrightarrow{\mu'} a_{i+1} \dots a_p \perp a_{q+1} \dots a_j M^\gamma \Rightarrow a_{i+1} \dots a_p \perp a_{q+1} \dots a_j \delta \xrightarrow{\theta} a_{i+1} \dots a_p \perp a_{q+1} \dots a_k,$$

donde  $\mu = \mu' + \theta$  y  $M^\gamma \rightarrow \delta \in \mathcal{P}(\gamma)$ ;

(i.ii.ii) si  $(p, q) \neq (-, -)$  y  $M^\gamma$  pertenece a la espina de  $\gamma$

$$\begin{aligned}
P^\gamma \nu M^\gamma &\xrightarrow{\mu} a_{i+1} \dots a_j M^\gamma \Rightarrow a_{i+1} \dots a_j \delta \xrightarrow{\iota'} \\
a_{i+1} \dots a_j \dots a_p \perp a_{q+1} \dots a_k &\xrightarrow{\theta} a_{i+1} \dots a_k, \\
\text{donde } \iota &= \iota' + \theta \text{ y } M^\gamma \rightarrow \delta \in \mathcal{P}(\gamma).
\end{aligned}$$

De (i) obtenemos que:

$$\begin{aligned}
\eta &= [N^\gamma \rightarrow P^\gamma \nu \bullet M^\gamma \omega, i, j \mid p', q'] \in \mathcal{W}_{\text{buLC}} \\
\zeta &= [M^\gamma \rightarrow \delta \bullet, j, k \mid p'', q''] \in \mathcal{W}_{\text{buLC}}.
\end{aligned}$$

Siempre teniendo en cuenta que la definición de adjunción obliga a que:

$$\begin{aligned}
\text{si } (p', q') &= (p'', q'') = (-, -) \text{ entonces } (p, q) = (-, -); \\
\text{si } (p', q') &\neq (-, -) \text{ entonces } (p'', q'') = (-, -), \text{ y } (p, q) = (p', q'); \\
\text{si } (p'', q'') &\neq (-, -) \text{ entonces } (p', q') = (-, -), \text{ y } (p, q) = (p'', q'').
\end{aligned}$$

El paso deductivo  $\{\eta, \zeta\} \vdash \xi$  existe y es  $\mathcal{D}_{\text{buLC}}^{\text{Comp}}$ .

Por definición de  $d$ :

$$\begin{aligned}
d(\xi) &= \min\{\mu + \iota + 1\}, \\
d(\eta) &= \min\{\mu\}, \\
d(\zeta) &= \min\{\iota\}.
\end{aligned}$$

Puesto que siempre  $\iota > 0$ , se cumple  $d(\xi) > d(\eta)$  y  $d(\xi) > d(\zeta)$ .

- $\xi = [\mathbf{F}^\beta \rightarrow \perp \bullet, k, l \mid k, l]$ .

De la viabilidad de  $\xi$  tenemos que:

$$\begin{aligned}
(i) \perp &\Rightarrow_f \nu \xrightarrow{\mu} a_{k+1} \dots a_l, \\
\text{donde } \beta &\in \text{adj}(M^\gamma) \text{ y } M^\gamma \rightarrow \nu \in \mathcal{P}(\gamma).
\end{aligned}$$

De (i) se implica que:

$$\eta = [M^\gamma \rightarrow \nu \bullet, k, l \mid p, q] \in \mathcal{W}_{\text{buLC}}.$$

El paso deductivo  $\eta \vdash \xi$  existe y es  $\mathcal{D}_{\text{buLC}}^{\text{FootComp}}$ .

Por definición de  $d$ :

$$\begin{aligned}
d(\xi) &= \min\{\mu + 1\}, \\
d(\eta) &= \min\{\mu\}.
\end{aligned}$$

Por lo que se cumple  $d(\xi) > d(\eta)$ .

- $\xi = [O^\gamma \rightarrow P^\gamma \bullet \nu, j, m \mid p, q]$  con  $\beta \in \text{adj}(P^\gamma)$ .

De la viabilidad de  $\xi$  y la condición inicial  $\beta \in \text{adj}(P^\gamma)$  tenemos que:

$$\begin{aligned}
(i.i) \text{ si } (p, q) &= (-, -) \\
P^\gamma &\Rightarrow_r \top \Rightarrow_r \mathbf{R}^\beta \xrightarrow{\mu} a_{j+1} \dots a_k \mathbf{F}^\beta a_{l+1} \dots a_m \Rightarrow_r \\
a_{j+1} \dots a_k \perp a_{l+1} \dots a_m &\Rightarrow_f a_{j+1} \dots a_k \delta a_{l+1} \dots a_m \xrightarrow{\iota} a_{j+1} \dots a_m, \\
\text{donde } P^\gamma &\rightarrow \delta \in \mathcal{P}(\gamma);
\end{aligned}$$

(i.ii) si  $(p, q) \neq (-, -)$

$$\begin{aligned}
P^\gamma &\Rightarrow_r \top \Rightarrow_r \mathbf{R}^\beta \xrightarrow{\mu} a_{j+1} \dots a_k \mathbf{F}^\beta a_{l+1} \dots a_m \Rightarrow_r a_{j+1} \dots a_k \perp a_{l+1} \dots a_m \Rightarrow_f \\
a_{j+1} \dots a_k \delta a_{l+1} \dots a_m &\xrightarrow{\iota'} a_{j+1} \dots a_k \dots a_p \perp a_{q+1} \dots a_m \xrightarrow{\theta} a_{j+1} \dots a_m, \\
\text{donde } \iota &= \iota' + \theta \text{ y } P^\gamma \rightarrow \delta \in \mathcal{P}(\gamma).
\end{aligned}$$

De (i) obtenemos que:

$$\eta = [\top \rightarrow \mathbf{R}^\beta \bullet, j, m \mid k, l] \in \mathcal{W}_{\text{buLC}}.$$

$$\zeta = [P^\gamma \rightarrow \delta \bullet, k, l \mid p, q] \in \mathcal{W}_{\text{buLC}}.$$

El paso deductivo  $\{\eta, \zeta\} \vdash \xi$  existe y es  $\mathcal{D}_{\text{buLC}}^{\text{LC}_{\text{cad}}}$ .

Por definición de  $d$ :

$$d(\xi) = \min\{\mu + \iota + 4\},$$

$$d(\eta) = \min\{\mu + \iota + 2\},$$

$$d(\zeta) = \min\{\iota\}.$$

Luego se cumple  $d(\xi) > d(\eta)$  y  $d(\xi) > d(\zeta)$ .

- $\xi = [N^\gamma \rightarrow P^\gamma \nu M^\gamma \bullet \omega, i, m \mid p, q]$  con  $\beta \in \text{adj}(M^\gamma)$ .

De la viabilidad de  $\xi$  y la condición inicial  $\beta \in \text{adj}(M^\gamma)$  tenemos que:

(i.i) si  $(p, q) = (-, -)$

$$P^\gamma \nu M^\gamma \xrightarrow{\mu} a_{i+1} \dots a_j M^\gamma \Rightarrow_r a_{i+1} \dots a_j \top \Rightarrow_r a_{i+1} \dots a_j \mathbf{R}^\beta \xrightarrow{\kappa}_r$$

$$a_{i+1} \dots a_k \mathbf{F}^\beta a_{l+1} \dots a_m \Rightarrow_r a_{i+1} \dots a_k \perp a_{l+1} \dots a_m \Rightarrow_f$$

$$a_{i+1} \dots a_k \delta a_{l+1} \dots a_m \xrightarrow{\iota}_f a_{i+1} \dots a_m,$$

donde  $P^\gamma \rightarrow \delta \in \mathcal{P}(\gamma)$ ;

(i.ii.i) si  $(p, q) \neq (-, -)$  y  $M^\gamma$  no pertenece a la espina de  $\gamma$

$$P^\gamma \nu M^\gamma \xrightarrow{\mu'} a_{i+1} \dots a_p \perp a_{q+1} \dots a_j M^\gamma \xrightarrow{\theta}_f a_{i+1} \dots a_j M^\gamma \Rightarrow_r$$

$$a_{i+1} \dots a_j \top \Rightarrow_r a_{i+1} \dots a_j \mathbf{R}^\beta \xrightarrow{\kappa}_r a_{i+1} \dots a_k \mathbf{F}^\beta a_{l+1} \dots a_m \Rightarrow_r$$

$$a_{i+1} \dots a_k \perp a_{l+1} \dots a_m \Rightarrow_f a_{i+1} \dots a_k \delta a_{l+1} \dots a_m \xrightarrow{\iota}_f a_{i+1} \dots a_m,$$

donde  $\mu = \mu' + \theta$  y  $P^\gamma \rightarrow \delta \in \mathcal{P}(\gamma)$ ;

(i.ii.ii) si  $(p, q) \neq (-, -)$  y  $M^\gamma$  pertenece a la espina de  $\gamma$

$$P^\gamma \nu M^\gamma \xrightarrow{\mu} a_{i+1} \dots a_j M^\gamma \Rightarrow_r a_{i+1} \dots a_j \top \Rightarrow_r a_{i+1} \dots a_j \mathbf{R}^\beta \xrightarrow{\kappa}_r$$

$$a_{i+1} \dots a_k \mathbf{F}^\beta a_{l+1} \dots a_m \Rightarrow_r a_{i+1} \dots a_k \perp a_{l+1} \dots a_m \Rightarrow_f$$

$$a_{i+1} \dots a_k \delta a_{l+1} \dots a_m \xrightarrow{\iota'}_f a_{i+1} \dots a_p \perp a_{q+1} \dots a_m \xrightarrow{\theta}_f a_{i+1} \dots a_m,$$

donde  $\iota = \iota' + \theta$  y  $P^\gamma \rightarrow \delta \in \mathcal{P}(\gamma)$ .

De (i) obtenemos que:

$$\psi = [N^\gamma \rightarrow P^\gamma \nu \bullet M^\gamma \omega, i, j \mid p', q']$$

$$\eta = [\top \rightarrow \mathbf{R}^\beta \bullet, j, m \mid k, l] \in \mathcal{W}_{\text{buLC}}$$

$$\zeta = [M^\gamma \rightarrow \delta \bullet \nu, k, l \mid p'', q''] \in \mathcal{W}_{\text{buLC}}.$$

Teniendo en cuenta que:

si  $(p', q') = (p'', q'') = (-, -)$  entonces  $(p, q) = (-, -)$ ;

si  $(p', q') \neq (-, -)$  entonces  $(p'', q'') = (-, -)$ , y  $(p, q) = (p', q')$ ;

si  $(p'', q'') \neq (-, -)$  entonces  $(p', q') = (-, -)$ , y  $(p, q) = (p'', q'')$ .

El paso deductivo  $\{\psi, \eta, \zeta\} \vdash \xi$  existe y es  $\mathcal{D}_{\text{buLC}}^{\text{AdjComp}}$ .

Por definición de  $d$ :

$$d(\xi) = \min\{\mu + \kappa + \iota + 4\},$$

$$d(\psi) = \min\{\mu\},$$

$$d(\eta) = \min\{\kappa + \iota + 2\},$$

$$d(\zeta) = \min\{\iota\}.$$

Por lo que se cumple  $d(\xi) > d(\eta)$ ,  $d(\xi) > d(\zeta)$  y  $d(\xi) > d(\psi)$ .

**Teorema A.1** *Corrección del analizador buLC*

Para cualquier gramática  $G \in TAG$  y una cadena de entrada  $a_1 \dots a_n$  se cumple que:

$$[\top \rightarrow \mathbf{R}^\alpha \bullet, 0, n \mid -, -] \in \mathcal{V}_{buLC} \text{ si y solo si } \top \Rightarrow \mathbf{R}^\alpha \xRightarrow{*} a_1 \dots a_n \wedge \alpha \in I$$

*Prueba*

Directamente de los lemas A.1 y A.2 y el corolario A.1.

## A.2 Prueba de corrección del esquema pLC

El analizador es correcto si, para una gramática  $G \in TAG$  y una cadena de entrada  $a_1 \dots a_n$  se cumple que

$$[\top; \top \rightarrow \mathbf{R}^\alpha \bullet, 0, n \mid -, -] \in \mathcal{V}_{pLC} \Leftrightarrow \top \Rightarrow \mathbf{R}^\alpha \xRightarrow{*} a_1 \dots a_n \wedge \alpha \in I$$

Sin embargo, no es un problema trivial el determinar el conjunto de ítems válidos  $\mathcal{V}_{pLC}$ . Vamos a proceder según el método descrito en los preliminares. Primero estableceremos un conjunto de *ítems viables*,  $\mathcal{W}_{pLC} \subseteq \mathcal{I}_{pLC}$ , que deben ser reconocidos por el sistema  $\mathbb{P}_{pLC}$  y, a continuación, probaremos que  $\mathcal{W}_{pLC} = \mathcal{V}_{pLC}$ .

**Definición A.2** *ítems pLC viables  $\mathcal{W}_{pLC}$*

Vamos a definir la viabilidad para cada tipo de ítem del sistema.

- Denotaremos con  $\Omega_{pLC}$  el conjunto de ítems predictivos viables. Un ítem predictivo  $[M^\gamma, j]$  es viable si existen  $C^\gamma, N^\gamma, \delta, \nu, i, p, q$  que cumplen las siguientes condiciones:
  - (i)  $M^\gamma \in V_N$
  - (ii)  $[C^\gamma; N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q] \in \Delta_{pLC}$
- Vamos a denotar como  $\Delta_{pLC}$  al conjunto de ítems left corner viables. Un ítem left corner  $[C^\gamma; M^\gamma \rightarrow \delta \bullet \nu, i, j \mid p, q]$  es viable si cumple:
  - (i)  $\gamma \in I \cup A$
  - (ii) En función del nodo ancla del ítem, vamos a distinguir cuatro casos para esta condición:
    - **Caso 1:** El nodo ancla ha sido predicho mediante un ítem predictivo:
      - (ii.i)  $[C^\gamma, i] \in \Omega_{pLC}$
      - (ii.ii)  $\mathbf{nil} \in \mathit{adj}(C^\gamma)$
    - **Caso 2:** Tal como hemos definido los ítems predictivos viables, los nodos top de los árboles auxiliares no pueden aparecer en el conjunto  $\Omega_{pLC}$ . Sin embargo, estos nodos son anclas en relaciones de esquina izquierda y esta situación se recoge en este caso:
      - (ii.i)  $\mathit{label}(C^\gamma) = \top$
      - (ii.ii)  $\gamma \in A$



(ii.iii) existe un  $N^\varphi$  tal que:

$$\begin{aligned} [N^\varphi, i] &\in \Omega_{\text{pLC}} \\ \gamma &\in \text{adj}(N^\varphi) \end{aligned}$$

- **Caso 3:** Si el nodo ancla domina el árbol escindido por una adjunción, dicho nodo tampoco ha sido predicho mediante un ítem predictivo (recordamos que el nodo que ha sido predicho es el nodo que domina la raíz del árbol auxiliar adjuntado). En este caso se debe cumplir que exista un árbol auxiliar que haya sido adjuntado en el nodo ancla y que dicho árbol auxiliar haya sido reconocido hasta su nodo bottom. Por tanto, debe existir un  $\beta$  tal que:

$$\begin{aligned} \beta &\in \text{adj}(C^\gamma) \\ [E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, i, i \mid -, -] &\in \Delta_{\text{pLC}} \end{aligned}$$

- **Caso 4:** Tal como hemos definido los ítems predictivos viables, los nodos top de los árboles iniciales no pueden aparecer en el conjunto  $\Omega_{\text{pLC}}$ . Sin embargo, al igual que ocurría en el caso 2 con los árboles auxiliares, estos nodos son anclas en relaciones de esquina izquierda y esta situación se recoge en este caso:

$$\begin{aligned} \text{(ii.i)} \quad \text{label}(C^\gamma) &= \top \\ \text{(ii.ii)} \quad \gamma &\in \mathbf{I} \end{aligned}$$

(iii)  $C^\gamma >_l^* M^\gamma$

(iv)  $\top \Rightarrow \mathbf{R}^\alpha \xrightarrow{*} a_1 \dots a_i M^\gamma \omega \Rightarrow a_1 \dots a_i \delta \nu \omega$

$$\text{con } \alpha = \gamma \Leftrightarrow \gamma \in \mathbf{I}$$

(v) Vamos a distinguir dos casos:

- **Caso 1:**  $(p, q) = (-, -)$ . Este ocurre cuando el nodo  $M^\gamma$  no está en la espina de un árbol auxiliar ( $\gamma \in \mathbf{I} \cup \mathbf{A}$ ) ó  $M^\gamma$  pertenece a la espina ( $\gamma \in \mathbf{A}$ ) pero no se ha completado el reconocimiento del nodo hijo que pertenece a la espina. Para este caso se debe cumplir que:

$$\delta \xrightarrow{*} a_{i+1} \dots a_j$$

- **Caso 2:**  $(p, q) \neq (-, -)$ . Este ocurre cuando el nodo  $M^\gamma$  pertenece a la espina ( $\gamma \in \mathbf{A}$ ) y lleva la información del subárbol que domina el nodo pie. Para este caso se deben cumplir que:

$$\delta \xrightarrow{*} a_{i+1} \dots a_p \perp a_{q+1} \dots a_j \Rightarrow_f a_{i+1} \dots a_p \omega a_{q+1} \dots a_j \xrightarrow{*}_f a_{i+1} \dots a_j,$$

donde  $N^\varphi \rightarrow \omega$  y  $\gamma \in \text{adj}(N^\varphi)$

Y definiremos el conjunto de ítems pLC viables como:

$$\mathcal{W}_{\text{pLC}} = \Omega_{\text{pLC}} \cup \Delta_{\text{pLC}}$$

- Las hipótesis  $[a, j-1, j]$  son viables si  $a = a_j$ .

### Corolario A.2

Un ítem  $[\top; \top \rightarrow \mathbf{R}^\alpha \bullet, 0, n \mid -, -] \in \mathcal{W}_{\text{pLC}}$  para una gramática  $G \in \text{TAG}$  y una cadena de entrada  $a_1 \dots a_n$  si y solo si  $\top \Rightarrow \mathbf{R}^\alpha \xrightarrow{*} a_1 \dots a_n$  y  $\alpha \in \mathbf{I}$ .

Ahora queda probar la equivalencia entre  $\mathcal{V}_{\text{pLC}}$  y  $\mathcal{W}_{\text{pLC}}$ .

**Lema A.3**

$\mathcal{V}_{\text{pLC}} \subseteq \mathcal{W}_{\text{pLC}}$ , por tanto, el sistema  $\mathbb{P}_{\text{pLC}}$  es consistente.

*Prueba*

Por inducción sobre  $\vdash$ :

- todas las hipótesis son viables;
- para cada paso deductivo de  $\mathcal{D}_{\text{pLC}}$ , se cumple que la viabilidad del consecuente se implica de la viabilidad de los antecedentes.

**Lema A.4**

$\mathcal{W}_{\text{pLC}} \subseteq \mathcal{V}_{\text{pLC}}$ , por tanto, el sistema  $\mathbb{P}_{\text{pLC}}$  es completo.

*Prueba*

Para esta prueba construiremos una función de longitud de derivación (dlf)  $d$  sobre  $\mathcal{H}_{\text{pLC}} \cup \mathcal{W}_{\text{pLC}}$ , que nos permita probar la completitud por inducción en  $d(\xi)$  para cualquier  $\xi \in \mathcal{H}_{\text{pLC}} \cup \mathcal{W}_{\text{pLC}}$ . Nuestra dlf asignará a cada ítem el número de arcos atravesados en el recorrido (de arriba a abajo y de izquierda a derecha) del árbol de análisis que se necesita para derivarlo. Puesto que pueden existir varios árboles de análisis para el mismo ítem, se debe tomar siempre el valor mínimo de todos los recorridos.

Vamos a hacer uso de la siguiente notación:

- $\pi$ : número de arcos atravesados en el recorrido hasta un ítem predictivo.
- $\rho$ : número de arcos atravesados en el recorrido hasta un ítem left corner.
- $\lambda$ : número de arcos atravesados en dirección descendente por la relación de esquina izquierda en un ítem left corner.
- $\mu$ : longitud de la derivación  $\delta \xrightarrow{H} a_{i+1} \dots a_j$  para ítems left corner de la forma  $[C^\gamma; M^\gamma \rightarrow \delta \bullet \nu, i, j \mid p, q]$ .

La función parcial  $d: \mathcal{H}_{\text{pLC}} \cup \mathcal{W}_{\text{pLC}} \rightarrow \mathbb{N}$  se define por:

- $d([M^\gamma, j]) = \min\{\rho + 1 \mid \rho = d([C^\gamma; N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q])\}$ .

Por la propia definición de  $\Omega_{\text{pLC}}$  se deduce que si  $[M^\gamma, j] \in \Omega_{\text{pLC}}$  entonces existe un  $[C^\gamma; N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q] \in \Delta_{\text{pLC}}$ .

Además, se le suma uno al número de arcos atravesados hasta el ítem left corner porque se genera un arco más en dirección descendente desde  $N^\gamma$  a  $M^\gamma$ .

- $d([C^\gamma; N^\gamma \rightarrow \delta \bullet \nu, i, j \mid p, q]) = \min\{\pi + \lambda + 2\mu \mid$ 
  - (i)  $\pi = d([C^\gamma, i])$
  - (ii)  $C^\gamma >_i^\lambda N^\gamma$
  - (iii) Distinguimos dos casos:
    - Caso 1:  $(p, q) \neq (-, -)$ 

$$\delta \xrightarrow{\mu'} a_{i+1} \dots a_p \perp a_{q+1} \dots a_j \xrightarrow{\mu''} a_{i+1} \dots a_j,$$
donde  $\mu = \mu' + \mu''$ ;
    - Caso 2:  $(p, q) = (-, -)$ 

$$\delta \xrightarrow{\mu} a_{i+1} \dots a_j \}.$$

Esta función sólo es aplicable cuando se cumple que:

- (i)  $[C^\gamma, i] \in \Omega_{\text{pLC}}$
- (ii)  $\text{nil} \in \text{adj}(C^\gamma)$ .

Por consiguiente, el número de arcos atravesados hasta este tipo de ítems viene dado por la suma de:

- el número de arcos atravesados hasta el nodo ancla  $C^\gamma$ , tanto en sentido descendente como ascendente;
  - el número de arcos atravesados en sentido descendente por la relación de esquina izquierda desde el nodo  $C^\gamma$  hasta el nodo  $N^\gamma$ ;
  - y el número de arcos atravesados en la derivación de  $\delta$  (condición (iii)). Para este caso, se cuentan  $2\mu$  porque cada arco es atravesado dos veces, una en dirección descendente y otra ascendente.
- $d([\top; O^\beta \rightarrow \delta \bullet \nu, j, k \mid p, q]) = \min\{\pi + 1 + \lambda + 2\mu \mid$ 
    - (i)  $\pi = d([M^\gamma, j])$
    - (ii)  $\top >_i^\lambda O^\beta$
    - (iii) Distinguimos dos casos:
      - Caso 1:  $(p, q) \neq (-, -)$ 

$$\delta \xrightarrow{\mu'} a_{j+1} \dots a_p \perp a_{q+1} \dots a_k \xrightarrow{\mu''} a_{j+1} \dots a_k,$$
donde  $\mu = \mu' + \mu''$ ;
      - Caso 2:  $(p, q) = (-, -)$ 

$$\delta \xrightarrow{\mu} a_{j+1} \dots a_k \}.$$

Esta función sólo es aplicable cuando se cumple que:

- (i)  $\beta \in \mathbf{A}$
- (ii) Existe un  $M^\gamma$  tal que:
  - $[M^\gamma, j] \in \Omega_{\text{pLC}}$
  - $\beta \in \text{adj}(M^\gamma)$ .

El número de arcos atravesados hasta este tipo de ítems viene dado por la suma de:

- el número de arcos atravesados, tanto en sentido descendente como ascendente, hasta el nodo  $M^\gamma$  donde se está llevando a cabo la adjunción del árbol auxiliar  $\beta$ ;
- el número de arcos atravesados en sentido descendente por la relación de esquina izquierda desde el nodo top del árbol auxiliar  $\beta$  hasta el nodo  $O^\beta$ ;
- y el número de arcos atravesados en la derivación de  $\delta$  (condición (iii)).

Obsérvese que la suma anterior se incrementa en uno, debido a que hemos convenido que existe un arco desde el nodo adjuntable  $M^\gamma$  hasta el nodo  $\top$  del árbol auxiliar  $\beta$ , el cual es atravesado en dirección descendente.

$$\bullet d([\top; O^\alpha \rightarrow \delta \bullet \nu, 0, j \mid -, -]) = \min\{\lambda + 2\mu \mid$$

$$(i) \top >_l^\lambda O^\alpha$$

$$(ii) \delta \xrightarrow{\mu} a_1 \dots a_j\}.$$

Esta función sólo es aplicable cuando se cumple que  $\alpha \in I$ .

El número de arcos atravesados hasta este tipo de ítems viene dado por la suma de:

- el número de arcos atravesados en sentido descendente por la relación de esquina izquierda desde el nodo top del árbol inicial  $\alpha$  hasta el nodo  $O^\alpha$ ;
- y el número de arcos atravesados en la derivación de  $\delta$  (condición (ii)).

$$\bullet d([M^\gamma; O^\gamma \rightarrow \delta \bullet \nu, k, l \mid p, q]) = \min\{\rho + \lambda + 2\mu \mid$$

$$(i) \rho = d([E^\beta; F^\beta \rightarrow \bullet \perp, k, k \mid -, -])$$

$$(ii) M^\gamma >_l^\lambda O^\gamma$$

(iii) Distinguimos dos casos:

- Caso 1:  $(p, q) \neq (-, -)$

$$\delta \xrightarrow{\mu} a_{k+1} \dots a_p \perp a_{q+1} \dots a_l \xrightarrow{\mu''} a_{k+1} \dots a_l,$$

$$\text{donde } \mu = \mu' + \mu'';$$

- Caso 2:  $(p, q) = (-, -)$

$$\delta \xrightarrow{\mu} a_{k+1} \dots a_l\}.$$

Esta función sólo es aplicable cuando se cumple que:

(i) Existe un árbol  $\beta \in \mathbf{A}$  tal que:

$$[E^\beta; F^\beta \rightarrow \bullet \perp, k, k \mid -, -] \in \Delta_{\text{pLC}}$$

(ii)  $\beta \in \text{adj}(M^\gamma)$ .

El número de arcos atravesados hasta este tipo de ítems viene dado por la suma de:

- el número de arcos atravesados, tanto en sentido descendente como ascendente, hasta el nodo bottom del árbol auxiliar  $\beta$  adjuntable en  $M^\gamma$ ;
- el número de arcos atravesados (en sentido descendente) por la relación de esquina izquierda desde el nodo  $M^\gamma$  hasta el nodo  $O^\gamma$ ;
- y el número de arcos atravesados en la derivación de  $\delta$  (condición (iii)).

- $d([a, j - 1, j]) = 0$

Es trivial demostrar que la función  $d$  está definida para cada  $\xi \in \mathcal{W}_{\text{pLC}}$ , ya que se puede verificar que para cada ítem viable hay al menos un  $\pi, \lambda, \mu, \rho$  para el cual se satisfacen las condiciones descritas en la definición y, por tanto, (por inducción en la definición de viabilidad)  $d$  está definida para todos los ítems viables.

Para probar que  $d$  es una dlf sólo nos queda establecer que para cada  $\xi \in \mathcal{W}_{\text{pLC}}$  hay algún  $\eta_1, \dots, \eta_k \vdash \xi \in \mathcal{D}_{\text{pLC}}$  tal que  $\{\eta_1, \dots, \eta_k\} \subseteq \mathcal{W}_{\text{pLC}}$  y  $d(\eta_i) < d(\xi)$  para  $1 \leq i \leq k$ . Para ello vamos distinguir los siguientes casos:

- $\xi = [M^\gamma, j]$ .

De la viabilidad de  $\xi$  tenemos que existen  $C^\gamma, N^\gamma, \delta, \nu, i, p, q$  tales que  $\eta = [C^\gamma; N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p, q] \in \mathcal{W}_{\text{pLC}}$ .

El paso deductivo  $\eta \vdash \xi$  existe y es  $\mathcal{D}_{\text{pLC}}^{\text{Pre}}$ .

Por definición de  $d$ :

$$d(\xi) = d(\eta) + 1.$$

Por lo que se cumple  $d(\xi) > d(\eta)$ .

- $\xi = [M^\gamma; O^\gamma \rightarrow P^\gamma \bullet \nu, j, j + 1 \mid -, -]$  con  $\text{label}(P^\gamma) = a$ .

De la viabilidad de  $\xi$  tenemos que:

$$(i) M^\gamma >_\ell^\lambda O^\gamma,$$

$$(ii) P^\gamma \Rightarrow a_{j+1}.$$

De (ii) obtenemos que:

$$\zeta = [a, j, j + 1] \in \mathcal{H}_{\text{pLC}}.$$

Por definición de  $d$ :

$$d(\zeta) = 0.$$

Distinguiremos los cuatro casos descritos en la condición (ii) de la definición de  $\Delta_{\text{pLC}}$ :

- **Caso 1:**  $\eta = [M^\gamma, j] \in \Omega_{\text{pLC}}$ .

El paso deductivo  $\{\eta, \zeta\} \vdash \xi$  existe y es  $\mathcal{D}_{\text{pLC}}^{\text{LC}_t}$ .

Por definición de  $d$ :

$$d(\xi) = \min\{\pi + \lambda + 2\mu\} \text{ con } \pi = d(\eta).$$

Por (ii) nos queda:

$$d(\xi) = \min\{\pi + \lambda + 2\}.$$

El incremento en dos significa que el reconocimiento de un símbolo de la cadena de entrada implica atravesar un arco dos veces, una en dirección descendente y otra ascendente.

Por lo que se cumple  $d(\xi) > d(\eta)$  y  $d(\xi) > d(\zeta)$ .

- **Caso 2:** existe  $N^\varphi$  tal que:

$$\eta = [N^\varphi, j] \in \Omega_{\text{pLC}},$$

$$\gamma \in \text{adj}(N^\varphi).$$

El paso deductivo  $\{\eta, \zeta\} \vdash \xi$  existe y es  $\mathcal{D}_{\text{pLC}}^{\text{LA}_t}$ .

Por definición de  $d$ :

$$d(\xi) = \min\{\pi + 1 + \lambda + 2\} \text{ con } \pi = d(\eta).$$

Por lo que se cumple  $d(\xi) > d(\eta)$  y  $d(\xi) > d(\zeta)$ .

– **Caso 3:** existe un  $\beta$  tal que:

$$\eta = [E^\beta; \mathbf{F}^\beta \rightarrow \bullet\perp, j, j \mid -, -] \in \Delta_{\text{pLC}},$$

$$\beta \in \text{adj}(M^\gamma).$$

El paso deductivo  $\{\eta, \zeta\} \vdash \xi$  existe y es  $\mathcal{D}_{\text{pLC}}^{\text{LF}_t}$ .

Por definición de  $d$ :

$$d(\xi) = \min\{\rho + \lambda + 2\} \text{ con } \rho = d(\eta).$$

Por lo que se cumple  $d(\xi) > d(\eta)$  y  $d(\xi) > d(\zeta)$ .

– **Caso 4:** El paso deductivo  $\zeta \vdash \xi$  existe y es  $\mathcal{D}_{\text{pLC}}^{\text{LI}_t}$ .

Por definición de  $d$ :

$$d(\xi) = \min\{\lambda + 2\}.$$

Por lo que se cumple  $d(\xi) > d(\zeta)$ .

•  $\xi = [M^\gamma; O^\gamma \rightarrow P^\gamma \bullet\nu, j, j \mid -, -]$  con  $\text{label}(P^\gamma) = \epsilon$ .

De la viabilidad de  $\xi$  tenemos que:

(i)  $M^\gamma >_\ell^\lambda O^\gamma$ .

Distinguiremos los cuatro casos descritos en la condición (ii) de la definición de  $\Delta_{\text{pLC}}$ :

– **Caso 1:**  $\eta = [M^\gamma, j] \in \Omega_{\text{pLC}}$ .

El paso deductivo  $\eta \vdash \xi$  existe y es  $\mathcal{D}_{\text{pLC}}^{\text{LC}_\epsilon}$ .

Por definición de  $d$ :

$$d(\xi) = \min\{\pi + \lambda + 2\} \text{ con } \pi = d(\eta).$$

El reconocimiento de  $\epsilon$ , considerado como símbolo de la entrada, implica también que se atraviere un arco dos veces.

Por lo que se cumple  $d(\xi) > d(\eta)$ .

– **Caso 2:** existe  $N^\varphi$  tal que:

$$\eta = [N^\varphi, j] \in \Omega_{\text{pLC}},$$

$$\gamma \in \text{adj}(N^\varphi).$$

El paso deductivo  $\eta \vdash \xi$  existe y es  $\mathcal{D}_{\text{pLC}}^{\text{LA}_\epsilon}$ .

Por definición de  $d$ :

$$d(\xi) = \min\{\pi + 1 + \lambda + 2\} \text{ con } \pi = d(\eta).$$

Por lo que se cumple  $d(\xi) > d(\eta)$ .

– **Caso 3:** existe un  $\beta$  tal que:

$$\eta = [E^\beta; \mathbf{F}^\beta \rightarrow \bullet\perp, j, j \mid -, -] \in \Delta_{\text{pLC}},$$

$$\beta \in \text{adj}(M^\gamma).$$

El paso deductivo  $\eta \vdash \xi$  existe y es  $\mathcal{D}_{\text{pLC}}^{\text{LF}_\epsilon}$ .

Por definición de  $d$ :



$$d(\xi) = \min\{\rho + \lambda + 2\} \text{ con } \rho = d(\eta).$$

Por lo que se cumple  $d(\xi) > d(\eta)$ .

– **Caso 4:**

El paso deductivo  $\vdash \xi$  existe y es  $\mathcal{D}_{\text{pLC}}^{\text{LI}\epsilon}$ .

Puesto que no tiene antecedentes, no es necesario demostrar el crecimiento de la función.

- $\xi = [M^\gamma; O^\gamma \rightarrow \bullet P^\gamma \nu, j, j \mid -, -]$  con  $P^\gamma >_\ell \Delta$  ó  $\text{label}(P^\gamma) = \perp$ .

De la viabilidad de  $\xi$  tenemos que:

$$M^\gamma >_\ell^\lambda O^\gamma.$$

Distinguiremos los cuatro casos descritos en la condición (ii) de la definición de  $\Delta_{\text{pLC}}$ :

– **Caso 1:**  $\eta = [M^\gamma, j] \in \Omega_{\text{pLC}}$ .

El paso deductivo  $\eta \vdash \xi$  existe y es  $\mathcal{D}_{\text{pLC}}^{\text{LCpre}}$ .

Por definición de  $d$ :

$$d(\xi) = \min\{\pi + \lambda\} \text{ con } \pi = d(\eta).$$

Por lo que se cumple  $d(\xi) > d(\eta)$ .

– **Caso 2:** existe  $N^\varphi$  tal que:

$$\eta = [N^\varphi, j] \in \Omega_{\text{pLC}},$$

$$\gamma \in \text{adj}(N^\varphi).$$

El paso deductivo  $\eta \vdash \xi$  existe y es  $\mathcal{D}_{\text{pLC}}^{\text{LApre}}$ .

Por definición de  $d$ :

$$d(\xi) = \min\{\pi + 1 + \lambda\} \text{ con } \pi = d(\eta).$$

Por lo que se cumple  $d(\xi) > d(\eta)$ .

– **Caso 3:** existe un  $\beta$  tal que:

$$\eta = [E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, j, j \mid -, -] \in \Delta_{\text{pLC}},$$

$$\beta \in \text{adj}(M^\gamma).$$

El paso deductivo  $\eta \vdash \xi$  existe y es  $\mathcal{D}_{\text{pLC}}^{\text{LFpre}}$ .

Por definición de  $d$ :

$$d(\xi) = \min\{\rho + \lambda\} \text{ con } \rho = d(\eta).$$

Por lo que se cumple  $d(\xi) > d(\zeta)$ .

– **Caso 4:** El paso deductivo  $\vdash \xi$  existe y es  $\mathcal{D}_{\text{pLC}}^{\text{LIpre}}$ .

Puesto que no tiene antecedentes, no es necesario demostrar el crecimiento de la función.

- $\xi = [M^\gamma; O^\gamma \rightarrow P^\gamma \delta N^\gamma \bullet \nu, i, j + 1 \mid p, q]$  con  $\text{label}(N^\gamma) = a$ .

De la viabilidad de  $\xi$  tenemos que:

(i)  $M^\gamma >_\ell^\lambda O^\gamma$

(ii.i) si  $(p, q) = (-, -)$

$$P^\gamma \delta N^\gamma \stackrel{\mu-1}{\Rightarrow} a_{i+1} \dots a_j N^\gamma \Rightarrow a_{i+1} \dots a_{j+1};$$

(ii.ii) si  $(p, q) \neq (-, -)$

$$P^\gamma \delta N^\gamma \xrightarrow{\mu'-1} a_{i+1} \dots a_p \perp a_{q+1} \dots a_j N^\gamma \xrightarrow{\theta}_f a_{i+1} \dots a_j N^\gamma \Rightarrow a_{i+1} \dots a_{j+1},$$

donde  $\mu = \mu' + \theta$ .

De (i)-(ii) se obtiene que:

$$\eta = [M^\gamma; O^\gamma \rightarrow P^\gamma \delta \bullet N^\gamma \nu, i, j \mid p, q] \in \Delta_{\text{pLC}}.$$

Y de la condición de partida que:

$$\zeta = [a, j, j+1] \in \mathcal{H}_{\text{pLC}}.$$

El paso deductivo  $\{\eta, \zeta\} \vdash \xi$  existe y es  $\mathcal{D}_{\text{pLC}}^{\text{Scan}}$ .

Por definición de  $d$ :

$$d(\xi) = d(\eta) + 2.$$

$$d(\zeta) = 0.$$

Por lo que se cumple  $d(\xi) > d(\zeta)$  y  $d(\xi) > d(\eta)$ .

- $\xi = [M^\gamma; O^\gamma \rightarrow P^\gamma \delta N^\gamma \bullet \nu, i, j \mid p, q]$  con  $\text{label}(N^\gamma) = \epsilon$ .

De la viabilidad de  $\xi$  tenemos que:

(i)  $M^\gamma >_\ell^\lambda O^\gamma$ ,

(ii.i) si  $(p, q) = (-, -)$

$$P^\gamma \delta N^\gamma \xrightarrow{\mu'-1} a_{i+1} \dots a_j N^\gamma \Rightarrow a_{i+1} \dots a_j;$$

(ii.ii) si  $(p, q) \neq (-, -)$

$$P^\gamma \delta N^\gamma \xrightarrow{\mu'-1} a_{i+1} \dots a_p \perp a_{q+1} \dots a_j N^\gamma \xrightarrow{\theta}_f a_{i+1} \dots a_j N^\gamma \Rightarrow a_{i+1} \dots a_j,$$

donde  $\mu = \mu' + \theta$ .

De (i)-(ii) se obtiene que:

$$\eta = [M^\gamma; O^\gamma \rightarrow P^\gamma \delta \bullet N^\gamma \nu, i, j \mid p, q] \in \Delta_{\text{pLC}}.$$

El paso deductivo  $\eta \vdash \xi$  existe y es  $\mathcal{D}_{\text{pLC}}^\epsilon$ .

Por definición de  $d$  y considerando el reconocimiento de  $\epsilon$  como el de un símbolo de la cadena de entrada, tenemos que:

$$d(\xi) = d(\eta) + 2.$$

Por lo que se cumple  $d(\xi) > d(\eta)$ .

- $\xi = [M^\gamma; O^\gamma \rightarrow P^\gamma \bullet \nu, j, k \mid p, q]$  con  $P^\gamma \in V_N$ ,  $\text{label}(P^\gamma) \neq \perp$  y  $O^\gamma >_\ell^* P^\gamma$ .

De la viabilidad de  $\xi$  tenemos que:

(i)  $M^\gamma >_\ell^\lambda O^\gamma$ ,

(ii.i) si  $(p, q) = (-, -)$

$$P^\gamma \Rightarrow \delta \xrightarrow{\mu'-1} a_{j+1} \dots a_k,$$

donde  $P^\gamma \rightarrow \delta \in \mathcal{P}(\gamma)$ ;

(ii.ii) si  $(p, q) \neq (-, -)$

$$P^\gamma \Rightarrow \delta \xrightarrow{\mu'-1} a_{j+1} \dots a_p \perp a_{q+1} \dots a_k \xrightarrow{\theta}_f a_{j+1} \dots a_k,$$

donde  $\mu = \mu' + \theta$  y  $P^\gamma \rightarrow \delta \in \mathcal{P}(\gamma)$ .

De (i)-(ii) y de la condición  $O^\gamma >_\ell^* P^\gamma$  se obtiene que:

$$\eta = [M^\gamma; P^\gamma \rightarrow \delta \bullet, j, k \mid p, q] \in \Delta_{\text{pLC}}.$$

De la viabilidad de  $\eta$  tenemos que:

$$M^\gamma >_{\ell}^{\lambda+1} P^\gamma,$$

El paso deductivo  $\eta \vdash \xi$  existe y es  $\mathcal{D}_{\text{pLC}}^{\text{LC}_n}$ .

Por definición de  $d$ :

$$d(\xi) = \min\{\tau + \lambda + 2\mu\},$$

$$d(\eta) = \min\{\tau + (\lambda + 1) + 2(\mu - 1)\},$$

Por tanto,  $d(\eta) = d(\xi) - 1$ .

Hemos denominado  $\tau$  al tamaño del recorrido hasta el nodo ancla que, como hemos visto en los casos anteriores, dependerá del tipo de ítem left corner (véase condición (ii) de la definición de  $\Delta_{\text{pLC}}$ ). No distinguimos entre los cuatro casos porque no es de interés, ya que se trata del mismo valor en  $\eta$  y en  $\xi$ .

Por lo que se cumple  $d(\xi) > d(\eta)$ .

- $\xi = [C^\gamma; O^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k \mid p, q]$  con  $M^\gamma \in V_N$  y  $\text{nil} \in \text{adj}(M^\gamma)$ .

De la viabilidad de  $\xi$  tenemos que:

(i)  $C^\gamma >_{\ell}^{\lambda} O^\gamma$ ,

(ii.i) si  $(p, q) = (-, -)$

$$\delta M^\gamma \xrightarrow{\mu} a_{i+1} \dots a_j M^\gamma \Rightarrow a_{i+1} \dots a_j \alpha \xrightarrow{\zeta} a_{i+1} \dots a_k,$$

donde  $M^\gamma \rightarrow \alpha \in \mathcal{P}(\gamma)$ ;

(ii.ii.i) si  $(p, q) \neq (-, -)$  y  $M^\gamma$  no pertenece a la espina de  $\gamma$

$$\delta M^\gamma \xrightarrow{\mu'} a_{i+1} \dots a_p \perp a_{q+1} \dots a_j M^\gamma \Rightarrow a_{i+1} \dots a_p \perp a_{q+1} \dots a_j \alpha \xrightarrow{\zeta}$$

$$a_{i+1} \dots a_p \perp a_{q+1} \dots a_k \xrightarrow{\theta} a_{i+1} \dots a_k,$$

donde  $\mu = \mu' + \theta$  y  $M^\gamma \rightarrow \alpha \in \mathcal{P}(\gamma)$ ;

(ii.ii.ii) si  $(p, q) \neq (-, -)$  y  $M^\gamma$  pertenece a la espina de  $\gamma$

$$\delta M^\gamma \xrightarrow{\mu} a_{i+1} \dots a_j M^\gamma \Rightarrow a_{i+1} \dots a_j \alpha \xrightarrow{\zeta'}$$

$$a_{i+1} \dots a_p \perp a_{q+1} \dots a_k \xrightarrow{\theta} a_{i+1} \dots a_k,$$

donde  $\iota = \iota' + \theta$  y  $M^\gamma \rightarrow \alpha \in \mathcal{P}(\gamma)$ .

De (i)-(ii) se obtiene que:

$$\eta = [C^\gamma; O^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p', q'] \in \Delta_{\text{pLC}}.$$

De la viabilidad de  $\eta$  se tiene que:

(iii)  $[M^\gamma, j] \in \Omega_{\text{pLC}}$ .

De (ii)-(iii) obtenemos:

$$\zeta = [M^\gamma; M^\gamma \rightarrow \alpha \bullet, j, k \mid p'', q''] \in \Delta_{\text{pLC}}.$$

Hay que tener en cuenta que:

si  $(p', q') = (p'', q'') = (-, -)$  entonces  $(p, q) = (-, -)$ ;

si  $(p', q') \neq (-, -)$  entonces  $(p'', q'') = (-, -)$ , y  $(p, q) = (p', q')$ ;

si  $(p'', q'') \neq (-, -)$  entonces  $(p', q') = (-, -)$ , y  $(p, q) = (p'', q'')$ .

El paso deductivo  $\{\eta, \zeta\} \vdash \xi$  existe y es  $\mathcal{D}_{\text{pLC}}^{\text{Comp}}$ .

Por definición de  $d$ :

$$d(\xi) = \min\{\pi + \lambda + 2(\mu + \iota + 1)\}$$

$$d(\eta) = \min\{\pi + \lambda + 2\mu\}$$

$$d(\zeta) = \min\{\pi' + 2\iota\}$$

Para evitar tener que considerar los distintos casos de nodos anclas, hemos denominado  $\pi$  al número de arcos atravesados hasta el nodo  $C^\gamma$  y  $\pi'$  al número de arcos atravesados hasta el nodo  $M^\gamma$ .

Por lo que se cumple  $d(\xi) > d(\eta)$ .

Por otro lado, como  $\pi' = d(\eta) + 1$ , entonces  $d(\zeta) = \min\{d(\eta) + 1 + 2\iota\} = \min\{\pi + \lambda + 2(\mu + \iota) + 1\} = d(\xi) - 1$ .

Por lo que se cumple  $d(\xi) > d(\zeta)$ .

- $\xi = [E^\gamma; \mathbf{F}^\gamma \rightarrow \perp \bullet, p, q \mid p, q]$ .

De la viabilidad de  $\xi$  tenemos que:

$$(i) E^\gamma >_i^\lambda \mathbf{F}^\gamma,$$

$$(ii) \perp \Rightarrow_f \nu \xrightarrow{\mu-1} f a_{p+1} \dots a_q,$$

donde  $N^\varphi \rightarrow \nu$  y  $N^\varphi \in \text{adj}(\gamma)$ .

Tal como hemos definido las TAG, ni el nodo bottom ni el nodo pie son nodos adjuntables. Ésto, junto a las condiciones (i)-(ii), nos lleva a que:

$$\zeta = [E^\gamma; \mathbf{F}^\gamma \rightarrow \bullet \perp, p, p \mid -, -] \in \Delta_{\text{pLC}}$$

De la viabilidad de  $\zeta$ , la condición (ii) y la definición de ítem left corner viable, se implica que:

$$\eta = [N^\varphi; N^\varphi \rightarrow \nu \bullet, p, q \mid k, l] \in \Delta_{\text{pLC}}.$$

El paso deductivo  $\{\eta, \zeta\} \vdash \xi$  existe y es  $\mathcal{D}_{\text{pLC}}^{\text{FootComp}}$ .

Por definición de  $d$ :

$$d(\xi) = \min\{d(\zeta) + 2\mu\}$$

$$d(\eta) = \min\{d(\zeta) + 2(\mu - 1)\}$$

Por lo que se cumple  $d(\xi) > d(\zeta)$  y  $d(\xi) > d(\eta)$ .

- $\xi = [C^\gamma; N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, m \mid p, q]$  con  $\beta \in \text{adj}(M^\gamma)$ .

De la viabilidad de  $\xi$  tenemos que:

$$(i) C^\gamma >_i^\lambda N^\gamma,$$

Teniendo en cuenta que sobre  $M^\gamma$  se ha completado una adjunción ( $\beta \in \text{adj}(M^\gamma)$ ), tenemos que:

$$(ii.i) \text{ si } (p, q) = (-, -)$$

$$\delta M^\gamma \xrightarrow{\mu} a_{i+1} \dots a_j M^\gamma \Rightarrow_r a_{i+1} \dots a_j \top \Rightarrow_r a_{i+1} \dots a_j \mathbf{R}^\beta \xrightarrow{\kappa}_r$$

$$a_{i+1} \dots a_k \mathbf{F}^\beta a_{l+1} \dots a_m \Rightarrow_r a_{i+1} \dots a_k \perp a_{l+1} \dots a_m \Rightarrow_f$$

$$a_{i+1} \dots a_k \omega a_{l+1} \dots a_m \xrightarrow{\lambda}_f a_{i+1} \dots a_m,$$

donde  $M^\gamma \rightarrow \omega \in \mathcal{P}(\gamma)$ ;

$$(ii.ii.i) \text{ si } (p, q) \neq (-, -) \text{ y } M^\gamma \text{ no pertenece a la espina de } \gamma$$

$$\delta M^\gamma \xrightarrow{\mu} a_{i+1} \dots a_p \perp a_{q+1} \dots a_j M^\gamma \xrightarrow{\theta}_f a_{i+1} \dots a_j M^\gamma \Rightarrow_r$$

$$a_{i+1} \dots a_j \top \Rightarrow_r a_{i+1} \dots a_j \mathbf{R}^\beta \xrightarrow{\kappa}_r a_{i+1} \dots a_k \mathbf{F}^\beta a_{l+1} \dots a_m \Rightarrow_r$$

$$a_{i+1} \dots a_k \perp a_{l+1} \dots a_m \Rightarrow_f a_{i+1} \dots a_k \omega a_{l+1} \dots a_m \xrightarrow{\lambda}_f a_{i+1} \dots a_m,$$



donde  $\mu = \mu' + \theta$  y  $M^\gamma \rightarrow \omega \in \mathcal{P}(\gamma)$ ;

(ii.ii.ii) si  $(p, q) \neq (-, -)$  y  $M^\gamma$  pertenece a la espina de  $\gamma$

$$\begin{aligned} \delta M^\gamma &\xrightarrow{\kappa} a_{i+1} \dots a_j M^\gamma \Rightarrow_r a_{i+1} \dots a_j \top \Rightarrow_r a_{i+1} \dots a_j \mathbf{R}^\beta \xrightarrow{\kappa} a_{i+1} \dots a_k \mathbf{F}^\beta a_{l+1} \dots a_m \Rightarrow_r a_{i+1} \dots a_k \perp a_{l+1} \dots a_m \Rightarrow_f \\ &a_{i+1} \dots a_k \omega a_{l+1} \dots a_m \xrightarrow{\iota'} a_{i+1} \dots a_p \perp a_{q+1} \dots a_m \xrightarrow{\theta} a_{i+1} \dots a_m, \end{aligned}$$

donde  $\iota = \iota' + \theta$  y  $M^\gamma \rightarrow \omega \in \mathcal{P}(\gamma)$ .

De (i)-(ii) se obtiene que:

$$(iii) \eta = [C^\gamma; N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j \mid p', q'] \in \Delta_{\text{pLC}}.$$

De (iii) tenemos que:

$$(iv) \sigma = [M^\gamma, j] \in \Omega_{\text{pLC}}.$$

De (ii), (iv) y de la condición  $\beta \in \text{adj}(M^\gamma)$  obtenemos (por el caso 2 de la definición de ítem predictivo viable) que:

$$(v) \zeta = [\top; \top \rightarrow \mathbf{R}^\beta \bullet, j, m \mid k, l] \in \Delta_{\text{pLC}}.$$

De (ii) y (v) obtenemos la viabilidad de:

$$(vi) \tau = [E^\beta; \mathbf{F}^\beta \rightarrow \bullet \perp, k, k \mid -, -] \in \Delta_{\text{pLC}}.$$

De (ii), (vi) y la condición de partida, tenemos que:

$$\psi = [M^\gamma; M^\gamma \rightarrow \omega \bullet, k, l \mid p'', q''] \in \Delta_{\text{pLC}}.$$

De nuevo hay que tener en cuenta que:

si  $(p', q') = (p'', q'') = (-, -)$  entonces  $(p, q) = (-, -)$ ;

si  $(p', q') \neq (-, -)$  entonces  $(p'', q'') = (-, -)$ , y  $(p, q) = (p', q')$ ;

si  $(p'', q'') \neq (-, -)$  entonces  $(p', q') = (-, -)$ , y  $(p, q) = (p'', q'')$ .

El paso deductivo  $\{\eta, \zeta, \psi\} \vdash \xi$  existe y es  $\mathcal{D}_{\text{pLC}}^{\text{AdjComp}}$ .

Por definición de  $d$  tenemos que:

$$d(\xi) = \min\{\pi + \lambda + 2(\mu + \kappa + \iota + 4)\}$$

$$d(\eta) = \min\{\pi + \lambda + 2\mu\}$$

$$d(\zeta) = \min\{\pi' + 2(\kappa + \iota + 2)\}$$

Para evitar tener que considerar los distintos casos de nodos anclas, hemos denominado  $\pi$  al número de arcos atravesados hasta el nodo  $C^\gamma$  y  $\pi' = d(\sigma)$ .

Por otro lado, también tenemos que:

$$d(\sigma) = d(\eta) + 1,$$

por lo cual,

$$d(\zeta) = \min\{d(\eta) + 1 + 2(\kappa + \iota + 2) = \min\{\pi + \lambda + 2(\mu + \kappa + \iota + 2) + 1\}.$$

Se cumple que  $d(\xi) > d(\eta)$  y  $d(\xi) > d(\zeta)$ .

Si hacemos el siguiente refinamiento en la condición (ii):

$$\begin{aligned} \dots \Rightarrow_r a_{i+1} \dots a_j \top \Rightarrow_r a_{i+1} \dots a_j \mathbf{R}^\beta \xrightarrow{\kappa'} a_{i+1} \dots a_k \mathbf{E}^\beta \xrightarrow{\lambda'} \\ a_{i+1} \dots a_k \mathbf{F}^\beta \xrightarrow{\kappa''} a_{i+1} \dots a_k \mathbf{F}^\beta a_{l+1} \dots a_m \Rightarrow_f \dots, \end{aligned}$$

con  $\mathbf{E}^\beta >_{\lambda'} \mathbf{F}^\beta$ ,  $V^\beta \rightarrow \omega' \mathbf{E}^\beta \omega'' \in \mathcal{P}(\beta)$  y  $\kappa = \kappa' + \lambda' + \kappa''$ ;

se puede establecer que:

$$d(\tau) = \min\{\pi' + \lambda' + 2\kappa'\}.$$

Por definición de  $d$ :

$$d(\psi) = \min\{d(\tau) + 1 + 2\iota\}.$$

Combinando ambos resultados, obtenemos que:

$$d(\psi) = \min\{\pi' + \lambda' + 2(\kappa' + \iota)\} = \min\{d(\eta) + 1 + \lambda' + 2(\kappa' + \iota)\} = \min\{\pi + \lambda + \lambda' + 2(\mu + \kappa' + \iota) + 1\}.$$

Puesto que  $\kappa \geq \kappa'$  y  $\kappa > \lambda'$ , se cumple  $d(\xi) > d(\psi)$ .

**Teorema A.2** *Corrección del analizador pLC*

Para cualquier gramática  $G \in TAG$  y una cadena de entrada  $a_1 \dots a_n$  se cumple que:

$$[\top; \top \rightarrow \mathbf{R}^{\alpha \bullet}, 0, n \mid -, -] \in \mathcal{V}_{pLC} \text{ si y solo si } \top \Rightarrow \mathbf{R}^{\alpha} \stackrel{*}{\Rightarrow} a_1 \dots a_n \wedge \alpha \in I$$

*Prueba*

Directamente de los lemas A.3 y A.4 y el corolario A.2.

### A.3 Prueba de corrección del esquema sLC

**Teorema A.3** *Corrección del analizador sLC*

Para cualquier gramática  $G \in TAG$  y una cadena de entrada  $a_1 \dots a_n$  se cumple que:

$$[\top \rightarrow \mathbf{R}^{\alpha \bullet}, 0, n \mid -, -] \in \mathcal{V}_{LC} \text{ si y solo si } \top \Rightarrow \mathbf{R}^{\alpha} \stackrel{*}{\Rightarrow} a_1 \dots a_n \wedge \alpha \in I$$

*Prueba*

Se deduce de la relación que mantiene con el esquema pLC, y que es la siguiente:

$$sLC \xrightarrow{\text{ir}} pLC$$

Por el corolario 2.1, la corrección de pLC implica la corrección de sLC.

## Apéndice B

# Adaptando los esquemas para TIG

El objetivo de este apéndice es redefinir algunos de los esquemas para TIG que introducimos en la memoria con objeto de facilitar su integración con sus homónimos para TAG y crear esquemas combinados eficientes, tal como se detalla en el capítulo 10. Básicamente las modificaciones que se aplicarán sobre los esquemas originales son dos:

1. Sólo se permitirá una adjunción, izquierda o derecha, sobre un nodo. Para garantizar esta restricción en los esquemas  $\text{CYK}^i$ ,  $\text{buE}^i$  y  $\text{dVH}^i$ , el parámetro  $code \subseteq \{L, R\}$  que existía en los ítems será sustituido por un parámetro booleano  $adj \in \{true, false\}$  que indique si se ha efectuado una adjunción en el nodo que domina la regla de producción del ítem. Donde los cambios son más significativos es en el esquema  $\text{E}^i$ , debido a que en él se permiten adjunciones múltiples. En este caso es necesario añadir el componente booleano y algunos cambios colaterales.
2. Los árboles auxiliares izquierdos y derechos ( $\mathbf{A}_L$  y  $\mathbf{A}_R$ ) se redennominan como árboles auxiliares fuertemente izquierdos y derechos ( $\mathbf{A}_{SL}$  y  $\mathbf{A}_{SR}$ ), respectivamente. Puesto que en TIG no existen árboles auxiliares wrapping, esta modificación es simplemente un cambio de notación.

No vamos a hacer una descripción detallada de los esquemas que se definen aquí, ya que no presentan un relevante interés práctico por sí solos. Y usaremos el superíndice  $j$  junto al nombre como nota distintiva de este conjunto de esquemas auxiliares.

### B.1 Esquema tipo CYK

El dominio del esquema  $\text{CYK}^j$  viene dado por:

$$\mathcal{I}_{\text{CYK}^j} = \{[M^\gamma, i, j, adj]\}$$

donde  $M^\gamma \in V_N$ ,  $\gamma \in I \cup \mathbf{A}_{SL} \cup \mathbf{A}_{SR}$ ,  $0 \leq i \leq j$  y el parámetro booleano  $adj$  toma el valor *true* cuando se ha completado una adjunción izquierda o derecha sobre el nodo  $M^\gamma$  y *false* en caso contrario.

Los pasos deductivos del esquema son:

$$\mathcal{D}_{\text{CYK}^j} = \mathcal{D}_{\text{CYK}^j}^{\text{Scan}} \cup \mathcal{D}_{\text{CYK}^j}^\varepsilon \cup \mathcal{D}_{\text{CYK}^j}^{\text{CompUna}} \cup \mathcal{D}_{\text{CYK}^j}^{\text{CompBin}} \cup \mathcal{D}_{\text{CYK}^j}^{\text{LAdj}} \cup \mathcal{D}_{\text{CYK}^j}^{\text{RAAdj}} \cup \mathcal{D}_{\text{CYK}^j}^{\text{Subs}}$$

$$\mathcal{D}_{\text{CYK}^j}^{\text{Scan}} = \frac{[a, j, j + 1]}{[N^\gamma, j, j + 1, \text{false}]} \quad N^\gamma \rightarrow a \in \mathcal{P}(\gamma)$$

$$\mathcal{D}_{\text{CYK}^j}^\varepsilon = \frac{}{[N^\gamma, j, j, \text{false}]} \quad \begin{array}{l} N^\gamma \rightarrow \varepsilon \in \mathcal{P}(\gamma) \\ \text{ó } N^\gamma \rightarrow \perp \end{array}$$

$$\mathcal{D}_{\text{CYK}^j}^{\text{CompUna}} = \frac{[O^\gamma, i, j, \text{adj}]}{[M^\gamma, i, j, \text{false}]} \quad M^\gamma \rightarrow O^\gamma \in \mathcal{P}(\gamma)$$

donde se debe cumplir:

( $\mathbf{nil} \in \text{adj}(O^\gamma)$  y  $adj = \text{false}$ ) o ( $\exists \beta \in \text{adj}(O^\gamma)$  y  $adj \neq \text{false}$ ).

$$\mathcal{D}_{\text{CYK}^j}^{\text{CompBin}} = \frac{[O_1^\gamma, i, j, \text{adj}] \quad [O_2^\gamma, j, k, \text{adj}']}{[M^\gamma, i, k, \text{false}]} \quad M^\gamma \rightarrow O_1^\gamma O_2^\gamma \in \mathcal{P}(\gamma)$$

donde se debe cumplir:

(i) ( $\mathbf{nil} \in \text{adj}(O_1^\gamma)$  y  $adj = \text{false}$ ) o ( $\exists \beta \in \text{adj}(O_1^\gamma)$  y  $adj \neq \text{false}$ ),

(ii) ( $\mathbf{nil} \in \text{adj}(O_2^\gamma)$  y  $adj = \text{false}$ ) o ( $\exists \beta \in \text{adj}(O_2^\gamma)$  y  $adj \neq \text{false}$ ).

$$\mathcal{D}_{\text{CYK}^j}^{\text{LAdj}} = \frac{[P^\beta, i, j, \text{false}] \quad [M^\gamma, j, k, \text{false}]}{[M^\gamma, i, k, \text{true}]} \quad \begin{array}{l} \text{label}(P^\beta) = \top \\ \beta \in \text{ladj}(M^\gamma) \end{array}$$

$$\mathcal{D}_{\text{CYK}^j}^{\text{RAAdj}} = \frac{[P^\beta, j, k, \text{false}] \quad [M^\gamma, i, j, \text{false}]}{[M^\gamma, i, k, \text{true}]} \quad \begin{array}{l} \text{label}(P^\beta) = \top \\ \beta \in \text{radj}(M^\gamma) \end{array}$$

$$\mathcal{D}_{\text{CYK}^j}^{\text{Subs}} = \frac{[P^\alpha, i, j, \text{false}]}{[M^\gamma, i, j, \text{false}]} \quad \begin{array}{l} \text{label}(P^\alpha) = \top \\ \alpha \in \text{subst}(M^\gamma) \end{array}$$

El conjunto de ítems finales se define mediante:

$$\mathcal{F}_{\text{CYK}^j} = \{[P^\alpha, 0, n, \text{false}] \mid \alpha \in I, \text{label}(P^\alpha) = \top, \text{label}(\mathbf{R}^\alpha) = S\}$$

## B.2 Esquema tipo Earley ascendente

El dominio del esquema  $\text{buE}^j$  es:

$$\mathcal{I}_{\text{buE}^j} = \mathcal{I}_{\text{buE}^j}^{(i)} \cup \mathcal{I}_{\text{buE}^j}^{(ii)}$$

$$\mathcal{I}_{\text{buE}^j}^{(i)} = \{[M^\gamma \rightarrow \delta \bullet \nu, i, j, \text{false}]\}$$

donde  $M^\gamma \rightarrow \delta \nu \in \mathcal{P}(\gamma)$ ,  $\gamma \in I \cup \mathbf{A}_{SL} \cup \mathbf{A}_{SR}$ ,  $\nu \neq \varepsilon$ ,  $0 \leq i \leq j$  y el parámetro *false* indica que no se ha completado ninguna adjunción sobre el nodo  $M^\gamma$ .

$$\mathcal{I}_{\text{buE}^j}^{(ii)} = \{[M^\gamma \rightarrow \nu \bullet, i, j, \text{adj}]\}$$

donde  $M^\gamma \rightarrow \nu \in \mathcal{P}(\gamma)$ ,  $\gamma \in I \cup \mathbf{A}_{SL} \cup \mathbf{A}_{SR}$ ,  $0 \leq i \leq j$  y el parámetro booleano *adj* toma el valor *true* cuando se ha completado una adjunción izquierda o derecha sobre el nodo  $M^\gamma$  y *false* en caso contrario.

Los pasos deductivos del esquema son:

$$\mathcal{D}_{\text{buE}^j} = \mathcal{D}_{\text{buE}^j}^{\text{Ini}} \cup \mathcal{D}_{\text{buE}^j}^{\text{Scan}} \cup \mathcal{D}_{\text{buE}^j}^\varepsilon \cup \mathcal{D}_{\text{buE}^j}^{\text{Comp}} \cup \mathcal{D}_{\text{buE}^j}^{\text{LAdj}} \cup \mathcal{D}_{\text{buE}^j}^{\text{RAdj}} \cup \mathcal{D}_{\text{buE}^j}^{\text{Subs}}$$

$$\mathcal{D}_{\text{buE}^j}^{\text{Ini}} = \frac{[N^\gamma \rightarrow \bullet \nu, i, i, \text{false}]}{[N^\gamma \rightarrow \bullet \nu, i, i, \text{false}]} \quad \gamma \in I \cup \mathbf{A}_{SL} \cup \mathbf{A}_{SR}$$

$$\mathcal{D}_{\text{buE}^j}^{\text{Scan}} = \frac{\begin{array}{c} [a, j, j+1] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{false}] \end{array}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, j+1, \text{false}]} \quad \text{label}(M^\gamma) = a$$

$$\mathcal{D}_{\text{buE}^j}^\varepsilon = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{false}]}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, j, \text{false}]} \quad \begin{array}{l} \text{label}(M^\gamma) = \varepsilon \\ \text{ó label}(M^\gamma) = \perp \end{array}$$

$$\mathcal{D}_{\text{buE}^j}^{\text{Comp}} = \frac{\begin{array}{c} [M^\gamma \rightarrow \omega \bullet, j, k, \text{adj}] \\ [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{false}] \end{array}}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k, \text{false}]}$$

donde se debe cumplir:

( $\text{nil} \in \text{adj}(M^\gamma)$  y  $\text{adj} = \text{false}$ ) o ( $\exists \beta \in \text{adj}(M^\gamma)$  y  $\text{adj} \neq \text{false}$ ).

$$\mathcal{D}_{\text{buE}^j}^{\text{LAdj}} = \frac{\begin{array}{c} [\top \rightarrow \mathbf{R}^\beta \bullet, i, j, \text{false}] \\ [M^\gamma \rightarrow \nu \bullet, j, k, \text{false}] \end{array}}{[M^\gamma \rightarrow \nu \bullet, i, k, \text{true}]} \quad \beta \in \text{ladj}(M^\gamma)$$

$$\mathcal{D}_{\text{buE}^j}^{\text{RAdj}} = \frac{\begin{array}{c} [\top \rightarrow \mathbf{R}^\beta \bullet, j, k, \text{false}] \\ [M^\gamma \rightarrow \nu \bullet, i, j, \text{false}] \end{array}}{[M^\gamma \rightarrow \nu \bullet, i, k, \text{true}]} \quad \beta \in \text{radj}(M^\gamma)$$

$$\mathcal{D}_{\text{buEj}}^{\text{Subs}} = \frac{[\top \rightarrow \mathbf{R}^{\alpha \bullet}, j, k, false] [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, false]}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k, false]} \quad \alpha \in \text{subst}(M^\gamma)$$

El conjunto de ítems finales es:

$$\mathcal{F}_{\text{buEj}} = \{[\top \rightarrow \mathbf{R}^{\alpha \bullet}, 0, n, false] \mid \alpha \in I, \text{label}(\mathbf{R}^\alpha) = S\}$$

### B.3 Esquema tipo de Vreught y Honig

El dominio del esquema  $\text{dVH}^j$  se define mediante ítems de la forma:

$$\mathcal{I}_{\text{dVH}^j} = \mathcal{I}_{\text{dVH}^j}^{(i)} \cup \mathcal{I}_{\text{dVH}^j}^{(ii)}$$

$$\mathcal{I}_{\text{dVH}^j}^{(i)} = \{[M^\gamma \rightarrow \nu \bullet \delta \bullet \omega, i, j, false]\}$$

donde  $M^\gamma \rightarrow \nu \delta \omega \in \mathcal{P}(\gamma), \gamma \in I \cup \mathbf{A}_{SL} \cup \mathbf{A}_{SR}, 0 \leq i \leq j, \nu \neq \varepsilon \text{ o } \omega \neq \varepsilon$  y el componente *false* indica que no se ha completado ninguna adjunción en el nodo  $M^\gamma$ .

$$\mathcal{I}_{\text{dVH}^j}^{(ii)} = \{[M^\gamma \rightarrow \bullet \delta \bullet, i, j, adj]\}$$

donde  $M^\gamma \rightarrow \nu \in \mathcal{P}(\gamma), \gamma \in I \cup \mathbf{A}_{SL} \cup \mathbf{A}_{SR}, 0 \leq i \leq j$  y el parámetro booleano *adj* toma el valor *true* cuando se ha completado una adjunción izquierda o derecha sobre el nodo  $M^\gamma$  y *false* en caso contrario.

El conjunto de pasos deductivos viene dado por:

$$\mathcal{D}_{\text{dVH}^j} = \mathcal{D}_{\text{dVH}^j}^{\text{Scan}} \cup \mathcal{D}_{\text{dVH}^j}^\varepsilon \cup \mathcal{D}_{\text{dVH}^j}^{\text{Comp}} \cup \mathcal{D}_{\text{dVH}^j}^{\text{Con}} \cup \mathcal{D}_{\text{dVH}^j}^{\text{LAdj}} \cup \mathcal{D}_{\text{dVH}^j}^{\text{RAAdj}} \cup \mathcal{D}_{\text{dVH}^j}^{\text{Subs}}$$

$$\mathcal{D}_{\text{dVH}^j}^{\text{Scan}} = \frac{[a, j, j+1]}{[N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, j, j+1, false]} \quad \text{label}(M^\gamma) = a$$

$$\mathcal{D}_{\text{dVH}^j}^\varepsilon = \frac{[N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, j, j, false]}{[N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, j, j, false]} \quad \text{label}(M^\gamma) = \varepsilon$$

$$\mathcal{D}_{\text{dVH}^j}^{\text{Comp}} = \frac{[M^\gamma \rightarrow \bullet \delta \bullet, i, j, adj]}{[N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, i, j, false]}$$

donde se debe cumplir:

( $\text{nil} \in \text{adj}(M^\gamma)$  y  $adj = false$ ) o ( $\exists \beta \in \text{adj}(M^\gamma)$  y  $adj \neq false$ ).

$$\mathcal{D}_{\text{dVH}^j}^{\text{Con}} = \frac{[N^\gamma \rightarrow \nu \bullet \delta \bullet \delta' \omega, i, j', false] [N^\gamma \rightarrow \nu \delta \bullet \delta' \bullet \omega, j', j, false]}{[N^\gamma \rightarrow \nu \bullet \delta \delta' \bullet \omega, i, j, false]}$$

$$\mathcal{D}_{\text{dVHj}}^{\text{LAdj}} = \frac{[\top \rightarrow \bullet \mathbf{R}^\beta \bullet, i, j, false] [M^\gamma \rightarrow \bullet \delta \bullet, j, k, false]}{[M^\gamma \rightarrow \bullet \delta \bullet, i, k, true]} \quad \beta \in \text{ladj}(M^\gamma)$$

$$\mathcal{D}_{\text{dVHj}}^{\text{RAdj}} = \frac{[\top \rightarrow \bullet \mathbf{R}^\beta \bullet, j, k, false] [M^\gamma \rightarrow \bullet \delta \bullet, i, j, false]}{[M^\gamma \rightarrow \bullet \delta \bullet, i, k, true]} \quad \beta \in \text{radj}(M^\gamma)$$

$$\mathcal{D}_{\text{dVHj}}^{\text{Subs}} = \frac{[\top \rightarrow \bullet \mathbf{R}^\alpha \bullet, j, k, false]}{[N^\gamma \rightarrow \nu \bullet M^\gamma \bullet \omega, i, k, false]} \quad \alpha \in \text{subst}(M^\gamma)$$

El conjunto de ítems finales se define mediante:

$$\mathcal{F}_{\text{dVHj}} = \{[\top \rightarrow \bullet \mathbf{R}^\alpha \bullet, 0, n, false] \mid \alpha \in I, \text{label}(\mathbf{R}^\alpha) = S\}$$

## B.4 Esquema tipo Earley

Este esquema es el resultado de la aplicación de las modificaciones indicadas al esquema tipo Earley propuesto originalmente por Schabes y Waters y descrito en forma de esquema de análisis en el capítulo 7. El hecho de añadir un componente para evitar las adjunciones múltiples sobre un nodo nos obliga a introducir un nuevo tipo de ítems en el dominio, quedando de la siguiente forma:

$$\begin{aligned} \mathcal{I}_{\text{Ej}} &= \mathcal{I}_{\text{Ej}}^{(i)} \cup \mathcal{I}_{\text{Ej}}^{(ii)} \\ \mathcal{I}_{\text{Ej}}^{(ii)} &= \mathcal{I}_{\text{buEj}}^{(ii)} \\ \mathcal{I}_{\text{Ej}}^{(i)} &= \{[M^\gamma \rightarrow \delta \bullet \nu, i, j, adj]\} \end{aligned}$$

donde  $M^\gamma \rightarrow \delta \nu \in \mathcal{P}(\gamma)$ ,  $\gamma \in I \cup \mathbf{A}_{SL} \cup \mathbf{A}_{SR}$ ,  $\nu \neq \varepsilon$ ,  $0 \leq i \leq j$  y el parámetro booleano  $adj$  toma el valor *true* cuando se ha completado una adjunción izquierda sobre el nodo  $M^\gamma$  y *false* en caso contrario.

$$\mathcal{I}_{\text{Ej}}^{(ii)} = \mathcal{I}_{\text{buEj}}^{(ii)}$$

Los pasos deductivos del esquema son:

$$\begin{aligned} \mathcal{D}_{\text{Ej}} &= \mathcal{D}_{\text{Ej}}^{\text{Ini}} \cup \mathcal{D}_{\text{Ej}}^{\text{Scan}} \cup \mathcal{D}_{\text{Ej}}^\varepsilon \cup \mathcal{D}_{\text{Ej}}^{\text{Pred}} \cup \mathcal{D}_{\text{Ej}}^{\text{Comp}} \cup \mathcal{D}_{\text{Ej}}^{\text{Foot}} \cup \mathcal{D}_{\text{Ej}}^{\text{LAdjPred}} \cup \\ &\quad \mathcal{D}_{\text{Ej}}^{\text{LAdjComp}} \cup \mathcal{D}_{\text{Ej}}^{\text{RAdjPred}} \cup \mathcal{D}_{\text{Ej}}^{\text{RAdjComp}} \cup \mathcal{D}_{\text{Ej}}^{\text{SubsPred}} \cup \mathcal{D}_{\text{Ej}}^{\text{SubsComp}} \end{aligned}$$

$$\mathcal{D}_{\text{Ej}}^{\text{Ini}} = \frac{}{[\top \rightarrow \bullet \mathbf{R}^\alpha \bullet, 0, 0, false]} \quad \begin{array}{l} \alpha \in I \\ \text{label}(\mathbf{R}^\alpha) = S \end{array}$$

$$\mathcal{D}_{Ej}^{\text{Scan}} = \frac{[a, j, j+1] \quad [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{adj}]}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, j+1, \text{adj}]} \quad \text{label}(M^\gamma) = a$$

$$\mathcal{D}_{Ej}^\varepsilon = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{adj}]}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, j, \text{adj}]} \quad \begin{array}{l} \text{label}(M^\gamma) = \varepsilon \\ \text{ó label}(M^\gamma) = \perp \end{array}$$

$$\mathcal{D}_{Ej}^{\text{Pred}} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{adj}]}{[M^\gamma \rightarrow \bullet \nu, j, j, \text{false}]}$$

donde se debe cumplir:

( $\text{nil} \in \text{adj}(M^\gamma)$ ) o ( $\exists \beta \in \text{adj}(M^\gamma)$ ).

$$\mathcal{D}_{Ej}^{\text{Comp}} = \frac{[M^\gamma \rightarrow \nu \bullet, j, k, \text{adj}] \quad [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{adj}']}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k, \text{adj}']}$$

donde se debe cumplir:

( $\text{nil} \in \text{adj}(M^\gamma)$  y  $\text{adj} = \text{false}$ ) o ( $\exists \beta \in \text{adj}(M^\gamma)$  y  $\text{adj} \neq \text{false}$ ).

$$\mathcal{D}_{Ej}^{\text{LAdjPred}} = \frac{[M^\gamma \rightarrow \bullet \nu, i, j, \text{false}]}{[\top \rightarrow \bullet \mathbf{R}^\beta, j, j, \text{false}]} \quad \beta \in \text{ladj}(M^\gamma)$$

$$\mathcal{D}_{Ej}^{\text{LAdjComp}} = \frac{[\top \rightarrow \mathbf{R}^\beta \bullet, j, k, \text{false}] \quad [M^\gamma \rightarrow \bullet \nu, i, j, \text{false}]}{[M^\gamma \rightarrow \bullet \nu, i, k, \text{true}]} \quad \beta \in \text{ladj}(M^\gamma)$$

$$\mathcal{D}_{Ej}^{\text{RAAdjPred}} = \frac{[M^\gamma \rightarrow \nu \bullet, i, j, \text{false}]}{[\top \rightarrow \bullet \mathbf{R}^\beta, j, j, \text{false}]} \quad \beta \in \text{radj}(M^\gamma)$$

$$\mathcal{D}_{Ej}^{\text{RAAdjComp}} = \frac{[\top \rightarrow \mathbf{R}^\beta \bullet, j, k, \text{false}] \quad [M^\gamma \rightarrow \nu \bullet, i, j, \text{false}]}{[M^\gamma \rightarrow \nu \bullet, i, k, \text{true}]} \quad \beta \in \text{radj}(M^\gamma)$$

$$\mathcal{D}_{Ej}^{\text{SubsPred}} = \frac{[N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{adj}]}{[\top \rightarrow \bullet \mathbf{R}^\alpha, j, j, \text{false}]} \quad \alpha \in \text{subst}(M^\gamma)$$

$$\mathcal{D}_{Ej}^{\text{SubsComp}} = \frac{[\top \rightarrow \mathbf{R}^\alpha \bullet, j, k, \text{false}] \quad [N^\gamma \rightarrow \delta \bullet M^\gamma \nu, i, j, \text{adj}]}{[N^\gamma \rightarrow \delta M^\gamma \bullet \nu, i, k, \text{adj}]} \quad \alpha \in \text{subst}(M^\gamma)$$

El conjunto de ítems finales viene dado por:

$$\mathcal{F}_{Ej} = \{[\top \rightarrow \mathbf{R}^\alpha \bullet, 0, n, \text{false}] \mid \alpha \in I, \text{label}(\mathbf{R}^\alpha) = S\}$$

# Bibliografía

- [1] A. Abeillé y M. Candito. FTAG: a lexicalized tree adjoining grammar for French. *Tree Adjoining Grammars: Formal Properties, Linguistics Theory and Applications*, CSLI, Stanford, USA. (2000)
- [2] A. Agustini y V. L. Strube de Lima. An experiment on synchronous TAGs for the construction of a transfer module. In Proc. of *4th International Workshop on Tree Adjoining Grammars and Related Frameworks (TAG+4)*, pp. 1–4, Philadelphia, PA, USA. (1998)
- [3] A. V. Aho. Indexed Grammars – an extensión of context-free grammars. *Journal of the Association for Computer Machinery*, 15(4), pp 647–671. (1968)
- [4] A. V. Aho y J. Ullman. The theory of parsing, translation and compiling. *Prentice Hall*, Englewood Cliffs, New Jersey, USA. (1972)
- [5] H. Alblas, R. del Akker, P. O. Luttighuis y K. Sikkel. A bibliography on parallel parsing. *SIGPLAN Notices*, 29(1), pp 54–65. (1994)
- [6] M. A. Alonso. Interpretación tabular de autómatas para lenguajes de adjunción de árboles. *Tesis doctoral*, Universidade da Coruña, España. (2000)
- [7] M. A. Alonso, D. Cabrero, E. de la Clergerie y M. Vilares. Algoritmos tabulares para el análisis de TAG. *Procesamiento del lenguaje Natural*, 23, pp 157–164. (1998)
- [8] M. A. Alonso, D. Cabrero, E. de la Clergerie y M. Vilares. Tabular algorithms for TAG parsing. In Proc. of *European Chapter of ACL, (EACL'99)*, pp 150–157, Bergen, Norway. (1999)
- [9] M. A. Alonso, V. Carrillo y V. J. Díaz. Análisis sintáctico combinado de TAGs y TIGs. *Procesamiento del lenguaje Natural*, 29, pp 65–72. (2002)
- [10] M. A. Alonso, V. Carrillo y V. J. Díaz. Mixed Parsing of Tree Insertion and Tree Adjoining Grammars. *Advances in Artificial Intelligence. IBERAMIA 2002, Vol 2527 of Lecture Notes in Artificial Intelligence*, pp 694–703, Springer-Verlag. (2002)
- [11] G. Barton, R. Berwick y E. Ristad. Computational complexity and natural language. *Computational Intelligence*, 10(4), MIT Press, Cambridge, MA. (1987)



- [12] R. E. Bellman. *Dynamic Programming*. Princeton University Press, Princeton, NJ. (1957)
- [13] P. Boullier. On TAG and multicomponent TAG parsing. *Rapport de Recherche num. 3668. INRIA Institute National de Recherche en Informatique et Automatique*, Roquencourt, France. (1999)
- [14] G. Brassard y P. Bratley. *Fundamentos de algoritmia*. Prentice Hall. (1997)
- [15] S. A. Carballo y E. Charniak. New figures of merit for best-first probabilistic chart parsing. *Computational Linguistics* 24, pp 275–298. (1998)
- [16] V. Carrillo. Esquemas basados en left corner sin prefijo válido para TAGs: Relaciones. Informe Interno LSI-2002-03, Departamento de Lenguajes y Sistemas Informáticos, Universidad de Sevilla. (2002)
- [17] V. Carrillo. Algoritmos de análisis para gramáticas de inserción de árboles: Relaciones. Informe Interno LSI-2003-01, Departamento de Lenguajes y Sistemas Informáticos, Universidad de Sevilla. (2003)
- [18] V. Carrillo, V. J. Díaz y A. Gómez. Gramática FTAG del castellano: árboles elementales asociados a los adjetivos. *Procesamiento del lenguaje Natural*, 17, pp 130–141. (1995)
- [19] V. Carrillo, V. J. Díaz y A. Gómez. Adjetivos comparativos en FTAG. *Lenguajes Naturales y lenguajes Formales*, 11, pp 367–372. (1995)
- [20] V. Carrillo, V. J. Díaz y M. Toro. Un método general de transformación de CFG a TAG. *Lenguajes Naturales y lenguajes Formales*, 12, pp 201–208. (1996)
- [21] V. Carrillo, V. J. Díaz y M. Toro. Aspectos formales en la lexicalización de CFGs mediante TAGs. *Actas de la II Jornadas Científicas en la Tecnología de la Información*, pp 121–126. (1996)
- [22] V. Carrillo, V. J. Díaz y M. Toro. Una gramática S-atribuida para TAG. *III Jornadas Científicas en Tecnologías de la Información*, pp 247–254, Cádiz, Spain. (1998)
- [23] V. Carrillo, V. J. Díaz y M. Toro. Un recorrido por los formalismos gramaticales lexicalizados basados en reescritura de árboles. *Novática: Lengua y Tecnología*, 133, pp 22–25, Spain. (1998)
- [24] V. Carrillo, V. J. Díaz y M. A. Alonso. Análisis sintáctico ascendente de TAGs guiado por la esquina izquierda. *Procesamiento del lenguaje Natural*, 27, pp 47–54. (2001)
- [25] V. Carrillo, V. J. Díaz y M. A. Alonso. Algoritmos de análisis para gramáticas de inserción de árboles. *Procesamiento del lenguaje Natural*, 29, pp 89–96. (2002)

- [26] V. Carrillo, V. J. Díaz y M. A. Alonso. A predictive left corner parser for tree adjoining grammars. *Revista electrónica en Tecnologías del Habla*, 1. (2003)
- [27] V. Carrillo, M. A. Alonso y V. J. Díaz. Análisis eficiente de gramáticas de inserción de árboles. *II Jornadas en Tecnologías del Habla*, Granada, España. (2002)
- [28] V. Carrillo y V. J. Díaz. Left corner parser for tree insertion grammars. *Artificial Intelligence: Methodology, Systems, and Applications - AIMSA'02, Vol 2443 of Lecture Notes in Artificial Intelligence, Subseries of Lecture Notes in Computer Science*, pp 142–151, Springer-Verlag. (2002)
- [29] J. Carroll. An island parsing interpreter for the full augmented transition network formalism. In *Proc. of 1st Conference of the European Chapter of the Association for Computational Linguistics*, pp 101–105, Pisa, Italy. (1983)
- [30] C. Chelba y F. Jelinek. Exploiting syntactic structure for language modelling. In *Proc. of 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics (ACL-COLING'98)*, pp 225–231, Montreal, Canada. (1998)
- [31] N. Chomsky. Syntactic structures. *Mouton and Co.*, The Hague, the Netherlands. (1957)
- [32] N. Chomsky. On certain formal properties of grammars. *Information and Control*, 2(2), pp 137–167. (1959)
- [33] N. Chomsky. Aspects of theory of syntax. *MIT Press*, Cambridge, Mass, USA. (1965)
- [34] D. Crystal. Who cares about english usage. Penguin. (1984)
- [35] V. J. Díaz. Gramáticas de adjunción de árboles: Un enfoque deductivo en el análisis sintáctico. *Tesis doctoral*, Universidad de Sevilla, España. (2000)
- [36] V. J. Díaz, M. Toro y V. Carrillo. Un algoritmo de lexicalización de CFG mediante TAG. *Procesamiento del Lenguaje Natural*, 19, pp 201–208. (1996)
- [37] V. J. Díaz, V. Carrillo y M. Toro. A review of Earley-based parser for TIG. *Methodology and Tools in Knowledge-Based Systems, Vol 1415 of Lecture Notes in Artificial Intelligence*, pp 732–738, Springer-Verlag. (1998)
- [38] V. J. Díaz, V. Carrillo y M. Toro. Elementary tree representation. In *Proc. of 1st Workshop on Tabulation in Parsing and Deduction (TAPD'98)*, pp 10–15, Paris, France. (1998)
- [39] V. J. Díaz, V. Carrillo y M. Toro. Análisis sintáctico de TAGs usando analizadores deductivos. *Procesamiento del Lenguaje Natural*, 23, pp 126–131. (1998)

- [40] V. J. Díaz, V. Carrillo y M. Toro. Revisando el reconocedor con prefijo válido para TAGs de Schabes. *Procesamiento del Lenguaje Natural*, 25, pp 59–65. (1999)
- [41] V. J. Díaz, V. Carrillo y M. A. Alonso. A bidirectional bottom-up parser for TAG. In Proc. of *6th International Workshop on Parsing Technologies (IWPT'2000)*, Trento, Italy. (2000)
- [42] V. J. Díaz, M. A. Alonso y V. Carrillo. Bidirectional parsing of TAGs without heads. In Proc. of *5th International Workshop on Tree Adjoining Grammars and Related Frameworks (TAG+5)*, Paris, France. (2000)
- [43] V. J. Díaz, M. A. Alonso y V. Carrillo. A Left Corner parser for TAGs. In Proc. of *6th International Workshop on Tree Adjoining Grammars and Related Frameworks (TAG+6)*, Venice, Italy. (2002)
- [44] V. J. Díaz y M. A. Alonso. Análisis sintáctico bidireccional de TAGs. *Procesamiento del Lenguaje Natural*, 27. (2000)
- [45] V. J. Díaz y M. A. Alonso. Comparing tabular parsers for tree adjoining grammars. In Proc. of *2nd International Workshop on Tabulation in Parsing and Deduction (TAPD'2000)*, Vigo, Spain. (2000)
- [46] V. J. Díaz y M. Toro. Parsing TAGs with Prolog. Proc of *APPIA-GULP-PRODE'97, Joint Conference on Declarative Programming*, pp 359–367, Grado, Italy. (1997)
- [47] C. Doran, D. Egedi, B. A. Hockey, B. Srinivas y M. Zaidel. XTAG system - a wide coverage grammar for english. In Proc. of *17th International Conference on Computational Linguistics (COLING'94)*, pp 922–928, Kyoto, Japan. (1994)
- [48] J. Earley. An efficient context-free parsing algorithm. *PhD Thesis*, Carnegie-Mellon University, Pittsburg, PA. (1968)
- [49] J. Earley. An efficient context-free parsing algorithm. *Communications of the ACM*, 13(2), pp 94-102. (1970)
- [50] R. Evans y D. Weir. A structure-sharing parser for lexicalized grammars. In Proc. of *36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics (ACL-COLING'98, volume I)*, pp 372–378, Montreal, Canada. (1998)
- [51] G. Gazdar. Applicability of indexed grammars to natural languages. *Technical Report CSLI-85-34. Center for the Study of Language and Information*, Stanford University. (1985)
- [52] G. Gazdar. Applicability of indexed grammars to natural languages. *Natural Language Parsing and Linguistic Theories*, pp 69–94, D.Reidel Publishing Company. (1987)

- [53] G. Gazdar, E. Klein, G. Pullum y I. Sag. *Generalized Phrase Structure Grammar*. Blackwell, Oxford, UK. (1985)
- [54] G. Gazdar y C. Mellish. *Natural language processing in PROLOG*. Addison Wesley. (1989)
- [55] S. Görn. Explicit definitions and linguistic dominoes. *Systems and Computer Science*, pp 77–115, University of Toronto Press, Toronto, Canada. (1967)
- [56] S. L. Graham, M. A. Harrison y W. L. Ruzzo. An improved context-free recognizer. *ACM Transactions on Programming Languages and Systems*, 2(3), pp 415–462. (1980)
- [57] Y. Guan. An  $O(n^5)$  recognition algorithm for coupled parenthesis rewriting systems. In *Proc. of 2nd International Workshop on Tree Adjoining Grammars and Related Frameworks (TAG+2)*, Philadelphia, USA. (1992)
- [58] A. Halber. Tree-grammar linear typing for unified super-tagging/probabilistic parsing models. In *Proc. of 4th International Workshop on tree adjoining grammars and related frameworks TAG+4*, pp 54–57, Philadelphia, USA. (1998)
- [59] K. Harbusch y J. Woch. Direct Parsing of Schema-TAGs. In *Proc. of 6th International Workshop on Parsing Technologies (IWPT'2000)*, pp 305–306, Trento, Italy. (2000)
- [60] Z. Harris. String analysis of sentence structure. *Mouton and Co.*, La Haya. (1962)
- [61] J. Honey. The language trap: race, class and the 'standard english'. *Issue in British Schools*, National Council for Academic Standards. (1983)
- [62] M. A. Harrison. Introduction to formal language theory. *Reading*, Addison-Wesley. (1978)
- [63] J. E. Hopcroft y J. D. Ullman. Introduction to automata theory, languages and computation. *Reading* Addison-Wesley. (1979)
- [64] R. Huybregts. The weak inadequacy of context-free phrase structure grammar. *Van Preriferie naar Kern*, pp 81–99, Foris, Dordrecht. (1985)
- [65] P. Jackson e I. Moulinier. *Natural language processing for online applications: text retrieval, extraction and categorization*. John Benjamins, Amsterdam. (2002)
- [66] K. Jensen, G. Heidorn, S. Richardson and N. Haas. PLNLP, PEG and CRITIQUE: three contributions to computing in the humanities. *IBM, Yorktown Heights, NY, Research Report RC 11841*. (1986)
- [67] A. K. Joshi. Tree adjoining grammars: How much context-sensitivity is required to provide reasonable structural descriptions?. In D. R. Dowty, L. Karttunen and A. M. Zwicky, editors, *Natural Language Parsing, Psychological, Computational and Theoretical Perspectives*, chapter 6, pp 206–250. Cambridge University Press. (1985)



- [68] A. K. Joshi. An introduction to tree adjoining grammars. In A. Manaster-Ramer, editor, *Mathematics of Language*, pp 87–115, Amsterdam, Netherlands. John Benjamins Publishing Co. (1987)
- [69] A. K. Joshi, L. S. Levy y M. Takahashi. Tree adjunct grammars. *Journal of Computer and System Sciences*, 10(1), pp 136–163. (1975)
- [70] A. K. Joshi, K. Vijay-Shanker y D. Weir. The convergence of mildly context-sensitive grammar formalisms. In Sells, Shieber, Wasow, editors, *Foundational Issues in Natural Language Processing*, pp 31–82, Cambridge, USA. (1991)
- [71] A. K. Joshi y Y. Schabes. Tree-adjoining grammars. In Grzegorz Rozenberg and Arto Salomaa, editors, *Handbook of Formal Languages. Vol 3: Beyond Words*, chapter 2, pp 69–123. Springer-Verlag, Berlin/Heidelberg/New York. (1997)
- [72] A. K. Joshi y B. Srinivas. Disambiguation of super parts of speech (Supertags): almost parsing. In Proc. of the 17th International Conference on Computational Linguistics (COLING'94), Kyoto, Japan. (1994)
- [73] R. Kaplan y J. Bresnan. Lexical-functional grammars. A formal system for grammatical representation. In J. Bresnan (ed.) *The Mental Representation of Grammatical Relations*, MIT Press, Cambridge, MA. (1983)
- [74] T. Kasami. An efficient recognition and syntax analysis algorithm for context-free languages. *Scientific Report AFCLR-65-758*, Air Force Cambridge Research Laboratory, Bedford, Mass. (1965)
- [75] M. Kay. Algorithm Schemata and Data Structures in syntactic processing. *Report CSL-80-12*, Xerox PARC, Palo Alto, Ca. (1980)
- [76] M. Kay. Parsing in Functional Unification Grammar. *Natural Language Parsing*, Cambridge University Press, Cambridge, UK. (1985)
- [77] M. Kay. Head driven parsing. In Proc. of 1st International Workshop on Parsing Technologies (IWPT'89), pp 52-62, Pittsburgh, PA. (1989)
- [78] A. Kroch. Unbounded dependencies and subjacency in a tree adjoining grammar. *Mathematics of Language*, pp 143–172, John Benjamins Publishing Company, Amsterdam/Philadelphia. (1987)
- [79] A. Kroch y A. K. Joshi. Linguistic relevance of tree adjoining grammars. *Technical Report MS-CIS-85-18*, Department of Computer and Information Science, University of Pennsylvania, PA, USA. (1985)

- [80] P. Lopez. Extended partial parsing for lexicalized grammars. In Proc. of *6th International Workshop on Parsing Technologies (IWPT'2000)*, pp 159–170, Trento, Italy. (2000)
- [81] C. D. Manning y H. Schutze. Foundations of statistical natural language processing. MIT Press, Cambridge, Massachusetts. (1999)
- [82] Y. Matsumoto, H. Tanaka, H. Hirakawa, H. Miyoshi y H. Yasukawa. BUP: a bottom-up parser embedded in Prolog. *New Generation Computing 1*, pp 145–158. (1983)
- [83] M. J. Nederhof. Generalized Left-Corner Parsing. In Proc. of *6th Meeting of the European Chapter of the Association of Computational Linguistics (EACL'93)*, pp 305–314, Utrecht. (1993)
- [84] M. J. Nederhof. Solving the correct-prefix property for TAGs. In Proc. of *5th Meeting on Mathematics of Language (MOL'5)*, pp 124–130, Schloss Dagstuhl, Germany. (1997)
- [85] M. J. Nederhof. An alternative LR algorithm for TAGs. In Proc. of *36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics, (ACL-COLING'98)*, volume II, pp 946–952, Montreal, Canada. (1998)
- [86] M. J. Nederhof. The computational complexity of the correct-prefix property for TAGs. *Computational Linguistics*, 25(3), pp 345–360. (1999)
- [87] M. J. Nederhof y G. Satta. An extended theory of head-driven parsing. In Proc. of *32nd Meeting of the Association of Computational Linguistics (ACL'94)*, pp 210–217, Las Cruces, New Mexico. (1994)
- [88] G. van Noord. Head corner parsing for TAG. *Computational Intelligence*, 10(4), pp 525–534. (1994)
- [89] T. Nurkkala y V. Kumar. A parallel parsing algorithm for natural language using tree adjoining grammars. In Proc. of *8th International Parallel Processing Symposium*, (1994)
- [90] F. C. N. Pereira y D. H. D. Warren. Definite clause grammars for language analysis. A survey of the formalism and a comparison with augmented transition networks. *Artificial Intelligence*, 13, pp 231–278. (1980)
- [91] C. Pollard. Generalized phrase structure grammars, head grammars and natural language. *PhD thesis*, Stanford University. (1984)
- [92] C. Pollard y I. A. Sag. Information-based syntax and semantics: volume 1–fundamentals. University of Chicago Press, Chicago, IL. (1987)



- [93] C. Pollard y I. A. Sag. Head-Driven phrase structure grammars. *Studies in Contemporary Linguistics*, University of Chicago Press, Chicago, IL. (1994)
- [94] P. Poller. Incremental parsing with LD/TLP-TAGs. *Computational Intelligence*, 10(4), pp 549–562. (1994)
- [95] P. Poller y T. Becker. Two-step TAG parsing revisited. In proc. of *4th International Workshop on Tree Adjoining Grammars and Related Frameworks (TAG'4)*, pp 143–146, Philadelphia, USA. (1998)
- [96] V. R. Pratt. LINGOL – A Progress Report. In proc. of *4th International Joint Conference on Artificial Intelligence (IJCAI'75)*, pp 422–428. (1975)
- [97] C. A. Prolo. An efficient LR parser generator for tree adjoining grammars. In Proc. of *6th International Workshop on Parsing Technologies (IWPT'2000)*, pp 207–218, Trento, Italy. (2000)
- [98] S. Rajasekaran y S. Yooseph. TAL recognition in  $\mathcal{O}(m(n^2))$  time. In Proc. of *33rd Annual Meeting of the Association for Computational Linguistics (ACL'95)*, pp 166–173, Cambridge, MA. (1995)
- [99] O. Rambow, K. Vijay-Shanker y D. Weir. D-Tree grammars. In proc. of *33rd Annual Meeting of the Association for Computational Linguistics (ACL'95)*, pp 151–158, Cambridge, USA. (1995)
- [100] K. Roach. Formal properties of head grammars. In A. Manester-Ramer, (ed.) *Mathematics of Language*, John Benjamins, Amsterdam. (1987)
- [101] J. Robinson. DIAGRAM:a grammar for dialogues. *Communications of the ACM*, 25(1), pp 27–47. (1982)
- [102] J. Rogers. Capturing CFLs with tree adjoining grammars. In Proc. of *32nd Annual Meeting of the Association for Computational Linguistics (ACL'94)*, Las Cruces, Mexico. (1994)
- [103] J. Rogers. A unified notion of derived and derivation structures in TAG In Proc. of *5th Meeting on Mathematics of Language (MOL'5)*, Schloss Dagstuhl, Germany. (1997)
- [104] D. J. Rosenkrantz y P. M. Lewis. Deterministic Left Corner Parsing. In Proc. of *11th Annual Symposium on Switching and automa theory*, pp 139–152. (1970)
- [105] G. Satta y W. Schuler. Restrictions on tree adjoining languages. In Proc. of *36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics (ACL-COLING'98)*, volume II, pp 1176–1182, Montreal, Canada. (1998)

- [106] G. Satta y O. Stock. Head-Driven bidirectional parsing: A tabular method. In Proc. of *International Workshop on Parsing Technologies (IWPT'89)*, pp 43–51, Pittsburg, Pa. (1989)
- [107] G. Satta y O. Stock. Bidirectional context-free grammar parsing for natural language processing. *Artificial Intelligence* 69, pp 123–164. (1994)
- [108] Y. Schabes. Mathematical and computational aspects of lexicalized grammars. *PdD thesis*, University of Pennsylvania, PA, USA. (1990)
- [109] Y. Schabes. The valid prefix property and left to right parsing of tree-adjoining grammar. In Proc. of *2nd International Workshop on Parsing Technologies (IWPT'91)*, pp 21–30, Cancun, Mexico. (1991)
- [110] Y. Schabes. Stochastic lexicalized tree adjoining grammars. In Proc. of the *15th International Conference on Computational Linguistics COLING'92*, pp 426–432, Nantes, France. (1992)
- [111] Y. Schabes y S. M. Shieber. An alternative conception of tree-adjoining derivation. *Computational Linguistics*, 20(1), pp 91–124. (1994)
- [112] Y. Schabes y K. Vijay-Shanker. Deterministic left to right parsing of tree adjoining languages. In Proc. of the *28th Annual Meeting of Association for Computational Linguistics (ACL'90)*, pp 276–283, Oittsburgh, Pennsylvania, USA. (1990)
- [113] Y. Schabes y R. C. Waters. Lexicalized context free grammars. In Proc. of the *21st Meeting of the Association for Computational Linguistics (ACL'93)*, pp 121–129, Columbus, OH. (1993)
- [114] Y. Schabes y R. C. Waters. Tree insertion grammar: a cubic-time, parsable formalism that lexicalizes context-free grammar without changing the trees produced. *Computational Linguistics*, 21(4), pp 479–515. (1995)
- [115] Y. Schabes y R. C. Waters. Stochastic lexicalized tree insertion grammars. In Harry Bunt and Masuru Tomita, (eds.), *Recent advances in parsing technologies, vol. 1, Text-speech and Language Technology, chapter 15*, pp 281–294. Kluwer Academic Publishers, Dordrecht/Boston/London. (1996)
- [116] K. Schneider. Algebraic construction of parsing schemata. In Proc. of *6th International Workshop on Parsing Technologies (IWPT'2000)*, pp 242–253, Trento, Italy. (2000)
- [117] S. M. Shieber. The design of a computer language for linguistic information. In Proc. of *10th International Conference on Computational Linguistics (COLING)*, pp 362–366, Stanford, CA. (1984)



- [118] S. M. Shieber. Evidence against the context-freeness of natural language. *Linguistics and philosophy*, 8, pp 333-343. (1985)
- [119] S. M. Shieber. Criteria for designing computer facilities for linguistic analysis. *Linguistics*, 23, pp 189-211. (1985)
- [120] S. M. Shieber. An introduction to unification-based approaches to grammar. *CSLI Lecture Notes 4. Stanford, California: Center for the study of language and information*. (1986)
- [121] S. M. Shieber. Separating linguistics analyses from linguistic theories. *Linguistic Theory and Computer Applications*, pp 1-36, Academic Press, London. (1987)
- [122] S. M. Shieber, Y. Schabes y F. C. N. Pereira. Principles and implementation of deductive parsing. *Journal of Logic Programming* 24(1&2), pp 3-36. (1995)
- [123] S. M. Shieber y Y. Schabes. Synchronous tree-adjoining grammars. In Proc. of the 13th International Conference on Computational Linguistics COLING'90, Helsinki, Finland. (1990)
- [124] S. M. Shieber y Y. Schabes. Generation and synchronous tree-adjoining grammars. *Computational Intelligence*, 7(4), pp 220-228. (1992)
- [125] K. Sikkel. A framework for parsing algorithm specification and analysis. In H. Trost (ed.) 2 Konferenz Verarbeitung Natürlicher Sprache (KONVES'94), pp 300-309, Wien, Informatik Xpress 6, Springer-Verlag, Berlin. (1994)
- [126] K. Sikkel. Parsing schemata and correctness of parsing algorithms. In Proc. of TWLT/AMAST Workshop on Algebraic Methods in Language Processing, pp 83-97, Enschede, The Netherlands. (1995)
- [127] K. Sikkel. Parsing schemata. A framework for specification and analysis of parsing algorithms. *Texts in Theoretical Computer Science*, Springer-Verlag. (1997)
- [128] M. Steedman. Combinators and grammars. In R. Oehrle, E. Bach and D. Wheeler (eds.) *Categorial Grammars and Natural Language Structures*, 5, pp 417-442, Foris, Dordrecht. (1987)
- [129] J. W. Thatcher. Characterizing derivation trees of context free grammars through a generalization of finite automata theory. *Journal of Computer and System Sciences* 5, pp 365-396.
- [130] M. Tomita. An efficient augmented context-free parsing algorithm. *Computational Linguistics* 13(1-2), pp 31-46. (1987)
- [131] M. Tomita. Generalized LR parsing. *Kluwer Academic Publishers*, London. (1991)

- [132] K. Vijay-Shanker. A study of tree adjoining grammars. *PhD thesis*. University of Pennsylvania, PA, USA. (1988)
- [133] K. Vijay-Shanker. The use of domination statements in TAG and related formalisms. *Lenguajes Naturales y Lenguajes Formales*, 12, pp 327–350. (1996)
- [134] K. Vijay-Shanker. D-tree grammars. TAG+4 Tutorial, Philadelphia, PA, USA. (1998)
- [135] K. Vijay-Shanker, D. J. Weir y A.K. Joshi. Tree adjoining and head wrapping. In Proc. of *11th International Conference on Computational Linguistics (COLING'86)*. (1986)
- [136] K. Vijay-Shanker y A. K. Joshi. Some computational properties of tree adjoining grammars. In Proc. of *23rd Annual Meeting of the Association for Computational Linguistics (ACL'85)*, pp 82–93, Chicago, USA. (1985)
- [137] K. Vijay-Shanker y A. K. Joshi. Feature structures based tree adjoining grammars. In Proc. of *12th International Conference on Computational Linguistics (COLING'88)*, vol. 2, pp 714–719, Budapest, Hungary. (1988)
- [138] K. Vijay-Shanker y D. J. Weir. The equivalence of four extensions of context-free grammars. *Mathematical Systems Theory*, 27(4), pp 511–545. (1994)
- [139] J. P. M de Vreught y H. J. Honig. A tabular bottom-up recognizer. *Report 89-78*. Department of Applied Mathematics and Informatics, Delft University of Technologies, Delft, The Netherlands. (1989)
- [140] J. P. M de Vreught y H. J. Honig. Slow and fast parallel recognition. In Proc. of *2nd International Workshop on Parsing Technologies, IWPT'91*, pp 127–135, Cancun, Mexico. (1991)
- [141] E. Wehrli. *L'analyse syntaxique des langues naturelles*. Masson, Paris. (1988)
- [142] D. J. Weir. Characterizing mildly context-sensitive grammar formalisms. *PhD Thesis*, University of Pennsylvania, PA, USA. (1988)
- [143] W. A. Woods. Transition network grammars for natural language. *Communications of the ACM*, 13, pp 591–606. (1970)
- [144] The XTAG Research Group. A lexicalized tree adjoining grammar for English. <http://www.cis.upenn.edu/~xtag>. Technical Report IRCS 95-03, IRCS, Institute for Research in Cognitive Science, University of Pennsylvania, Philadelphia PA, EE.UU. (1995)
- [145] D. H. Younger. Recognition of context-free languages in time  $n^3$ . *Information and control*, 10, pp 189–208. (1967)
- [146] R. Zajac. Inheritance and constraint-based grammar formalisms. *Computational Linguistics*, 18(2), pp 159–180. (1992)

