Lógica epistémica: sintaxis y semántica, completitud, expresabilidad.

Carlos David Martínez Rial

Lógica epistémica: sintaxis y semántica, completitud, expresabilidad.

Carlos David Martínez Rial

Memoria presentada como parte de los requisitos para la obtención del título de Grado en Matemáticas por la Universidad de Sevilla.

Tutorizada por

Prof. Andrés Cordón Franco

Índice general

Resumen

A lo largo de este trabajo nos centraremos en presentar, desarrollar y estudiar diferentes lógicas epistémicas y algunas de sus propiedades más destacadas. En el primer capítulo hablaremos del sistema básico de la lógica epistémica multiagente, asentaremos la base del lenguaje, la semántica que utilizaremos en el resto del trabajo. Nos centraremos sobre todo en el lenguaje $\mathcal{S}5$ y su axiomatización **S5**, aquí también definiremos los conceptos de derivación y derivación con premisas que nos servirán para dar una prueba del Teorema de Deducción. En el segundo capítulo añadiremos el operador de conocimiento común, definiremos el nuevo lenguaje y el nuevo sistema de axiomas. Y en el tercer capítulo haremos lo mismo pero incorporando la noción de anuncios públicos, estudiando de esta forma *PA* y *PAC*. Además en cada capítulo probaremos el Teorema de Completitud de cada lógica que vayamos introduciendo, y comprobaremos sus capacidades de expresividad, utilizando los conceptos de Bisimulación y Juegos para comparar modelos y de este modo podamos decir si dos lógicas tienen el mismo poder expresivo o no.

Abstract

Throughout this work we will focus on presenting, developing and studying different epistemic logics and some of their most outstanding properties. In the first chapter we will talk about the basic system of multiagent epistemic logic, we will establish the basis of language, the semantics that we will use in the rest of the work. We will focus mainly on the language $\mathcal{S}5$ and its axiomatization **S5**, here we will also define the concepts of derivation and derivation with premises that will help us to give a proof of the Deduction Theorem. In the second chapter we will add the common knowledge operator, define the new language and the new system of axioms. And in the third chapter we will do the same but incorporating the notion of public announcements, studying in this way *PA* and *PAC*. In addition, in each chapter we will prove the Completeness Theorem of each logic that we introduce, and we will check its expressivity, using the concepts of Bisimulation and Games to compare models and in this way we can tell if two logics have the same expressive power or not.

1 Introducción

El presente trabajo se enmarca dentro del campo de la lógica modal y, a lo largo de él, nos centraremos en presentar, desarrollar y estudiar distintas *lógicas epistémicas multiagente* y algunas de sus propiedades matemáticas más destacadas; poniendo especial énfasis en el estudio de las propiedades de completitud y expresabilidad de cada lógica.

La lógica modal es, en líneas muy generales, el estudio del comportamiento deductivo de las expresiones 'es necesario que' y 'es posible que'. No existe un único sistema de lógica modal privilegiado sino que, justo al contrario, el término es comúnmente usado para englobar una familia de lógicas diversas, pero con ciertas propiedades comunes: típicamente son extensiones de la lógica proposicional mediante nuevos *operadores modales*. La siguiente lista, tomada de [Gar18], describe algunas de las lógicas modales más conocidas:

Lógica	**Símbolo**	**Expresiones formalizadas**
Lógica Modal básica	$\Box$	Es necesario que...
	$\Diamond$	Es posible que...
Lógica Deóntica	O	Es obligatorio que...
	P	Está permitido que...
	F	Está prohibido que...
Lógica Temporal	G	Siempre será el caso que...
	F	Será alguna vez el caso que...
	H	Ha sido siempre el caso que...
	P	Fue alguna vez el caso que...
Lógica Doxástica	B_a	El razonador a cree que...
Lógica Epistémica	K_a	El razonador a sabe que...
	C	Es conocimiento común que...
	$\vdots$	

En el presente trabajo, nos centraremos en la última interpretación de los operadores modales señalada en la lista (lógica epistémica o del conocimiento) y permitiremos la presencia de varios agentes razonadores distintos (lógica epistémica multiagente).

Históricamente, fue Aristóteles en '*De Interpretatione*' el primero en hablar de una forma sistemática de la lógica modal. Aristóteles no solo observó que "necesidad" implica "posibilidad" (y no al revés), sino que, entre otras propiedades básicas, mostró que ambas nociones son interdefinibles. Es decir, la proposición '*es posible p*' puede ser definida como '*no es necesario* $\neg p$'. Y de igual modo, '*es necesario p*' puede ser definida como '*no es posible* $\neg p$'. Después de Aristóteles, muchos otros filósofos hicieron nuevas aportaciones al campo pero no fue hasta el siglo XX cuando se reanudó el interés por la lógica modal desde un punto de vista *lógico-matemático*. Caben destacar dos trabajos pioneros. C. I. Lewis comenzó la búsqueda de un sistema de axiomas para caracterizar la 'implicación estricta', [Lew12; Lew18; LL32], y, en este contexto, Lewis definió cinco sistemas: $\mathcal{S}1 - \mathcal{S}5$. Dos de ellos, $\mathcal{S}4$ y $\mathcal{S}5$, se siguen utilizando hoy en día, y los definiremos en este trabajo. Por otra parte, en el estudio de la teoría de modelos de las relaciones de consecuencia lógica, R. Carnap ([Car42; Car47]) introduce ciertas nociones de descripciones de estado que anticipan el concepto fundamental de *semántica de mundos posibles.* Aunque los trabajos de Hintikka ([Hin57]) y Kanger ([Kan57b; Kan57a]) ya desarrollan las ideas de la semántica de mundos posibles, a dicha semántica también se la conoce como *semántica de Kripke* después de que Saul Kripke dedicara sus primeros trabajos a la semántica de la lógica modal ([Kri59]). En líneas muy generales, Kripke considera un dominio de mundos posibles relacionados entre sí por una relación de accesibilidad y propone la definición:

> "'*Necesariamente p*' es cierto en un mundo o estado s si y solo si p es cierto en cada mundo s' accesible desde s".

Esto es, en el contexto de la lógica epistémica, en cada estado o mundo posible s cada agente razonador a tiene asociado otros estados o mundos que son los que él o ella considera posibles desde s y diremos que el agente a sabe o conoce una propiedad φ en s si dicha propiedad es cierta en todos los mundos que él o ella considera posibles desde s. La semántica de Kripke será la base del presente trabajo y se materializa en la noción fundamental de *modelo de Kripke* para interpretar las fórmulas de la lógica epsitémica.

Como hemos observado, la *lógica epistémica* es un subcampo de la lógica modal que se centra en el estudio de operadores modales interpretados como operadores de conocimiento. El operador básico modal para expresar conocimiento será $K_a\varphi$, donde el subíndice a denota al agente razonador al que se refiere el operador, con el significado: "el agente a sabe o conoce la propiedad φ". No será el único operador que veamos en este trabajo. También trabajaremos con el operador $C_B\varphi$, que sirve para expresar *conocimiento común* de φ en un grupo de agentes B. Más adelante, añadiremos también una componente dinámica a nuestro lenguaje mediante el concepto de *anuncios públicos* y consideraremos el operador $[\varphi]\psi$ para significar: "tras la actualización de modelo con el anuncio público de la fórmula φ, la fórmula ψ es el caso".

En este trabajo daremos algunas de las axiomatizaciones más importantes de la lógica epistémica multiagente. Partiremos de la axiomatización básica **K** e iremos añadiendo nuevos axiomas y reglas de inferencia para definir los sistemas: lógica epistémica multiagente **S5**, lógica epistémica multiagente con conocimiento común **S5C**, lógica de anuncios públicos **PA** y lógica de anuncios públicos con conocimiento común **PAC**. A continuación presentaremos unas tablas que representan estas axiomatizaciones. Antes es necesario comentar que las axiomatizaciones en la tabla están formados por los axiomas y reglas que la acompañan y los anteriores, es decir, cada axiomatización de una tabla es una extensión de la anterior. Además, todas las axiomatizaciones incluyen por defecto todas las instancias de las tautologías proposicionales (nuestros sistemas son extensiones de la lógica proposicional clásica).

Axiomatización	**Axiomas**	
K	• $K_a(\varphi \to \psi) \to (K_a\varphi \to K_a\psi)$	(Distribución de K_a sobre $\to$)
	• $\varphi, \varphi \to \psi \vDash \psi$	(Modus Ponens)
	• $\varphi \vDash K_a\varphi$	(Necesitación para K_a)
T	• $K_a\varphi \to \varphi$	(Axioma de verdad)
S4	• $K_a\varphi \to K_aK_a\varphi$	(Introspección positiva)
S5	• $\neg K_a\varphi \to K_a\neg K_a\varphi$	(Introspección negativa)
S5C	• $C_B(\varphi \to \psi) \to (C_B\varphi \to C_B\psi)$	(Distribución de C_B sobre $\to$)
	• $C_B\varphi \to (\varphi \land E_BC_B\varphi)$	(Axioma mixto de conocimiento común)
	• $C_B(\varphi \to E_B\varphi) \to (\varphi \to C_B\varphi)$	(Inducción de conocimiento común)
	• $\varphi \vDash C_B\varphi$	(Necesitación para C_B)

Axiomatización	**Axiomas**	
PA	• $K_a(\varphi \to \psi) \to (K_a\varphi \to K_a\psi)$	(Distribución de K_a sobre $\to$)
	• $\varphi, \varphi \to \psi \vDash \psi$	(Modus Ponens)
	• $\varphi \vDash K_a\varphi$	(Necesitación para K_a)
	• $K_a\varphi \to \varphi$	(Axioma de verdad)
	• $K_a\varphi \to K_aK_a\varphi$	(Introspección positiva)
	• $\neg K_a\varphi \to K_a\neg K_a\varphi$	(Introspección negativa)
	• $[\varphi]p \leftrightarrow (\varphi \to p)$	(Permanencia atómica)
	• $[\varphi]\neg\psi \leftrightarrow (\varphi \to \neg[\varphi]\psi)$	(Anuncio y negación)
	• $[\varphi](\psi \wedge \chi) \leftrightarrow ([\varphi]\psi \wedge [\varphi]\chi)$	(Anuncio y conjunción)
	• $[\varphi]K_a\psi \leftrightarrow (\varphi \to K_a[\varphi]\psi)$	(Anuncio y conocimiento)
	• $[\varphi][\psi]\chi \leftrightarrow [\varphi \wedge [\varphi]\psi]\chi$	(Composición de anuncios)
PAC	• $C_B(\varphi \to \psi) \to (C_B\varphi \to C_B\psi)$	(Distribución de C_B sobre $\to$)
	• $C_B\varphi \to (\varphi \wedge E_BC_B\varphi)$	(Axioma mixto de conocimiento común)
	• $C_B(\varphi \to E_B\varphi) \to (\varphi \to C_B\varphi)$	(Inducción de conocimiento común)
	• $\varphi \vDash C_B\varphi$	(Necesitación para C_B)
	• $\varphi \vDash [\psi]\varphi$	(Necesitación para $[\psi]$)
	• $\chi \to [\varphi]\psi, (\chi \wedge \varphi) \to E_B\chi \vDash \chi \to [\varphi]C_B\psi$	(Anuncios y conocimiento común)

La presente memoria está dividida en cuatro capítulos. El primero es introductorio y los tres restantes estudian los sistemas arriba indicados. Excepto en la presente Introducción, en cada capítulo seguimos la misma estructura: primero definimos la sintaxis y la semántica de la lógica estudiada; en segundo lugar, introducimos una axiomatización de la lógica; más tarde, demostramos que la axiomatización dada cumple el Teorema de Completitud; y concluiremos cada capítulo hablando de la expresividad de la lógica.

Con un poco más de detalle, en el Capítulo 2 hablaremos del sistema básico de la lógica epistémica multiagente y asentaremos la base del lenguaje y la semántica que utilizaremos en el resto del trabajo. En este capítulo nos centraremos, sobre todo, en el lenguaje $\mathcal{S}5$ y su axiomatización **S5**. Aquí también definiremos los conceptos de derivación y derivación con premisas, que nos servirán para dar una prueba del Teorema de la Deducción. La base para la prueba de Completitud para el sistema **S5** será el concepto de *modelo canónico*, cuyo dominio está formado por todos los conjuntos consistentes maximales de la lógica. Necesitaremos pues también utilizar el Lema de Lindenbaum para construir conjuntos consistentes maximales. En la última sección del capítulo, hablaremos del poder expresivo de los lenguajes. Para ello, necesitaremos definir algunos conceptos básicos de equivalencia. Continuaremos definiendo la bisimulación, una herramienta para comparar modelos de Kripke que consiste en crear una relación entre dos modelos, de tal forma que los estados que estén relacionados sean similares en cuanto a las proposiciones atómicas válidas en estos estados, y en cuanto a los otros estados que estén relacionados con ellos. Veremos que la condición de que dos estados sean bisimilares es suficiente para la equivalencia de modelos, pero no es necesaria; esto último lo probaremos con un contraejemplo. Otra herramienta de comparación de modelos que utilizaremos en este trabajo será la de juegos *spoiler-duplicador*. Estos juegos consistirán en enfrentar a dos jugadores: duplicador, con el objetivo de demostrar que los modelos son iguales, y spoiler, que intentará probar que son diferentes.

En el capítulo 3, comenzaremos definiendo el operador modal de conocimiento común, C_B, y el lenguaje que se crea añadiendo este operador junto con la clase de modelos epistémicos $\mathcal{S5C}$. En la segunda sección, definiremos el sistema de axiomas **S5C**. Para demostrar la completitud, en este capítulo, tenemos el problema de que la clase $\mathcal{S5C}$ no es compacta y ello nos impide repetir la construcción del modelo canónico del capítulo anterior. Para solucionarlo, utilizaremos la función de clausura de una fórmula epistémica, que nos ayudará a tomar una región finita donde poder seguir con la construcción del modelo canónico. La última sección de este capítulo empezará comparando el lenguaje de este capítulo con el lenguaje del capítulo anterior en términos de poder expresivo. Obtendremos como resultado que al hablar de lenguajes con un único agente obtenemos una equivalencia entre ambos lenguajes, pero en cambio, cuando tomamos los lenguajes multiagente obtenemos que el lenguaje epistémico con conocimiento común tiene más poder expresivo. Describiremos también cómo adaptar los juegos spoiler-duplicador a este contexto.

En el capítulo 4, estudiaremos dos lógicas, **PA** y **PAC**. Comenzaremos con un ejemplo para introducir informalmente el concepto de anuncio público y después definiremos el nuevo lenguaje y cómo se adapta la semántica de los modelos de Kripke a este contexto *dinámico*. En la segunda sección del capítulo, definiremos los sistemas de axiomas que definen ambas lógicas. Para probar la completitud de **PA** vamos a definir una función de traslación que lleva las fórmulas del lenguaje de anuncios públicos al lenguaje epistémico básico que definimos en el primer capítulo. También daremos el concepto de complejidad que nos servirá para ordenar las fórmulas al realizar la demostración. Probaremos que todas las fórmulas son equivalentes a su traslación, y ahora sí, demostraremos el Teorema de Completitud de **PA**. Para la prueba del Teorema de Completitud de **PAC** seguiremos un procedimiento similar a la prueba de **S5C**.

Definiremos la clausura para tomar un espacio finito donde trabajar, definiremos el modelo canónico de este lenguaje y probaremos el Lema de Lindenbaum. En la última sección de este capítulo, sobre *PA* veremos que su lenguaje es equivalente al epistémico sin conocimiento común. Esto es sorprendente porque vemos que el operador de anuncios públicos no añade capacidad expresiva al lenguaje de la lógica epistémica multiagente. Esto es debido a la función de traslación, aunque conlleva el inconveniente de que la longitud de la traducción puede ser exponencialmente más grande que la longitud de la fórmula original. En cuanto a *PAC*, veremos que su lenguaje posee más poder expresivo que el lenguaje de $\mathcal{S}5C$.

En cuanto a las referencias bibliográficas, hemos basado gran parte de este trabajo en el libro de Hans van Ditmarsch, Wiebe van der Hoek y Bartfeld Kooi titulado "Dynamic Epistemic Logic" [DHK93], complementando algunos aspectos concretos con otros trabajos que se referencian en la bibliografía.

2 Sistema básico de la lógica epistémica multiagente

En este primer capítulo, aparte de desarrollar las nociones básicas de la lógica epistémica, vamos a centrarnos en el sistema lógico epistémico más popular **S5**. Estudiaremos su lenguaje y su semántica en la primera sección. Para ello definiremos los modelos de Kripke, esenciales para representar la semántica. La segunda sección la utilizaremos para caracterizar la lógica sintácticamente, aquí definiremos **S5** de manera formal, además probaremos el Teorema de Deducción y comentaremos brevemente la controversia que ha provocado este resultado en el campo. En la tercera sección demostraremos la completitud de la lógica que estamos estudiando. La última sección de este capítulo la utilizaremos para ver la expresabilidad de la lógica.

2.1 Lenguaje y semántica

2.1.1 Sintaxis

El lenguaje de la lógica epistémica es una extensión del lenguaje de la lógica proposicional clásica, de donde tomaremos el concepto de las proposiciones atómicas y sus operaciones básicas, mediante operadores modales unarios cuyo propósito es representar el conocimiento de un cierto agente.

El punto de partida es un conjunto numerable de variables proposicionales o átomos que llamaremos P:

$$P = \{p_0, p_1, p_2, \dots\}$$

Escribiremos también $p, q, r, \dots$ para denotar dichas variables proposicionales. A esto le añadiremos un conjunto *finito* de agentes A:

$$A = \{a_1, \dots, a_n\},$$

con $n \geq 1$, donde cada a_i es un nombre para un agente. Escribiremos también $a, b, c, \dots$ para denotar dichos agentes.

En general, escribiremos p para denotar un elemento cualquiera del conjunto de átomos P, y escribiremos a para denotar a un elemento cualquiera del conjunto de agentes A. La definición formal del lenguaje que vamos a utilizar sería:

Definición 2.1 (Lenguaje). *Sean P un conjunto de variables atómicas y A un conjunto finito de agentes. El lenguaje de la lógica epistémica multiagente basada en P y A consta de los siguientes símbolos:*

1. *Variables: los elementos de P.*
2. *Conectivas proposicionales: $\neg$, $\wedge$, $\vee$, $\rightarrow$*
3. *Operadores modales: K_a, para cada agente $a \in A$.*

Definición 2.2 (Fórmulas). *Sean P un conjunto de variables atómicas y A un conjunto finito de agentes. El conjunto de las fórmulas de la lógica epistémica multiagente basada en P y A, $\mathcal{L}_K$, se define como sigue:*

1. *Las variables proposicionales son fórmulas.*
2. *Si φ es una fórmula, también lo es $\neg\varphi$.*
3. *Si φ y ψ son fórmulas, también lo son $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$ y $(\varphi \rightarrow \psi)$.*
4. *Si φ es una fórmula, también los es $K_a\varphi$ para cada agente $a \in A$.*

Esto es, el conjunto de fórmulas está generado por la siguiente notación BNF:

$$\varphi \ ::= \ p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid (\varphi \vee \varphi) \mid (\varphi \rightarrow \varphi) \mid K_a\varphi$$

Notación 2.1. A lo largo de la memoria, usaremos una serie de abreviaturas y convenios notacionales habituales para hacer más legibles nuestras fórmulas. Por ejemplo, las constantes booleanas $\top$ (verdadero) y $\bot$ (falso) denotarán, respectivamente, $(p \vee \neg p)$ y $\neg(p \vee \neg p)$; y $(\varphi \leftrightarrow \psi)$ es una abreviatura para $((\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi))$. Además, si no genera confusión, omitiremos los paréntesis más externos a la hora de escribir nuestras fórmulas y haremos uso de los criterios de precedencia usuales entre las distintas conectivas proposicionales. Aunque el lenguaje *oficial* sea el especificado en la Definición anterior, usaremos estas extensiones del lenguaje cuando nos convenga (de hecho, como es bien conocido, podríamos incluir en el lenguaje oficial solo $\neg, \wedge$ y considerar que $\vee, \rightarrow$ son también definidos).

También consideraremos las siguientes extensiones modales del lenguaje oficial.

El operador de conocimiento 'K_a' representará que, para el agente a, la fórmula que sigue al operador es cierta. Leeremos $K_a\varphi$ como "el agente a sabe o conoce o tiene la certeza de φ". Cuando digamos "el agente a no sabe o no tiene la certeza de $\neg\varphi$" escribiremos $\neg K_a\neg\varphi$, que también podemos interpretarlo como "φ es consistente con el conocimiento de a" o "el agente a considera que φ es posible", y lo denotaremos como $\hat{K}_a\varphi$. Esto es:

Definición 2.3 (Operador de Posibilidad). *Definimos el operador modal de posibilidad $\hat{K}_a$ como sigue:*

$$\hat{K}_a\varphi = \neg K_a\neg\varphi.$$

Si consideramos un grupo de agentes $B \subseteq A$, podemos considerar expresiones del tipo "todos los agentes de B saben la propiedad φ". Introducimos un operador modal unario de conocimiento grupal E_B para capturar esa noción.

Definición 2.4 (Operadores de conocimiento grupal, E_B y $\hat{E}_B$). *Sea B un subconjunto del conjunto de agentes A. Definimos operador de conocimiento grupal para B como sigue:*

$$E_B\varphi = \bigwedge_{b\in B} K_b\varphi.$$

De forma análoga a lo anterior, denotamos por $\hat{E}_B\varphi$ la fórmula $\neg E_B\neg\varphi$, que leeremos como "al menos un agente de B considera φ posible".

Veamos un ejemplo de formalización usando el lenguaje de la lógica epistémica.

Ejemplo 2.1. Tenemos tres agentes, Ángel, Beatriz y Celia. Además definimos dos átomos, p ("Ángel tiene una hermana") y q ("Ángel tiene un hermano"). A continuación, escribiremos las siguientes frases en el lenguaje que hemos definido:

- Si Ángel tiene una hermana, él lo sabría.
  $p \to K_a p$
- Beatriz sabe que Ángel sabe si él tiene una hermana.
  $K_b(K_a p \vee K_a\neg p)$
- Celia sabe que si Ángel tiene un hermano, no tiene una hermana.
  $K_c(q \to \neg p)$
- Ángel considera que es posible que Beatriz no sepa que él tiene una hermana.
  $\hat{K}_a\neg K_b p$
- Todo el grupo sabe que Ángel no tiene una hermana o un hermano, si él no sabe que los tiene.
  $E_{\{a,b,c\}}(\neg K_a(p \vee q) \to (\neg p \wedge \neg q))$

- Ángel sabe que si hay alguien en el grupo que no sabe que él tiene una hermana, debería ser Beatriz.
  $K_a(\neg E_{\{a,b,c\}} p \to \neg K_b p)$

Con el siguiente ejemplo vamos a ver, en la práctica, algunos razonamientos lógicos y cómo funciona las iteraciones del conocimiento de los agentes. Razonaremos de manera informal, a falta del concepto fundamental de modelo de Kripke que introduciremos en la siguiente subsección.

Ejemplo 2.2 (Números consecutivos). Tenemos dos agentes, Ángel y Beatriz, que denotaremos como a y b respectivamente. Ángel y Beatriz están sentados uno frente al otro, y cada uno de ellos tiene un papel con un número natural en su frente, de manera que no pueden ver el número que tienen ellos mismos, pero sí pueden ver el número que tiene el otro agente. Saben que son *números consecutivos*. Vamos a notar como a_n a la expresión "Ángel tiene el número n en su cabeza" y como b_n a la expresión "Beatriz tiene el número n en su cabeza".
En este ejemplo, vamos a tomar como premisas a_3 y b_2. Es decir, suponemos que estamos en la situación en la que Ángel tiene el número 3 y Beatriz tiene el número 2. Describamos, con el lenguaje descrito anteriormente, el conocimiento de los dos agentes.

1. Como cada uno puede ver el número de la cabeza del otro, tenemos $K_a b_2$, y análogamente $K_b a_3$.
2. Como ambos agentes saben que los dos números son consecutivos, entonces $K_a(a_1 \vee a_3)$ y $K_b(b_2 \vee b_4)$.
3. Razonando un poco, Ángel se da cuenta de que en el caso de que él tenga en la cabeza el 1, Beatriz dudará entre si tiene el 0 o el 2, y en el caso en el que tenga el 3 sobre su cabeza, Beatriz dudará entre el 2 o el 4. Entonces se podría decir que $K_a K_b(b_0 \vee b_2 \vee b_4)$, y análogamente $K_b K_a(a_1 \vee a_3 \vee a_5)$.
4. Ángel sabe perfectamente los números entre los que puede dudar Beatriz, porque o bien, $K_b K_a(b_0 \vee b_2)$, o bien $K_b K_a(b_2 \vee b_4)$, entonces podemos decir que $K_a K_b K_a(b_0 \vee b_2 \vee b_4)$.
5. Ángel piensa que es posible que tenga tanto el 1 como el 3 en su cabeza, y Beatriz piensa que es posible que tenga el 2 o el 4. Tenemos entonces $\hat{K}_a a_1 \wedge \hat{K}_a a_3$ y $\hat{K}_b b_2 \wedge \hat{K}_b b_4$.

A partir de aquí vamos a suponer que el objetivo de esta situación es adivinar el número que tienen en su cabeza. Denotaremos, por tanto, win_a si tenemos que "Ángel sabe el número que tiene en su cabeza", igualmente si "Beatriz conoce el número de su cabeza" escribiremos win_b .

6. La situación solo permite que un individuo gane si ve el 0, en ese caso tendrá claro que tiene el 1 y ganará, y por el contrario, podrá pensar que el otro individuo puede ganar si ve en su cabeza el 1. Entonces ambos agentes tienen claro que ninguno de los dos puede ganar, $K_a(\neg win_a \wedge \neg win_b)$ y $K_b(\neg win_a \wedge \neg win_b)$.
7. Beatriz sabe que Ángel no tiene el 5 porque puede ver que tiene el 3, y Ángel sabe que él no tiene el 5 porque puede ver que Beatriz tiene el 2. Pero no todo el mundo sabe que todo el mundo sabe que Ángel no tiene el 5, es decir, Beatriz no sabe que Ángel sabe que no tiene el 5, porque ella tiene como un caso posible tener el 4 en su cabeza, entonces para Beatriz cabe la posibilidad de que Ángel tenga como un caso posible tener un 5 en su cabeza ($\hat{K}_b\hat{K}_a a_5$). Con la notación vista anteriormente se tiene $E_{\{a,b\}}\neg a_5 \wedge \neg E_{\{a,b\}}E_{\{a,b\}}\neg a_5$.
8. Tenemos que todo el mundo sabe que ninguno de los dos puede ganar, pero también tenemos que Ángel piensa que puede tener el 1, y por tanto para Ángel cabe la posibilidad de que a Beatriz le quepa la posibilidad de que Ángel tenga el 1 y sepa que tiene el 1, es decir, $\hat{K}_a\hat{K}_b(a_1 \wedge K_a a_1)$. Por tanto no todo el mundo sabe que todo el mundo sabe que ninguno de los agentes puede ganar, es decir, $E_{\{a,b\}}(\neg win_a \wedge \neg win_b) \wedge \neg E_{\{a,b\}}E_{\{a,b\}}(\neg win_a \wedge \neg win_b)$.

2.1.2 Semántica

Con este ejemplo hemos conocido más acerca del lenguaje que vamos a utilizar a partir de ahora, pero es de vital importancia, para el trato formal de la lógica del conocimiento, la semántica que usa. Necesitamos una estructura donde poder situar, representar y estudiar la lógica epistémica con mayor facilidad. Para ello vamos a utilizar los modelos de Kripke.

Definición 2.5 (Modelo de Kripke). *Sean un conjunto numerable de variables atómicas, P, y un conjunto finito de agentes, A. Un modelo de Kripke es una 3-upla $M = \langle S, R^A, V^P \rangle$, donde:*

- *S es un conjunto de estados. También llamado dominio de M, $\mathcal{D}(M)$.*
- *$R^A : A \to 2^{S\times S}$ es una función que asocia a cada agente $a \in A$ un conjunto $R^A(a) \subseteq S\times S$ (que interpretaremos como la relación de accesibilidad entre estados para el agente a).*

- $V^P : P \to 2^S$ *es una función de validación, que a cada variable proposicional* $p \in P$ *le asigna un conjunto de estados* $V^P(p) \subseteq S$ *(que interpretaremos como los estados en los cuales* p *es cierto).*

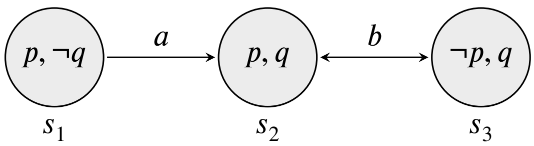
Para simplificar la notación, escribiremos $M = \langle S, R, V \rangle$ si los conjuntos P y A están claros por el contexto y escribiremos simplemente $R(a)$ o R_a en lugar de $R^A(a)$ y escribiremos $V(p)$ o V_p en lugar de $V^P(p)$. Por otra parte, para indicar que el par de estados (s, t) está en la relación binaria R_a, una veces usaremos notación prefija $R_a st$ y otras veces escribiremos $sR_a t$, según convenga.

Observación 2.1 (Modelos epistémicos). Si tenemos que R_a es una relación de equivalencia para todo agente a, diremos que M es un *modelo epistémico*, podemos utilizar $\sim_a$ en lugar de R_a en esta situación y representaremos el modelo como $M = \langle S, \sim, V \rangle$.

Ejemplo 2.3. Sea el conjunto de agentes $A = \{a, b\}$ y el conjunto de proposiciones atómicas $P = \{p, q\}$. Daremos el siguiente ejemplo de modelo de Kripke, $M_1 = \langle S_1, R_1, V_1 \rangle$, tal que:

- $S_1 = \{s_1, s_2, s_3\}$
- $R_1(a) = \{(s_1, s_2)\}$
- $R_1(b) = \{(s_2, s_3), (s_3, s_2)\}$
- $V_1(p) = \{s_1, s_2\}$
- $V_1(q) = \{s_2, s_3\}$

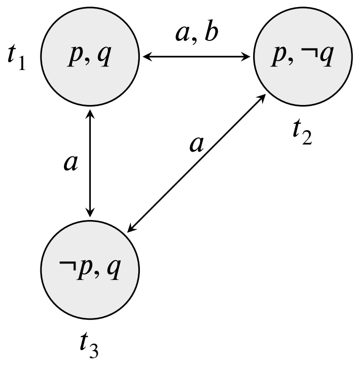
Representamos el modelo mediante la siguiente figura:



Ahora damos otro ejemplo de modelo de Kripke, pero en este caso el modelo es epistémico, puesto que la relación es de equivalencia. Damos, por tanto, el modelo $M_2 = \langle S_2, \sim, V_2 \rangle$, tal que:

- $S_2 = \{t_1, t_2, t_3\}$
- $\sim_a = \{(t_1, t_1), (t_1, t_2), (t_1, t_3), (t_2, t_1), (t_2, t_2), (t_2, t_3), (t_3, t_1), (t_3, t_2), (t_3, t_3)\}$
- $\sim_b = \{(t_1, t_1), (t_1, t_2), (t_2, t_1), (t_2, t_2), (t_3, t_3)\}$
- $V_2(p) = \{t_1, t_2\}$
- $V_2(q) = \{t_1, t_3\}$

Al representar el modelo cabe destacar que para mayor claridad en la representación no escribiremos las relaciones reflexivas.



Las fórmulas epistémicas serán evaluadas en pares (M, s), siendo $M = \langle S, R, V \rangle$ un modelo de Kripke y s un estado $s \in \mathcal{D}(M)$. Si M es un modelo epistémico, entonces decimos que (M, s) es un estado epistémico.

Definición 2.6 (Relación de Validez). *Dados un modelo de Kripke $M = \langle S, R, V \rangle$ y una fórmula φ, la relación de validez, $\models$, se define inductivamente como sigue:*

$M, s \models p$	*sii $s \in V(p)$*
$M, s \models \neg\varphi$	*sii no es cierto que $M, s \models \varphi$*
$M, s \models (\varphi \wedge \psi)$	*sii $M, s \models \varphi$ y $M, s \models \psi$*
$M, s \models (\varphi \vee \psi)$	*sii $M, s \models \varphi$ o $M, s \models \psi$*
$M, s \models (\varphi \rightarrow \psi)$	*sii no es cierto que $M, s \models \varphi$ o $M, s \models \psi$*
$M, s \models K_a\varphi$	*sii para todo t tal que sR_at, se tiene que $M, t \models \varphi$*
$M, s \models \hat{K}_a\varphi$	*sii existe t tal que sR_at y $M, t \models \varphi$*

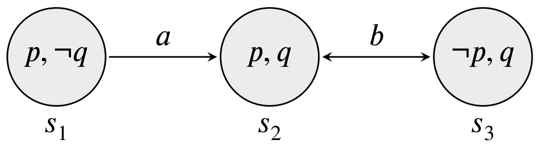
Observación 2.2.

1. Usaremos la notación $M, s \not\models \varphi$ para representar: "no es cierto que $M, s \models \varphi$".
2. La definición de la relación de validez para el operador de posibilidad $\hat{K}_a$ es la que se obtiene de desarrollar la equivalencia $\hat{K}_a\varphi \equiv \neg K_a \neg\varphi$.

Ejemplo 2.4. Sea el conjunto de agentes $A = \{a, b\}$ y el conjunto de proposiciones atómicas $P = \{p, q\}$, vamos a representar de nuevo el modelo de Kripke $M_1 = \langle S_1, R_1, V_1 \rangle$, definido por:

- $S_1 = \{s_1, s_2, s_3\}$
- $R_1(a) = \{(s_1, s_2)\}$
- $R_1(b) = \{(s_2, s_3), (s_3, s_2)\}$

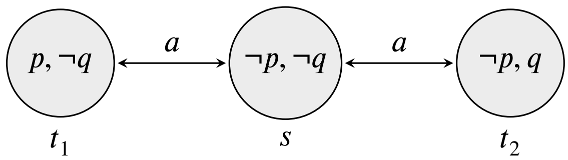
- $V_1(p) = \{s_1, s_2\}$
- $V_1(q) = \{s_2, s_3\}$



Veamos algunas formulas epistémicas que se verifican en este modelo:

1. Se tiene $M_1, s_1 \vDash p$, ya que $s_1 \in V_1(p)$.
2. Se tiene que $M_1, s_2 \vDash K_b q$, puesto que $R_1(b) = \{(s_2, s_3), (s_3, s_2)\}$ y $M_1, s_3 \vDash q$.
3. Se tiene $M_1, s_2 \vDash p \wedge K_b \neg p$, puesto que $s_2 \in V_1(p)$ y, por otra parte, se tiene que $\{(s_2, s_3), (s_3, s_2)\} = R_1(b)$ y $M, s_3 \vDash \neg p$.
4. Se tiene $M_1, s_1 \vDash K_a K_b q$, como consecuencia de que $R_1(a) = \{(s_1, s_2)\}$ y $M, s_2 \vDash K_b q$.

Ejemplo 2.5. Una de las particularidades de la semántica de Kripke es que puede mostrar con facilidad como el operador de conocimiento K_a conmuta con la conjunción, pero no lo hace con la disjunción. Es decir, se tiene $K_a(p \wedge q) \equiv K_a p \wedge K_a q$, pero, en cambio, $K_a(p \vee q) \not\equiv K_a p \vee K_a q$. (Aquí estamos usando el símbolo $\equiv$ de manera informal, más adelante daremos una definición formal de este símbolo.) La equivalencia $K_a(p \wedge q) \equiv K_a p \wedge K_a q$ se obtiene viendo que ambas fórmulas son ciertas en exactamente los mismos estados por la definición de la relación de validez. Para probar $K_a(p \vee q) \not\equiv K_a p \vee K_a q$, utilizaremos el siguiente contraejemplo.



En este modelo, podemos afirmar que se verifica $(M, s) \vDash K_a(p \vee \neg p)$, pero tanto $K_a p$ como $K_a \neg p$ son falsas en el estado s.

Esto también ocurre con el operador de posibilidad $\hat{K}_a$, pero al contrario. El operador $\hat{K}_a$ conmuta con la disyunción y no lo hace con la conjunción, es decir, $\hat{K}_a(p \vee q) \equiv \hat{K}_a p \vee \hat{K}_a q$, pero, en cambio, $\hat{K}_a(p \wedge q) \not\equiv \hat{K}_a p \wedge \hat{K}_a q$. Para probar la equivalencia $\hat{K}_a(p \vee q) \equiv \hat{K}_a p \vee \hat{K}_a q$, de nuevo basta con utilizar la definición de la relación de validez. Para $\hat{K}_a(p \wedge q) \not\equiv \hat{K}_a p \wedge \hat{K}_a q$ utilizaremos el mismo contraejemplo anterior.

En este caso tenemos $\hat{K}_a p \wedge \hat{K}_a q$ en el estado s, pero de ningún modo podemos obtener la afirmación $\hat{K}_a(p \wedge q)$ en s.

Definición 2.7 (Validez y Satisfacibilidad). *Sean M un modelo de Kripke, $\mathcal{X}$ una clase de modelos de Kripke y φ una fórmula epistémica.*

1. *Diremos que φ es válida o verdadera en (M, s) si $M, s \vDash \varphi$.*
2. *Diremos que φ es válida en el modelo M si para todo estado $s \in \mathcal{D}(M)$ se tiene que $M, s \vDash \varphi$; y lo escribiremos $M \vDash \varphi$.*
3. *Diremos que φ es válida en la clase $\mathcal{X}$ si para todo modelo $M \in \mathcal{X}$ y para todo estado $s \in \mathcal{D}(M)$ se tiene que $M, s \vDash \varphi$; y lo escribiremos $\mathcal{X} \vDash \varphi$.*
4. *Diremos que φ es válida si para todo modelo de Kripke M se tiene que φ es válida en M; y lo escribiremos $\vDash \varphi$.*
5. *Diremos que φ es satisfacible en el modelo M si existe algún estado $s \in \mathcal{D}(M)$ tal que $M, s \vDash \varphi$.*
6. *Diremos que φ es satisfacible en la clase $\mathcal{X}$ si existe algún modelo $M \in \mathcal{X}$ y algún estado $s \in \mathcal{D}(M)$ tal que $M, s \vDash \varphi$.*
7. *Diremos que φ es satisfacible si existe algún modelo de Kripke M y algún estado $s \in \mathcal{D}(M)$ tal que $M, s \vDash \varphi$.*

La siguiente definición recoge las principales clases de modelos de Kripke $\mathcal{X}$ que se emplean en la lógica epistémica. En este trabajo nos centraremos en la clase $\mathcal{K}$ de todos los modelos de Kripke, y en la clase $\mathcal{S}5$ de los modelos epistémicos.

Definición 2.8. *Sea R_a la relación de cada agente $a \in A$ en un modelos de Kripke $M = \langle S, R, V \rangle$.*

1. *La clase de todos los modelos de Kripke se denota como $\mathcal{K}$. Nótese que $\mathcal{K} \vDash \varphi$ es equivalente a decir $\vDash \varphi$.*
2. *R_a es sobreyectiva si para todo s existe t tal que $R_a st$. La clase de modelos de Kripke sobreyectivos (es decir, con R_a sobreyectiva para todo a) se denota como $\mathcal{KD}$.*
3. *R_a es reflexiva si para todo s, $R_a ss$. La clase de modelos de Kripke reflexivos se denota como $\mathcal{T}$.*
4. *R_a es transitiva si $\forall s, t, u$, se tiene que si $R_a st$ y $R_a tu$ entonces $R_a su$. La clase de modelos de Kripke transitivos se denota por $\mathcal{K}4$. Y la clase de los que además de transitivos son reflexivos se denomina $\mathcal{S}4$.*
5. *R_a es euclídea si para todo s, t, u, se tiene que si $R_a st$ y $R_a su$ entonces $R_a tu$. La clase de modelos de Kripke que son euclídeos y transitivos se llama $\mathcal{K}45$. La clase de los modelos que además son sobreyectivos se denota por $\mathcal{KD}45$.*
6. *R_a es una relación de equivalencia si R_a es reflexiva, transitiva y simétrica ($\forall s, t$, si $R_a st$ entonces $R_a ts$). Equivalentemente, R_a es relación de equivalencia si tenemos*

que R_a es reflexiva, transitiva y euclídea. La clase de los modelos de Kripke con R_a una relación de equivalencia para todo agente $a \in A$ se denota por $\mathcal{S}5$.

Los modelos de Kripke cumplen una serie de propiedades muy útiles para el desarrollo de la semántica. A estas propiedades se las denomina *omnisciencia lógica* y se recogen en la siguiente proposición. Estas propiedades ponen de manifiesto que estamos considerando agentes inteligentes idealizados que se comportan como perfectos razonadores lógicos.

Proposición 2.1 (Omnisciencia lógica). Sean φ, ψ fórmulas de $\mathcal{L}_K$, y sea K_a un operador epistémico para $a \in A$.

- $\mathcal{K} \vDash K_a\varphi \wedge K_a(\varphi \to \psi) \to K_a\psi$ LO1
- $\mathcal{K} \vDash \varphi \Rightarrow \quad \mathcal{K} \vDash K_a\varphi$ LO2
- $\mathcal{K} \vDash \varphi \to \psi \Rightarrow \quad \mathcal{K} \vDash K_a\varphi \to K_a\psi$ LO3
- $\mathcal{K} \vDash \varphi \leftrightarrow \psi \Rightarrow \quad \mathcal{K} \vDash K_a\varphi \leftrightarrow K_a\psi$ LO4
- $\mathcal{K} \vDash (K_a\varphi \wedge K_a\psi) \to K_a(\varphi \wedge \psi)$ LO5
- $\mathcal{K} \vDash K_a\varphi \to K_a(\varphi \vee \psi)$ LO6
- $\mathcal{S}5 \vDash \neg(K_a\varphi \wedge K_a\neg\varphi)$ LO7

Demostración. A continuación probaremos algunas de las propiedades.

LO1. Probaremos $\mathcal{K} \vDash K_a\varphi \wedge K_a(\varphi \to \psi) \to K_a\psi$. Para ello, en primer lugar suponemos $M, s \vDash K_a\varphi \wedge K_a(\varphi \to \psi)$. Por un lado, tenemos:

$$M, s \vDash K_a\varphi \text{ si y solo si } \forall t \text{ tal que } (s, t) \in R_a, \text{ se tiene que } M, t \vDash \varphi.$$

Por otro lado, tenemos:

$$M, s \vDash K_a(\varphi \to \psi) \text{ si y solo si } \forall t \text{ tal que } (s, t) \in R_a, \text{ se tiene } M, t \vDash \varphi \to \psi.$$

Tenemos como consecuencia, $M, t \vDash \varphi$ y $M, t \vDash \varphi \to \psi$, $\forall t$ tal que $(s, t) \in R_a$. De aquí se obtiene que $M, t \vDash \psi, \forall t$ tal que $(s, t) \in R_a$ y, por tanto, $M, s \vDash K_a\psi$. En consecuencia, podemos decir que: $\mathcal{K} \vDash K_a\varphi \wedge K_a(\varphi \to \psi) \to K_a\psi$.

LO3. Vamos a probar $\mathcal{K} \vDash \varphi \to \psi \Rightarrow \mathcal{K} \vDash K_a\varphi \to K_a\psi$. Partimos de la hipótesis $\mathcal{K} \vDash \varphi \to \psi$. Suponemos $M, s \vDash K_a\varphi$, lo cual se verifica si y solo si $\forall t$ tal que $(s, t) \in R_a$, $M, t \vDash \varphi$. Por hipótesis, $M, t \vDash \varphi \to \psi$. Luego, $\forall t$ tal que $(s, t) \in R_a$, $M, t \vDash \psi$ y, por tanto, $M, s \vDash K_a\psi$.

LO7. Vamos a probar $\mathcal{S}5 \vDash \neg(K_a\varphi \wedge K_a\neg\varphi)$. Lo haremos por reducción al absurdo, suponiendo que existe algún *modelo epistémico* M tal que $M, s \vDash K_a\varphi \wedge K_a\neg\varphi$. Puesto que M es un modelo epistémico, se tiene que, en particular, $\sim_a$ es una relación reflexiva y todo estado está relacionado consigo mismo. Por tanto, de la definición de la semántica de K_a se tendría $M, s \vDash \varphi$ y $M, s \vDash \neg\varphi$, lo cual es una contradicción. ∎

Otorgando ciertas restricciones al conjunto de modelos de Kripke que estamos estudiando, se pueden obtener algunas propiedades o fórmulas interesantes, como por ejemplo, el axioma de la verdad.

Proposición 2.2. Sea $M = \langle S, R, V \rangle$ tal que R_a es reflexiva para un agente $a \in A$. Entonces M satisface el axioma de verdad para el operador K_a: $M \vDash K_a\varphi \to \varphi$. En consecuencia, $\mathcal{T} \vDash K_a\varphi \to \varphi$.

Demostración. Tenemos que demostrar que $K_a\varphi \to \varphi$ es válida en M. Esto es equivalente a probar $M, s \vDash \neg K_a\varphi \vee \varphi$ para todo estado s. Para ello, sea $s \in \mathcal{D}(M)$ cualquiera. Vamos a distinguir los siguientes casos.
Caso 1: $M, s \nvDash K_a\varphi$. Nada que probar.
Caso 2: $M, s \vDash K_a\varphi$. Puesto que R_a es reflexiva, $(s, s) \in R_a$ y, por la definición 2.6, se tiene que $M, s \vDash \varphi$.

∎

2.2 Axiomatización

Una axiomatización es, en esencia, una manera sintáctica de caracterizar las fórmulas válidas de una lógica. En este apartado vamos a definir las axiomatizaciones más relevantes en el contexto de la lógica epistémica multiagente. La primera que veremos es el sistema **K**, sistema básico pues estará contenido en el resto de lógicas que consideremos.

2.2.1 El sistema K

Definición 2.9 (El sistema K). *Sea A un conjunto finito de agentes y sea K_a un operador modal para cada $a \in A$. La lógica epistémica básica* **K** *está compuesta por los axiomas:*

instanciaciones de las tautologías proposicionales *(Prop)*

$K_a(\varphi \to \psi) \to (K_a\varphi \to K_a\psi)$ *(Axioma K)*

y las reglas de inferencia

$\varphi, \varphi \to \psi \vDash \psi$ *(MP)*

$\varphi \vDash K_a\varphi$ *(Nec)*

Observación 2.3.

1. Por comodidad, hemos introducido todas las instanciaciones de las tautologías de la lógica proposicional como axiomas de nuestro sistema **K**. Alternativamente, podría considerarse una de las axiomatizaciones usuales de la lógica proposicional y considerar solo las instanciaciones de esas fórmulas como axiomas del sistema.
2. El sistema **K** incluye no solo las tautologías en el lenguaje de la lógica proposicional sino que incluye también todas las instanciaciones de las tautologías en el lenguaje de la lógica epistémica. Esto es, no solo las fórmulas $p \vee \neg p$ o $p \to (q \to p)$ son axiomas de **K** sino que también lo son, por ejemplo, $K_aK_b\neg q \vee \neg K_aK_b\neg q$ o $K_ap \to (K_b(p \wedge \neg K_bK_bq) \to K_ap)$.
3. Obsérvese que la regla de necesitación (Nec) nos dice que si φ es válida, entonces $K_a\varphi$ también lo es. Ahora bien, ello no implica que el axioma $\varphi \to K_a\varphi$ sea válido. Por ejemplo, la fórmula $p \to K_ap$ no es un teorema del sistema **K**.Basta considerar un modelo con dos estados s_1 y s_2 conectados entre sí para el agente a y tales que en s_1 sea cierto p y en s_2 sea cierto $\neg p$.

En el futuro, trabajaremos con lógicas derivadas de **K** añadiendo nuevas reglas de inferencia. Es por ello que damos la siguiente definición general de *demostración* o *derivación* en un sistema lógico epistémico.

Definición 2.10 (Derivaciones). *Sea **X** una axiomatización cualquiera generada por los axiomas $Ax_1, Ax_2, ..., Ax_n$ y las reglas $Ru_1, Ru_2, ..., Ru_k$, donde cada regla Ru_j (con $j \leq k$), es de la forma "De $\varphi_1, ..., \varphi_{j_{ar}}$ se obtiene φ_j". Diremos que j_{ar} es la aridad de la regla. Una derivación o demostración para φ dentro de **X** es una sucesión finita $\varphi_1, ..., \varphi_m$ de fórmulas tales que:*

1. *$\varphi_m = \varphi$;*
2. *cada φ_i en la sucesión es:*
   a) *o bien, uno de los axiomas $Ax_1, ..., Ax_n$,*
   b) *o bien, el resultado de aplicar una regla Ru_j a j_{ar} fórmulas que aparecen en la sucesión antes de φ_i.*

*Si existe una derivación para φ en **X** escribiremos $\vdash_X \varphi$ y diremos que φ es un teorema de **X**, o que **X** prueba φ. Si el sistema **X** está claro por el contexto, a veces escribiremos solo $\vdash \varphi$.*

Ejemplo 2.6. Veamos algunas derivaciones para que quede más clara la definición.

1. $\vdash \varphi \to \psi \Rightarrow \vdash K_a\varphi \to K_a\psi$.

   Suponemos $\vdash \varphi \to \psi$. Por la regla de necesitación, obtenemos $\vdash K_a(\varphi \to \psi)$. Por el axioma K y aplicando modus ponens nos resulta $\vdash K_a\varphi \to K_a\psi$.

2. $\vdash K_a(\varphi \wedge \psi) \leftrightarrow (K_a\varphi \wedge K_a\psi)$.

   Solo probaremos la implicación de derecha a izquierda.

1 $\varphi \to (\psi \to (\varphi \wedge \psi))$	Prop
2 $K_a\varphi \to K_a(\psi \to (\varphi \wedge \psi))$	1, apartado anterior
3 $(K_a\varphi \wedge K_a\psi) \to K_a\varphi$	Prop
4 $(K_a\varphi \wedge K_a\psi) \to K_a(\psi \to (\varphi \wedge \psi))$	Silogismo Hipotético, 3, 2
5 $K_a(\psi \to (\varphi \wedge \psi)) \to (K_a\psi \to K_a(\varphi \wedge \psi))$	K
6 $(K_a\varphi \wedge K_a\psi) \to (K_a\psi \to K_a(\varphi \wedge \psi))$	SH, 4, 5
7 $((K_a\varphi \wedge K_a\psi) \to (K_a\psi \to K_a(\varphi \wedge \psi))) \to ((K_a\varphi \wedge K_a\psi) \to K_a(\varphi \wedge \psi))$	Prop
8 $(K_a\varphi \wedge K_a\psi) \to K_a(\varphi \wedge \psi)$	MP, 6,7

   La regla de Silogismo Hipotético SH es fácilmente derivable en **K** aplicando MP a una instancia de $(Prop)$ adecuada.

A continuación abordamos la definición de derivación o demostración *con premisas* en un sistema de la lógica epistémica. Dicha definición requiere un poco de cuidado debido a la presencia de las *reglas de necesitación* en el sistema.

Definición 2.11. *Dado un operador modal arbitrario, $\Box$, una regla de inferencia Ru es llamada una regla de necesitación para $\Box$ si es de la forma "De φ, se obtiene $\Box\varphi$". Sea $\boldsymbol{X}$ una axiomatización cualquiera generada por los axiomas $Ax_1, Ax_2, ..., Ax_n$ y las reglas $Ru_1, Ru_2, ..., Ru_k$. Se define la clausura bajo las reglas de necesitación de $\boldsymbol{X}$, $Cl_{Nec}(\boldsymbol{X})$, como el menor conjunto tal que:*

1. *$\{Ax_1, ..., Ax_n\} \subseteq Cl_{Nec}(\boldsymbol{X})$; y*
2. *para toda $\psi \in Cl_{Nec}(\boldsymbol{X})$ y para cualquier regla de necesitación para $\Box$ de $\boldsymbol{X}$, se tiene que $\Box\psi \in Cl_{Nec}(\boldsymbol{X})$.*

Definición 2.12 (Derivaciones con premisas). *Sea $\boldsymbol{X}$ una axiomatización con axiomas $Ax_1, Ax_2, ..., Ax_n$ y reglas $Ru_1, Ru_2, ..., Ru_k$, donde cada regla $Ru_j (j \leq k)$ es de* *la forma "De $\varphi_1, ..., \varphi_{j_{ar}}$ se obtiene φ_j", y su respectiva clausura bajo regla de necesitación, $Cl_{Nec}(\boldsymbol{X})$. Sea también Γ un conjunto de fórmulas epistémicas (premisas). Una derivación para una fórmula φ a partir de Γ es una sucesión finita $\varphi_1, ..., \varphi_m$ de fórmulas tal que:*

1. *$\varphi_m = \varphi$;*
2. *cada φ_i en la sucesión es:*
   a) *un elemento de $Cl_{Nec}(\boldsymbol{X})$, o bien*
   b) *un elemento de Γ, o bien*
   c) *el resultado de aplicar una regla $Ru_j(j \leq k)$ que NO es de necesitación, a j_{ar} fórmulas que aparecen previamente en la sucesión.*

Si existe una derivación a partir de Γ para φ en $\boldsymbol{X}$, escribiremos $\Gamma \vdash_{\boldsymbol{X}} \varphi$.

La diferencia entre la Definición 2.10 y la Definición 2.12 se encuentra en las fórmulas de partida que utilizamos. Mientras en la primera definición partimos de los axiomas de la axiomatización **X**, en la segunda a esto le añadimos un conjunto de fórmulas externo Γ, que podremos tomar como premisas. Estas fórmulas externas deben quedar fuera del conjunto de clausura bajo regla de necesitación de **X**: si una fórmula no pertenece a los axiomas prefijados, no podemos aplicar una regla de necesitación para una fórmula que dependa de ella. Esta restricción es importante porque, como veremos, en caso contrario no se verifica el Teorema de la Deducción.

Teorema 2.1 (Teorema de la Deducción). *Sea $\boldsymbol{X}$ una axiomatización que extiende al sistema **K** con axiomas $Ax_1, Ax_2, ..., Ax_n$ y con las mismas reglas de inferencia: (MP) y (Nec). Sea Γ un conjunto de fórmulas y sean φ, ψ fórmulas.*

$$\text{Si } \varphi, \Gamma \vdash_{\boldsymbol{X}} \psi \text{ , entonces } \Gamma \vdash_{\boldsymbol{X}} \varphi \to \psi.$$

Demostración. Puesto que $\varphi, \Gamma \vdash_{\mathbf{X}} \psi$, existe una derivación a partir de las premisas $\Gamma \cup \{\varphi\}$ de la fórmula ψ en el sistema **X**. La prueba es por inducción en la longitud de una tal derivación. Distinguimos varios casos.

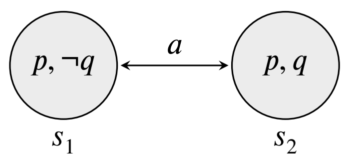
1. Si $\psi = \varphi$, entonces $\varphi \to \psi$ es una instancia de una tautología proposicional y el resultado se sigue de $(\boldsymbol{Prop})$.
2. Supongamos que $\psi \in \Gamma$. Entonces, $\Gamma \vdash_{\mathbf{X}} \psi$, y $\Gamma \vdash_{\mathbf{X}} \psi \to (\varphi \to \psi)$ por $(\boldsymbol{Prop})$. Una aplicación de Modus Ponens nos da $\Gamma \vdash_{\mathbf{X}} \varphi \to \psi$.
3. Supongamos que $\psi \in Cl_{Nec}(\mathbf{X})$. Entonces, $\Gamma \vdash_{\mathbf{X}} \psi$, y $\Gamma \vdash_{\mathbf{X}} \psi \to (\varphi \to \psi)$ por $(\boldsymbol{Prop})$. Una aplicación de Modus Ponens nos da $\Gamma \vdash_{\mathbf{X}} \varphi \to \psi$.
4. Supongamos la última regla que utilizamos en la prueba de ψ es Modus Ponens. (Nótese que por la definición dada de derivación con premisas, las reglas de necesitación no se pueden usar fuera del conjunto $Cl_{Nec}(\mathbf{X})$). Entonces, debe

existir una fórmula θ tal que las fórmulas θ y $\theta \to \psi$ aparecen previamente en la derivación de ψ. Por hipótesis de inducción, tenemos que $\Gamma \vdash_{\mathbf{X}} \varphi \to \theta$ y $\Gamma \vdash_{\mathbf{X}} \varphi \to (\theta \to \psi)$. Ahora bien, por los axiomas (***Prop***) se tiene que $\Gamma \vdash_{\mathbf{X}} (\varphi \to \theta) \to ((\varphi \to (\theta \to \psi)) \to (\varphi \to \psi))$ al ser una instancia de una tautología proposicional. Una aplicación de Modus Ponens nos proporciona $\Gamma \vdash_{\mathbf{X}} (\varphi \to (\theta \to \psi)) \to (\varphi \to \psi)$; y una segunda aplicación de dicha regla nos da $\Gamma \vdash_{\mathbf{X}} \varphi \to \psi$, como queríamos demostrar.

Esto concluye la prueba. ∎

El Teorema de la Deducción se verifica tanto para **K** como para la lógica epistémica multiagente **S5**. Ahora bien, si no consideramos la definición de derivación con premisas y permitimos aplicar la regla de necesitación a cualquier fórmula de una prueba, no se verificaría dicho Teorema. Es llamativo que este resultado ha generado históricamente cierta controversia en el campo. Algunos autores han argumentado que el Teorema de la Deducción es correcto para las lógicas modales, mientras que en otras publicaciones se ha argumentado que dicho Teorema no es cierto para las lógicas modales. Dicha controversia se analiza en detalle en el artículo Hakli-Negri ([HS12]). En él se comenta que el estudio del Teorema de la Deducción comenzó con una prueba de una versión primitiva del Teorema realizada por Jacques Herbrand en su tesis doctoral para axiomatizaciones de lógica proposicional y de primer orden. Más tarde, P. Bernays da de forma explícita el teorema, dándole el nombre con el que ahora lo conocemos y una prueba en [HB34]. En el contexto modal, en Fagin et al.(1995) se observa que el teorema no se cumple si no se añade una definición de derivaciones con premisas ([Fag+95]). Por su parte, en [Smo84], Smorynski expone un teorema de deducción sin restricciones para las derivaciones que no usan la regla de necesitación, sugiere obviar dicha regla y dar como axioma, cada axioma precedido del operador.

Para aclarar este argumento, veamos el siguiente ejemplo. Si tenemos como premisa $p \wedge q$, ignoramos la Definición 2.12 y aplicamos la regla de necesitación sobre $p \wedge q$, nos resulta la afirmación $p \wedge q \vdash K_a(p \wedge q)$. Pero consideremos el siguiente ejemplo:



Entonces, en el estado s_2 se tiene $p \wedge q$ y no se tiene $K_a(p \wedge q)$. Por tanto, se incumpliría dicho Teorema. Esto ha ocurrido por aplicar la regla de necesitación sobre una premisa, si hubieramos considerado la Definición 2.12, el Teorema se habría verificado.

2.2.2 Los sistemas **T**, **S4** y **S5**

Denotamos por $\mathbf{X} + \varphi$ al sistema de axiomas constituido por los axiomas de $\mathbf{X}$ y φ (sin variar las reglas de inferencia). Denotamos por $\mathbf{X} - \varphi$ al sistema de axiomas constituido por los axiomas de $\mathbf{X}$ eliminado φ si estuviera (sin variar las reglas de inferencia). También diremos que dos fórmulas A y B son equivalentes con respecto a $\mathbf{X}$ si $\vdash_{(\mathbf{X}-A)+B} A$ y $\vdash_{(\mathbf{X}-B)+A} B$.

Vamos a definir tres axiomas que van a ser importantes, y nos servirán para definir tres axiomatizaciones partiendo de $\mathbf{K}$:

- $K_a\varphi \rightarrow \varphi$ (Axioma T)
- $K_a\varphi \rightarrow K_aK_a\varphi$ (introspección positiva)
- $\neg K_a\varphi \rightarrow K_a\neg K_a\varphi$ (introspección negativa)

A los axiomas de introspección positiva y negativa los llamaremos también, respectivamente, Axioma 4 y Axioma 5.

Definición 2.13 (Los sistemas T, S4 y S5). *Sea $\boldsymbol{K}$ la axiomatización vista en la Definición 2.9. Dados los axiomas $T, 4, 5$ que hemos introducido justo arriba, definimos los siguientes sistemas lógicos:*

- $\boldsymbol{T} = \boldsymbol{K} + T$
- $\boldsymbol{S4} = \boldsymbol{T} + 4$
- $\boldsymbol{S5} = \boldsymbol{S4} + 5$

De los sistemas que acabos de introducir, el más importante para el presente trabajo será el sistema **S5**, pues tradicionalmente es utilizado como base de los distintos sistemas lógicos para modelar la lógica epistémica multiagente.

Un resultado fundamental de la lógica epistémica establece que los sistemas lógicos que acabamos de introducir capturan, exactamente, aquellas fórmulas válidas en la correspondiente clase de modelos de Kripke. Más formalmente:

Teorema 2.2. *Dados las axiomatizaciones **K**, **T**, **S4**, **S5**, las hacemos corresponder con las clases semánticas $\mathcal{K}, \mathcal{T}, \mathcal{S}4, \mathcal{S}5$, respectivamente. Se cumplen las siguientes propiedades:*

1. *(Adecuación y Completitud)*
   Los sistema de axiomas son adecuados y completos con respecto a su clase semántica. Es decir, para cada fórmula φ, tenemos $\vdash_{\mathbf{X}} \varphi$ si y solo si $\mathcal{X} \models \varphi$, siendo $\mathbf{X}$ una de estos sistemas de axiomas y $\mathcal{X}$ su clase semántica correspondiente.
2. *(Propiedad de Modelos Finitos)*
   Una fórmula φ es satisfacible en una clase $\mathcal{X}$, si y solo si es satisfacible en un modelo finito de esa clase.
3. *(Decidilidad)*
   Las clases anteriores son decidibles, es decir, existe un proceso de decisión que determina, en un tiempo finito, para cada φ, si es satisfacible en $\mathcal{X}$ o no.

En la siguiente sección daremos una prueba detallada del teorema de completitud para la lógica **S5**, un resultado teórico fundamental para es estudio de este sistema.

A la vista de nuestra definición de derivaciones con premisas, podemos extender los teoremas de completitud para los sitemas anteriores a la siguiente versión *fuerte*.

Definición 2.14. *Dados una clase de modelos de Kripke C, un conjunto de fórmulas Γ, y una fórmula φ:*
Diremos que el sistema $\mathbf{X}$ es fuertemente adecuado con respecto a C, si:

$$\Gamma \vdash_{\mathbf{X}} \varphi \Rightarrow \forall M \in C, s \in M : M, s \models \Gamma \text{ implica } M, s \models \varphi.$$

Diremos que $\mathbf{X}$ es fuertemente completo con respecto a C, si:

$$\Gamma \vdash_{\mathbf{X}} \varphi \Leftarrow \forall M \in C, s \in M : M, s \models \Gamma \text{ implica } M, s \models \varphi.$$

Teorema 2.3 (Completitud fuerte). *Los sistemas de axiomas **K**, **T**, **S4**, **S5** son fuertemente adecuados y fuertemente completos con respecto a las clases $\mathcal{K}, \mathcal{T}, \mathcal{S}4, \mathcal{S}5$, respectivamente.*

2.3 Completitud para S5

En esta sección vamos a probar la completitud del sistema **S5**, es decir, probaremos que toda fórmula válida en cualquier modelo epistémico es un teorema de **S5**. Para ello, nos hará falta definir algunos conceptos previos, como los conjuntos de fórmulas que se denominan consistentes maximales, o el modelo canónico, cuyo dominio estará formado por estos conjuntos de fórmulas maximales.

La completitud es uno de los temas principales en lógica. Generalmente, se busca asegurar que la noción semántica de validez coincida con la noción teórica de la prueba de validez. Existe una gran variedad de pruebas de completitud, aunque suelen seguir el mismo esquema. La prueba de completitud para **S5** que seguiremos en este trabajo es la presentada por Blackburn, de Rijke y Venema en su libro "Modal logic" ([BRV01]) para la completitud de una lógica modal normal.

Definición 2.15 (Conjuntos consistentes). *Sean $\Gamma \subseteq \mathcal{L}_K$ un conjunto de fórmulas y $\boldsymbol{X}$ un sistema lógico.*

1. *Diremos que Γ es $\boldsymbol{X}$-inconsistente si $\Gamma \vdash_{\boldsymbol{X}} \bot$.*
2. *Diremos que Γ es $\boldsymbol{X}$-consistente si no es $\boldsymbol{X}$-inconsistente, es decir, si no hay ninguna prueba de la inconsistencia $\bot$ a partir de Γ en el sistema $\boldsymbol{X}$.*

Nótese que, por ejemplo, el conjunto $\Gamma = \{K_a p, \neg p\}$ es **K**-consistente pero no es **S5**-consistente. Si el sistema **X** está claro por el contexto, lo omitiremos y diremos simplemente consistente o inconsistente.

Proposición 2.3. Sea $\Gamma \subseteq \mathcal{L}_K$ un conjunto de fórmulas. Entonces

$$\Gamma \vdash_{\mathbf{S5}} \varphi \text{ si y solo si } \Gamma \cup \{\neg\varphi\} \text{ es } \mathbf{S5}\text{-inconsistente.}$$

Demostración. Supongamos primero que $\Gamma \vdash_{\mathbf{S5}} \varphi$. Entonces, es claro que $\Gamma + \neg\varphi \vdash_{\mathbf{S5}} \varphi$ y $\Gamma + \neg\varphi \vdash_{\mathbf{S5}} \neg\varphi$. Pero $\varphi \to (\neg\varphi \to \bot)$ es una tautología proposicional. Luego, aplicando (MP) dos veces, se obtiene $\Gamma + \neg\varphi \vdash_{\mathbf{S5}} \bot$.

Supongamos ahora que $\Gamma \cup \{\neg\varphi\}$ es **S5**-inconsistente, esto es, $\Gamma + \neg\varphi \vdash_{\mathbf{S5}} \bot$. Por el teorema de la Deducción, $\Gamma \vdash_{\mathbf{S5}} \neg\varphi \to \bot$. Ahora bien, $(\neg\varphi \to \bot) \to \varphi$ es una tautología proposicional y el resultado se tiene por (MP).

Proposición 2.4. Sean $\Gamma \subseteq \mathcal{L}_K$ un conjunto **S5**-consistente y φ una fórmula. Entonces, o bien $\Gamma + \varphi$ o bien $\Gamma + \neg\varphi$ es **S5**-consistente.

Demostración. Supongamos lo contrario. Entonces, $\Gamma \vdash_{\mathbf{S5}} \varphi$ y $\Gamma \vdash_{\mathbf{S5}} \neg\varphi$. Ahora bien, $\varphi \to (\neg\varphi \to \bot)$ es una tautología proposicional y se sigue por (MP) que $\Gamma \vdash_{\mathbf{S5}} \bot$. Lo cual está en contradicción con la consistencia de Γ.

Definición 2.16 (Conjuntos maximales). *Diremos que $\Gamma \subseteq \mathbf{L}_K$ es un conjunto **S5**-consistente maximal si:*

1. *Γ es **S5**-consistente, es decir, $\Gamma \nvdash_{S5} \bot$; y*
2. *Γ es maximal, es decir, no existe $\Gamma' \subseteq \mathbf{L}_K$ tal que $\Gamma \subset \Gamma'$ y $\Gamma' \nvdash_{S5} \bot$.*

Definición 2.17 (El modelo canónico). *Sean un conjunto numerable de variables atómicas, P, y un conjunto finito de agentes, A. El modelo canónico para la lógica **S5** basado en P y A, $M^c = \langle S^c, \sim^c, V^c \rangle$, es el modelo de Kripke definido por:*

- *$S^c = \{\Gamma \mid \Gamma$ es **S5**-consistente maximal $\}$.*
- *$\Gamma \sim_a^c \Delta$ sii $\{K_a\varphi \mid K_a\varphi \in \Gamma\} = \{K_a\varphi \mid K_a\varphi \in \Delta\}$ para cada $a \in A$.*
- *$V_p^c = \{\Gamma \in S^c \mid p \in \Gamma\}$ para cada $p \in P$.*

Con estos conceptos, obtenemos varios resultados relevantes, como el Lema de Lindenbaum o el Lema de la verdad, que nos ayudarán próximamente para probar el resultado de completitud en $S5$.

Lema 2.1 (Lindenbaum). *Todo conjunto **S5**-consistente de fórmulas es subconjunto de un conjunto **S5**-consistente maximal.*

Demostración. Sea Δ un conjunto **S5**-consistente de fórmulas cualquiera. Puesto que $\mathcal{L}_K$ contiene un número finito de agentes y el conjunto de variables atómicas es numerable, podemos establecer una enumeración de las fórmulas de $\mathcal{L}_K$ de forma $\{\varphi_n \ : \ n \in \mathbb{N}\}$. Definimos la siguiente sucesión de subconjuntos de fórmulas:

$$
\begin{aligned}
\Gamma_0 &= \Delta \\
\Gamma_{n+1} &= \begin{cases} \Gamma_n \cup \{\varphi_n\} & \text{si } \Gamma_n \cup \{\varphi_n\} \text{ es } \mathbf{S5}\text{-consistente} \\ \Gamma_n & \text{en otro caso} \end{cases}
\end{aligned}
$$

Sea $\Gamma = \bigcup_{n\in\mathbb{N}} \Gamma_n$. Tenemos que $\Delta \subseteq \Gamma$. Falta probar que Γ es maximal y **S5**-consistente.

- **Maximal.** Tomamos φ_n una fórmula cualquiera tal que $\varphi_n \notin \Gamma$. Entonces, $\varphi_n \notin \Gamma_{n+1}$. Por tanto, $\Gamma_n \cup \{\varphi_n\}$ es **S5**-inconsistente y, en consecuencia, $\Gamma \cup \{\varphi_n\}$ también lo es. Entonces, no existe Γ' **S5**-consistente tal que $\Gamma \subset \Gamma'$.
- **Consistente.** Probaremos por inducción sobre $n \geq 0$ que cada Γ_n es un conjunto **S5**-consistente. Esto es suficiente porque, si Γ fuese **S5**-inconsistente, también lo sería una parte finita de Γ y, en particular, algún Γ_n sería **S5**-inconsistente.
  En el caso base, tenemos por hipótesis que $\Gamma_0 = \Delta$ es **S5**-consistente. Damos por hecho que Γ_n es consistente. Entonces, por la construcción se sigue que, o bien $\Gamma_{n+1} = \Gamma_n \cup \{\varphi_n\}$ (que es **S5**-consistente por definición de la sucesión), o bien $\Gamma_{n+1} = \Gamma_n$ que es **S5**-consistente por hipótesis de inducción. ∎

Por tanto, el modelo canónico está bien definido y su dominio no es vacío. Más aún, existe una cantidad no numerable (de hecho, de la misma cardinalidad que el conjunto $\mathbb{R}$ de los números reales) de conjuntos consistentes maximales; luego, el dominio del modelo canónico es un conjunto *grande* (no numerable) que contiene tantos estados como números reales. Los conjuntos maximales consistentes verifican una serie de propiedades que nos servirán en la prueba del Lema de la verdad, que a su vez será necesario para demostrar el teorema de completitud.

Lema 2.2. *Sean Γ y Δ dos conjuntos **S5**-consistentes maximales. Entonces:*

1. *Γ es deductivamente cerrado, es decir, cerrado bajo **S5**-derivaciones.*
2. *$\varphi \in \Gamma$ si y solo si $\neg\varphi \notin \Gamma$.*
3. *$(\varphi \wedge \psi) \in \Gamma$ si y solo si $\varphi \in \Gamma$ y $\psi \in \Gamma$*
4. *$(\varphi \vee \psi) \in \Gamma$ si y solo si $\varphi \in \Gamma$ o $\psi \in \Gamma$.*
5. *$(\varphi \to \psi) \in \Gamma$ si y solo si $\varphi \notin \Gamma$ o $\psi \in \Gamma$.*
6. *$\Gamma \sim_a^c \Delta$ si y solo si $\{K_a\varphi \mid K_a\varphi \in \Gamma\} \subseteq \Delta$.*
7. *$\{K_a\varphi \mid K_a\varphi \in \Gamma\} \vdash_{\mathbf{S5}} \psi$ si y solo si $\{K_a\varphi \mid K_a\varphi \in \Gamma\} \vdash_{\mathbf{S5}} K_a\psi$.*

Demostración. Dados Γ y Δ, probemos cada apartado.

1. Suponemos que $\Gamma \vdash_{\mathbf{S5}} \varphi$. Entonces, puesto que Γ es **S5**-consistente, $\Gamma \cup \{\varphi\}$ también lo es. Además, por ser Γ maximal, $\varphi \in \Gamma$.
2. Por un lado, por ser Γ consistente, de $\varphi \in \Gamma$ se sigue que $\neg\varphi \notin \Gamma$.
Por otro lado, suponemos que $\neg\varphi \notin \Gamma$. Por ser Γ maximal, $\Gamma \cup \{\neg\varphi\} \vdash_{\mathbf{S5}} \bot$. Por tanto, $\Gamma \vdash_{\mathbf{S5}} \varphi$. En consecuencia, $\varphi \in \Gamma$ ya que Γ es deductivamente cerrado.
3. Supongamos primero que $(\varphi \wedge \psi) \in \Gamma$. Entonces, $\Gamma \vdash_{\mathbf{S5}} \varphi$ y $\Gamma \vdash_{\mathbf{S5}} \psi$. En efecto, basta aplicar (MP) pues $(\varphi \wedge \psi) \to \varphi$ y $(\varphi \wedge \psi) \to \psi$ son sendas instancias de tautologías proposicionales. Puesto que Γ es deductivamente cerrado, $\varphi \in \Gamma$ y $\psi \in \Gamma$.
Por otro lado, supongamos que $\varphi \in \Gamma$ y $\psi \in \Gamma$. Entonces, $\Gamma \vdash_{\mathbf{S5}} (\varphi \wedge \psi)$ pues $\varphi \to (\psi \to (\varphi \wedge \psi))$ es una instancia de una tautología proposicional. Puesto que Γ es deductivamente cerrado, $(\varphi \wedge \psi) \in \Gamma$.
4. Supongamos primero que $(\varphi \vee \psi) \in \Gamma$. Por reducción al absurdo, supongamos que $\varphi \notin \Gamma$ y $\psi \notin \Gamma$. Razonando como en 2, se tiene que $\neg\varphi \in \Gamma$ y $\neg\psi \in \Gamma$. Ahora bien, $(\varphi \vee \psi) \to (\neg\varphi \to (\neg\psi \to \bot))$ es una instancia de una tautología proposicional. Aplicando (MP) tantas veces como se a necesario, se obtiene $\Gamma \vdash_{\mathbf{S5}} \bot$, una contradicción pues Γ es consistente.
Por otro lado, supongamos, por ejemplo, que $\varphi \in \Gamma$. Entonces, es fácil comprobar que $\Gamma \vdash_{\mathbf{S5}} (\varphi \vee \psi)$ pues $\varphi \to (\varphi \vee \psi)$ es una tautología. Puesto que Γ es deductivamente cerrado, $(\varphi \vee \psi) \in \Gamma$.

5. Similar a la prueba de 4.
6. Fijemos $a \in A$. Por un lado, la implicación $\Gamma \sim_a^c \Delta \Rightarrow \{K_a\varphi \mid K_a\varphi \in \Gamma\} \subseteq \Delta$ se obtiene de forma directa por la definición de $\sim^c$.
   Para la implicación opuesta, suponemos que $\{K_a\varphi \mid K_a\varphi \in \Gamma\} \subseteq \Delta$. Suponemos además que $K_a\varphi \in \Delta$. Entonces, por el apartado 2 tenemos que $\neg K_a\varphi \notin \Delta$ y, por tanto, $K_a\neg K_a\varphi \notin \Delta$ (Δ es cerrado bajo **S5**-derivaciones). Además, por nuestra suposición y la definición de $\sim_a^c$, $K_a\neg K_a\varphi \notin \Gamma$, y por la introspección negativa $\neg K_a\varphi \notin \Gamma$. Por tanto, aplicando de nuevo el apartado 2, $K_a\varphi \in \Gamma$.
7. De derecha a izquierda se demuestra fácilmente usando que $\vdash_{\mathbf{S5}} K_a\psi \rightarrow \psi$.
   De izquierda a derecha, suponemos que $\{K_a\varphi \mid K_a\varphi \in \Gamma\} \vdash_{\mathbf{S5}} \psi$. Por la definición de derivación con premisas, existe un subconjunto finito $\Gamma' \subseteq \{K_a\varphi \mid K_a\varphi \in \Gamma\}$ tal que $\vdash_{\mathbf{S5}} \bigwedge \Gamma' \rightarrow \psi$. Además, aplicando la regla de necesitación (Nec) y el axioma de distribución, se obtiene que $\vdash_{\mathbf{S5}} K_a \bigwedge \Gamma' \rightarrow K_a\psi$. Por tanto, $\vdash_{\mathbf{S5}} \bigwedge\{K_a\chi \mid \chi \in \Gamma'\} \rightarrow K_a\psi$ (nótese que K_a y $\wedge$ conmutan demostrablemente en **S5**). Puesto que $\vdash_{\mathbf{S5}} K_a\varphi \rightarrow K_aK_a\varphi$, se tiene que $\{K_a\varphi \mid \varphi \in \Gamma\} \vdash_{\mathbf{S5}} \bigwedge\{K_a\chi \mid \chi \in \Gamma'\}$. Luego, $\{K_a\varphi \mid K_a\varphi \in \Gamma\} \vdash_{\mathbf{S5}} K_a\psi$.

Ahora podemos demostrar el Lema de la verdad:

Lema 2.3 (Lema de la verdad). *Para cada $\varphi \in \mathcal{L}_K$ y para cada conjunto **S5**-consistente maximal Γ, se tiene que:*

$$\varphi \in \Gamma \text{ si y solo si } (M^c, \Gamma) \models \varphi.$$

Demostración. Vamos a demostrarlo por inducción sobre φ.
Caso Base. Suponemos que φ es una variable proposicional, p. Por definición de V^c, se tiene que $p \in \Gamma$ si y solo si $\Gamma \in V_p^c$, lo cual es equivalente a decir que $(M^c, \Gamma) \models p$ por la propia definición de la semántica de los modelos de Kripke.
Hipótesis de inducción: "Para cada conjunto **S5**-consistente maximal Γ y dados φ y ψ, se tiene que $\varphi \in \Gamma$ si y solo si $(M^c, \Gamma) \models \varphi$, y $\psi \in \Gamma$ si y solo si $(M^c, \Gamma) \models \psi$".
Paso de inducción. Distinguimos varios casos:

- **El caso para $\neg\varphi$.** Si tenemos $\neg\varphi \in \Gamma$, entonces, por el Lema 2.2, $\varphi \notin \Gamma$, que, por la hipótesis de inducción, es equivalente a $(M^c, \Gamma) \not\models \varphi$; lo cual es equivalente a $(M^c, \Gamma) \models \neg\varphi$ por la propia definición de la semántica de los modelos de Kripke.
- **El caso para $\varphi \wedge \psi$.** Nótese que $(\varphi \wedge \psi) \in \Gamma$ es equivalente a $\varphi \in \Gamma$ y $\psi \in \Gamma$, por el Lema 2.2. Con esto y aplicando la hipótesis de inducción, tenemos que $(\varphi \wedge \psi) \in \Gamma$ es equivalente a $(M^c, \Gamma) \models \varphi$ y $(M^c, \Gamma) \models \psi$; lo cual es equivalente a decir que $(M^c, \Gamma) \models \varphi \wedge \psi$.

- **El caso para $\varphi \vee \psi$.** Nótese que $(\varphi \vee \psi) \in \Gamma$ es equivalente a $\varphi \in \Gamma$ o $\psi \in \Gamma$, por el Lema 2.2. Con esto y aplicando la hipótesis de inducción, tenemos que $(\varphi \vee \psi) \in \Gamma$ es equivalente a $(M^c, \Gamma) \vDash \varphi$ o $(M^c, \Gamma) \vDash \psi$; lo cual es equivalente a decir que $(M^c, \Gamma) \vDash \varphi \vee \psi$.
- **El caso para $\varphi \to \psi$.** Nótese que $(\varphi \to \psi) \in \Gamma$ es equivalente a $\varphi \notin \Gamma$ o $\psi \in \Gamma$, por el Lema 2.2. Con esto y aplicando la hipótesis de inducción, tenemos que $(\varphi \to \psi) \in \Gamma$ es equivalente a $(M^c, \Gamma) \nvDash \varphi$ o $(M^c, \Gamma) \vDash \psi$; lo cual es equivalente a decir que $(M^c, \Gamma) \vDash \varphi \to \psi$.
- **El caso para $K_a\varphi$.** En primer lugar probaremos la implicación: $K_a\varphi \in \Gamma \Rightarrow (M^c, \Gamma) \vDash \varphi$. Supongamos que $K_a\varphi \in \Gamma$. Tomamos un conjunto **S5**-consistente maximal cualquiera Δ tal que $\Gamma \sim_a^c \Delta$. Entonces, $K_a\varphi \in \Delta$ por la definición de $\sim_a^c$. Por el Lema 2.2, como $\vdash_{\mathbf{S5}} K_a\varphi \to \varphi$ y Δ es deductivamente cerrado, se tiene que $\varphi \in \Delta$. Por la hipótesis de inducción, se tiene pues que $(M^c, \Delta) \vDash \varphi$. Ahora bien, puesto que Δ era fijo pero arbitrario, hemos demostrado que $(M^c, \Delta) \vDash \varphi$ para todo Δ tal que $\Gamma \sim_a^c \Delta$ y, por la semántica de los modelos de Kripke, ello es equivalente a decir que $(M^c, \Gamma) \vDash K_a\varphi$, como queríamos demostrar.
  Por otro lado, vamos a probar la implicación contraria, esto es: $(M^c, \Gamma) \vDash K_a\varphi \Rightarrow K_a\varphi \in \Gamma$. Supongamos que $(M^c, \Gamma) \vDash K_a\varphi$. Entonces, $(M^c, \Delta) \vDash \varphi$ para todo Δ tal que $\Gamma \sim_a^c \Delta$. Vamos a probar que $\{K_a\chi \mid K_a\chi \in \Gamma\} \vdash_{\mathbf{S5}} \varphi$. Lo haremos por reducción al absurdo. Supongamos, por deducción al absurdo, que el conjunto

  $$\Lambda = \{\neg\varphi\} \cup \{K_a\chi \mid K_a\chi \in \Gamma\}$$

  es **S5**-consistente. Por el Lema de Lindenbaum, Λ puede extenderse a un conjunto **S5**-consistente maximal, al que llamaremos Θ. Entonces, es evidente que $\{K_a\chi | K_a\chi \in \Gamma\} \subseteq \Theta$. Por lo tanto, por el Lema 2.2, tenemos que $\Gamma \sim_a^c \Theta$. Puesto que $\neg\varphi \in \Theta$ y Θ es **S5**-consistente maximal, $\varphi \notin \Theta$. Por la hipótesis de inducción, tenemos que $(M^c, \Theta) \nvDash \varphi$. Ahora bien, esto contradice que $(M^c, \Delta) \vDash \varphi$ para todo Δ tal que $\Gamma \sim_a^c \Delta$. Por lo tanto, $\{K_a\chi \mid K_a\chi \in \Gamma\} \vdash_{\mathbf{S5}} \varphi$. Volviendo a aplicar el Lema 2.2, tenemos que $\Gamma \vdash_{\mathbf{S5}} K_a\varphi$. Queda probado por tanto que $K_a\varphi \in \Gamma$ puesto que Γ es deductivamente cerrado. ∎

Antes de probar la completitud de **S5**, necesitamos demostrar una propiedad más del modelo canónico para **S5**: el modelo canónico es un modelo epistémico, es decir, las relaciones de accesibilidad $\sim_a^c$ son, después de todo, relaciones de equivalencia. Este es el objeto del siguiente lema.

Lema 2.4. *El modelo canónico para **S5** es reflexivo, transitivo y euclídeo (y, por tanto, es un modelo epistémico).*

Demostración. La prueba se sigue de la definición de $\sim_a^c$:

$$\Gamma \sim_a^c \Delta \text{ si y solo si } \{K_a\varphi \mid K_a\varphi \in \Gamma\} = \{K_a\varphi \mid K_a\varphi \in \Delta\}.$$

- **Reflexivo.** $\Gamma \sim_a^c \Gamma$. Es evidente, ya que
$$\{K_a\varphi \mid K_a\varphi \in \Gamma\} = \{K_a\varphi \mid K_a\varphi \in \Gamma\}.$$
- **Transitivo.** Teniendo $\Gamma \sim_a^c \Delta$ y $\Delta \sim_a^c \Theta$, debemos llegar a $\Gamma \sim_a^c \Theta$. A la vista de la definición de $\sim_a^c$, se tiene que
$$K_a\varphi \in \Gamma \Leftrightarrow K_a\varphi \in \Delta\,,\ K_a\varphi \in \Delta \Leftrightarrow K_a\varphi \in \Theta$$
Entonces, $K_a\varphi \in \Gamma \Leftrightarrow K_a\varphi \in \Theta$, es decir, $\{K_a\varphi \mid K_a\varphi \in \Gamma\} = \{\varphi \mid K_a\varphi \in \Delta\}$ y por tanto $\Gamma \sim_a^c \Theta$.
- **Euclídeo.** Ahora tenemos que probar que si $\Gamma \sim_a^c \Delta$ y $\Gamma \sim_a^c \Theta$, entonces $\Delta \sim_a^c \Theta$. Es fácil probarlo siguiendo el mismo razonamiento anterior.

Estamos preparados por fin para dar la prueba del teorema de completitud para la lógica **S**5.

Teorema 2.4 (Adecuación y Completitud de S5). *Para cada $\varphi \in \mathcal{L}_K$,*

$$\mathcal{S}5 \vDash \varphi \textit{ si y solo si } \vdash_{\mathbf{S5}} \varphi.$$

Demostración.
(Adecuación): Supongamos que $\vdash_{\mathbf{S5}} \varphi$. Vamos a probar por inducción en la longitud de una **S5**-derivación de la fórmula φ que φ es válida para la clase $\mathcal{S}5$:

- **Caso base:** Sea A un axioma de **S5**. Es fácil comprobar que A es válido en la clase de los modelos epistémicos teniendo en cuenta que, para todo agente a, la relación R_a es de equivalencia.
- **Paso de inducción:**
  - **Caso (MP):** Dado $\vdash_{\mathbf{S5}} \varphi \Rightarrow \mathcal{S}5 \vDash \varphi$ y dado $\vdash_{\mathbf{S5}} \varphi \to \psi \Rightarrow \mathcal{S}5 \vDash \varphi \to \psi$, tenemos que probar $\vdash_{\mathbf{S5}} \psi \Rightarrow \mathcal{S}5 \vDash \psi$.
    Tenemos:
    $$\mathcal{S}5 \vDash \varphi \text{ si y solo si } \forall M \ \mathbf{S5}\text{-modelo y } s \text{ cualquiera}\ , M, s \vDash \varphi$$
    Y por otro lado:
    $$\mathcal{S}5 \vDash \varphi \to \psi \text{ si y solo si } \forall M \ \mathbf{S5}\text{-modelo y } s \text{ cualquiera}\ , M, s \vDash \neg\varphi \vee M, s \vDash \psi$$
    Como consecuencia tenemos $\forall M$ **S5**-modelo y s cualquiera , $M, s \vDash \psi$, y esto se tiene si y solo si $\mathcal{S}5 \vDash \psi$.

- **Caso (Nec):** Sea $\vdash_{\mathbf{S5}} \varphi \Rightarrow \mathcal{S}5 \models \varphi$, tenemos que probar $\vdash_{\mathbf{S5}} K_a\varphi \Rightarrow \mathcal{S}5 \models K_a\varphi$. Basta probar $\forall M$ **S5**-modelo y s cualquiera , $M, s \models K_a\varphi$ y esto se tiene si y solo si $\forall t$ tal que $(s,t) \in R_a$, se tiene $M, t \models \varphi$. Esto se verifica pues

$$\mathcal{S}5 \models \varphi \text{ si y solo si } \forall M \text{ } \mathbf{S5}\text{-modelo y } t \text{ cualquiera } , M, t \models \varphi.$$

(Completitud): Supongamos que $\mathcal{S}5 \models \varphi$. Esto es, para todo modelo M de **S5** y para todo estado $s \in M$ se tiene que $M, s \models \varphi$. Tenemos que probar que $\vdash_{\mathbf{S5}} \varphi$. Lo haremos por reducción al absurdo. Supongamos que $\nvdash_{\mathbf{S5}} \varphi$. Entonces, tendremos que $\{\neg\varphi\}$ es un conjunto **S5**-consistente. Aplicando el Lema de Lindenbaum, $\{\neg\varphi\}$ puede extenderse a un conjunto **S5**-consistente maximal, Γ.
Por el lema de la Verdad, se tiene que $(M^c, \Gamma) \models \neg\varphi$. Pero el Lema 2.4 muestra que el modelo canónico M^c es un **S5**-modelo. Por tanto, hemos encontrado un modelo de **S5** y un estado $\Gamma \in M^c$ tal que $(M^c, \Gamma) \not\models \varphi$. Luego, $\mathcal{S}5 \not\models \varphi$, lo cual contradice nuestra hipótesis. ∎

De la prueba del teorema anterior se sigue el siguiente corolario.

Corolario 2.1. Sea $\varphi \in \mathcal{L}_K$. Los siguientes son equivalentes:

1. $\{\varphi\}$ es **S5**-consistente.
2. φ es satisfacible en el modelo canónico M^c.
3. φ es $\mathcal{S}5$-satisfacible.

Esto es, si una fórmula epistémica puede satisfacerse en algún modelo arbitrario de **S5**, entonces puede también satisfacerse en el modelo canónico. Este resultado también es cierto para un conjunto de fórmulas $\Gamma \subseteq \mathcal{L}_K$.

Otra consecuencia importante del Teorema de Completitud es la *compacidad* para la clase $\mathcal{S}5$. Un conjunto de fórmulas Γ es $\mathcal{X}$-satisfacible si existe un modelo M de $\mathcal{X}$ y un estado $s \in M$ tal que $M, s \models \varphi$ para toda $\varphi \in \Gamma$.

Teorema 2.5 (Compacidad de $\mathcal{S}5$). *Sea $\Gamma \subseteq \mathcal{L}_K$ un conjunto de fórmulas. Si cada subconjunto finito de Γ es $\mathcal{S}5$-satisfacible, Γ también lo es.*

Demostración. Supongamos que cada subconjunto finito de Γ es $\mathcal{S}5$-satisfacible. Entonces, Γ es **S5**-consistente. En caso contrario, existiría algún subconjunto finito $\Gamma' \subseteq \Gamma$ tal que $\Gamma' \vdash_{\mathbf{S5}} \perp$ y, en consecuencia, Γ' no sería **S5**-consistente. Por el lema de Lindenbaum, Γ puede extenderse a un conjunto **S5**-consistente maximal, Δ. Ahora bien, por el lema de la Verdad, $(M^c, \Delta) \models \varphi$ para toda $\varphi \in \Gamma$. Luego, Γ es $\mathcal{S}5$-satisfacible. ∎

Por último, usando la Proposición 2.3, podemos demostrar que, de hecho, la lógica **S5** es *fuertemente* completa.

Teorema 2.6 (Adecuación y Completitud fuerte de S5). *Sean $\Gamma \subseteq \mathcal{L}_K$ y $\varphi \in \mathcal{L}_K$. Son equivalentes:*

1. $\Gamma \vdash_{S5} \varphi$.
2. *Para cada M modelo de **S5** y cada $s \in M$: $M, s \models \Gamma \Rightarrow M, s \models \varphi$.*

Demostración. Probaremos la equivalencia como dos implicaciones por separado:

- **De 1 a 2:** (Adecuación) Suponemos $\Gamma \vdash_{\mathbf{S5}} \varphi$. Sea M un modelo y s un estado, cualesquiera, tales que $M, s \models \Gamma$, debemos probar que $M, s \models \varphi$. Por inducción sobre la longitud de una **S5**-derivación a partir de Γ de la fórmula φ, tal y como hicimos en el Teorema 2.4 se prueba que φ es válida en (M, s) un estado arbitrario.
- **De 2 a 1:** (Completitud) Suponemos que para todo M modelo de **S5** y para cada $s \in M$ se tiene $M, s \models \Gamma$ implica $M, s \models \varphi$. Tenemos que probar $\Gamma \vdash_{\mathbf{S5}} \varphi$. Lo haremos por reducción al absurdo, para ello, suponemos $\Gamma \nvdash_{\mathbf{S5}} \varphi$. Entonces, $\{\neg\varphi\} \cup \Gamma$ es **S5**-consistente. Aplicando el Lema de Lindenbaum, $\{\neg\varphi\} \cup \Gamma$ puede extenderse a un conjunto **S5**-consistente maximal, Δ. Por el Lema de la verdad, $(M^c, \Delta) \models \neg\varphi$ y $(M^c, \Delta) \models \Gamma$. Pero, por el Lema 2.4, M^c es un **S5**-modelo. Lo cual contradice nuestra hipótesis.

Esto concluye la prueba. ∎

2.4 Expresividad

En esta sección vamos a estudiar la expresividad del lenguaje de la lógica epistémica $\mathcal{S}5$. Para ello vamos a utilizar herramientas para comparar dos lenguajes y poder decidir si son o no igualmente expresivos. Pero antes, vamos a definir algunos conceptos previos para aclarar qué significa que un lenguaje sea más expresivo que otro.

2.4.1 Conceptos básicos

En primer lugar vamos a definir el concepto de equivalencia entre fórmulas para poder determinar que dos fórmulas diferentes expresan la misma propiedad de un modelo.

Definición 2.18. *Diremos que dos fórmulas φ y ψ son equivalentes si son verdaderas en los mismos estados, esto es, para todo modelo M y para todo $s \in M$:*

$$M, s \models \varphi \Longleftrightarrow M, s \models \psi.$$

Lo denotaremos como: $\varphi \equiv \psi$.

Diremos que dos fórmulas φ y ψ son equivalentes sobre una clase de modelos $\mathcal{X}$ si para todo modelo $M \in \mathcal{X}$ y para todo $s \in M$:

$$M, s \models \varphi \Longleftrightarrow M, s \models \psi.$$

Lo denotaremos también como $\varphi \equiv \psi$ y entenderemos la clase $\mathcal{X}$ por el contexto.

Notación 2.2. En la siguiente definición, daremos un nuevo uso del símbolo $\equiv$. También usaremos $\equiv$ para decir que dos lenguajes son igualmente expresivos.

Definición 2.19. *Dados dos lenguajes $\mathcal{L}_1$ y $\mathcal{L}_2$, que son interpretados en la misma clase de modelos $\mathcal{X}$, podemos comparar su poder expresivo como sigue:*

- *$\mathcal{L}_2$ es al menos tan expresivo como $\mathcal{L}_1$ si y solo si para cada fórmula $\varphi_1 \in \mathcal{L}_1$ existe una fórmula $\varphi_2 \in \mathcal{L}_2$ tal que $\varphi_1 \equiv \varphi_2$. Se denota como $\mathcal{L}_1 \preceq \mathcal{L}_2$.*
- *$\mathcal{L}_1$ y $\mathcal{L}_2$ son igualmente expresivos si y solo si $\mathcal{L}_1 \preceq \mathcal{L}_2$ y $\mathcal{L}_2 \preceq \mathcal{L}_1$. Se denota como $\mathcal{L}_1 \equiv \mathcal{L}_2$.*
- *$\mathcal{L}_2$ es más expresivo que $\mathcal{L}_1$ si y solo si $\mathcal{L}_1 \preceq \mathcal{L}_2$ pero $\mathcal{L}_2 \not\preceq \mathcal{L}_1$. Se denota como $\mathcal{L}_1 \prec \mathcal{L}_2$.*

Para poner en práctica estos conceptos, en el siguiente ejemplo vamos a definir dos lenguajes para la lógica proposicional y los compararemos.

Ejemplo 2.7. Consideramos un conjunto numerable de variables proposicionales P. Definimos el lenguaje de la lógica proposicional, $\mathcal{L}_{\mathrm{PL}}$, dado por la siguiente BNF:

$$\varphi \quad ::= \quad p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid (\varphi \vee \varphi) \mid (\varphi \rightarrow \varphi) \mid (\varphi \leftrightarrow \varphi)$$

El segundo lenguaje que vamos a definir es $\mathcal{L}_{\mathrm{NAND}}$ y viene dado por la siguiente BNF:

$$\varphi \quad ::= \quad p \mid (\varphi \barwedge \varphi)$$

donde el operador NAND ($\barwedge$) está definido con la siguiente tabla de verdad:

φ	ψ	$(\varphi \barwedge \psi)$
0	0	1
0	1	1
1	0	1
1	1	0

El operador NAND lo podemos definir con el lenguaje $\mathcal{L}_{\text{PL}}$, de hecho se tiene la equivalencia:

$$\varphi \barwedge \psi \equiv \neg(\varphi \wedge \psi)$$

Lo mismo podemos hacer con los operadores de $\mathcal{L}_{\text{PL}}$. Los podemos escribir utilizando solo el operador NAND gracias a las equivalencias:

$$\begin{aligned}
\neg\varphi &\equiv \varphi \barwedge \varphi \\
\varphi \wedge \psi &\equiv (\varphi \barwedge \psi) \barwedge (\varphi \barwedge \psi) \\
\varphi \vee \psi &\equiv (\varphi \barwedge \varphi) \barwedge (\psi \barwedge \psi) \\
\varphi \to \psi &\equiv \varphi \barwedge (\psi \barwedge \psi) \\
\varphi \leftrightarrow \psi &\equiv (\varphi \barwedge \psi) \barwedge (\varphi \barwedge \varphi) \barwedge (\psi \barwedge \psi)
\end{aligned}$$

Ahora vamos a comparar los lenguajes. Para ello intentaremos probar $\mathcal{L}_{\text{NAND}} \preceq \mathcal{L}_{\text{PL}}$ y $\mathcal{L}_{\text{PL}} \preceq \mathcal{L}_{\text{NAND}}$ por separado.

- $\boldsymbol{\mathcal{L}_{\text{NAND}} \preceq \mathcal{L}_{\text{PL}}}$: Para probar esta propiedad, definimos una función de traslación $t_1 : \mathcal{L}_{\text{NAND}} \to \mathcal{L}_{\text{PL}}$ de forma que, utilizando la equivalencia $\varphi \barwedge \psi \equiv \neg(\varphi \wedge \psi)$, lleve cada fórmula de $\mathcal{L}_{\text{NAND}}$ a una de $\mathcal{L}_{\text{PL}}$:

$$\begin{aligned}
t_1(p) &= p \\
t_1(\varphi \barwedge \psi) &= \neg(t_1(\varphi) \wedge t_1(\psi))
\end{aligned}$$

Ahora bastaría probar, por inducción sobre φ, que $\varphi \equiv t_1(\varphi)$, para cada fórmula de $\mathcal{L}_{\text{NAND}}$. Como $t_1(\varphi) \in \mathcal{L}_{\text{PL}}$, entonces podemos decir que $\mathcal{L}_{\text{NAND}} \preceq \mathcal{L}_{\text{PL}}$.

- $\boldsymbol{\mathcal{L}_{\text{PL}} \preceq \mathcal{L}_{\text{NAND}}}$: Para probar esto, el procedimiento es semejante al anterior, pero ahora definimos la función de traslación $t_2 : \mathcal{L}_{\text{PL}} \to \mathcal{L}_{\text{NAND}}$ tal que:

$$\begin{aligned}
t_2(p) &= p \\
t_2(\neg\varphi) &= t_2(\varphi) \barwedge t_2(\varphi) \\
t_2(\varphi \wedge \psi) &= (t_2(\varphi) \barwedge t_2(\psi)) \barwedge (t_2(\varphi) \barwedge t_2(\psi)) \\
t_2(\varphi \vee \psi) &= (t_2(\varphi) \barwedge t_2(\varphi)) \barwedge (t_2(\psi) \barwedge t_2(\psi)) \\
t_2(\varphi \to \psi) &= t_2(\varphi) \barwedge (t_2(\psi) \barwedge t_2(\psi)) \\
t_2(\varphi \leftrightarrow \psi) &= (t_2(\varphi) \barwedge t_2(\psi)) \barwedge (t_2(\varphi) \barwedge t_2(\varphi)) \barwedge (t_2(\psi) \barwedge t_2(\psi))
\end{aligned}$$

De nuevo por inducción sobre φ y usando las equivalencias anteriores, podemos probar que $\varphi \equiv t_2(\varphi)$ para cada fórmula de $\mathcal{L}_{\text{PL}}$. Como $t_2(\varphi) \in \mathcal{L}_{\text{NAND}}$, entonces queda probado que $\mathcal{L}_{\text{PL}} \preceq \mathcal{L}_{\text{NAND}}$.

Al haber probado que ambas implicaciones son ciertas, tenemos: $\mathcal{L}_{\text{NAND}} \equiv \mathcal{L}_{\text{PL}}$. Es decir, hemos demostrado que ambos lenguajes para la lógica proposicional son igualmente expresivos.

2.4.2 Bisimulación

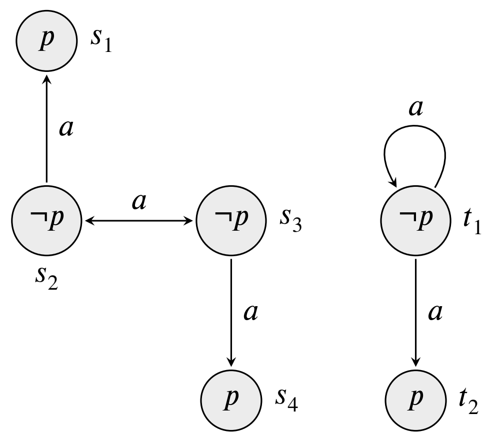
Ahora definimos la noción fundamental de *bisimulación* entre modelos de Kripke, una herramienta que se basa en crear una relación que empareja estados *similares* de dos modelos de Kripke. Esto nos ayudará a poder comparar dos modelos distintos de un mismo lenguaje y decidir si pueden considerarse equivalentes.

Definición 2.20 (Bisimulación). *Fijemos un conjunto de variables atómicas P y un conjunto finito de agentes A. Consideramos dos modelos $M = \langle S, R, V \rangle$ y $M' = \langle S', R', V' \rangle$. Una relación $\mathfrak{R} \subseteq S \times S'$, no vacía, es una bisimulación si y solo si para todo $s \in S$ y para todo $s' \in S'$ con $(s, s') \in \mathfrak{R}$ se cumple:*

- *$s \in V(p)$ si y solo si $s' \in V'(p)$ para todo $p \in P$*
- *$\forall a \in A, \forall t \in S$, si $(s, t) \in R_a$, entonces $\exists t' \in S'$ tal que $(s', t') \in R'_a$ y $(t, t') \in \mathfrak{R}$*
- *$\forall a \in A, \forall t' \in S$, si $(s', t') \in R'_a$, entonces $\exists t \in S$ tal que $(s, t) \in R_a$ y $(t, t') \in \mathfrak{R}$*

Escribiremos $(M, s) \leftrightarrows (M', s')$ si hay una bisimulación entre M y M' que asocia los estados s y s. Diremos entonces que (M, s) y (M', s') son bisimilares.

Ejemplo 2.8. Para ver de forma más clara la definición de bisimulación, veamos un ejemplo de dos modelos bisimilares, M_1 y M_2. Utilizaremos el ejemplo de [Fig].



Sea la bisimulación:

$$\mathfrak{R} = \{(s_1, t_2), (s_2, t_1), (s_3, t_1), (s_4, t_2)\}$$

Veamos si cumple las tres condiciones de la definición. Para la primera, basta con observar que $s_1, s_4 \in V_1(p)$ y están relacionadas con $t_2 \in V_2(p)$, y $s_2, s_3 \notin V_1(p)$ y están relacionadas con $t_1 \notin V_2(p)$. Para la segunda y la tercera, vamos a comprobarlo estado por estado:

- Los casos de s_1, s_4, t_2 se verifican por no haber ningún par $(s_1, s_i) \in R_a$, $(s_4, s_i) \in R_a$, $(t_2, t_i) \in R'_a$, para ningún $i = 1, 2, 3, 4$.
- **Caso s_2:**
  $(s_2, s_1) \in R_a$ y $\exists t_2 \in S'$ tal que $(t_1, t_2) \in R'_a$ y $(s_1, t_2) \in \mathfrak{R}$
  $(s_2, s_3) \in R_a$ y $\exists t_1 \in S'$ tal que $(t_1, t_1) \in R'_a$ y $(s_3, t_1) \in \mathfrak{R}$
- **Caso s_3:**
  $(s_3, s_2) \in R_a$ y $\exists t_1 \in S'$ tal que $(t_1, t_1) \in R'_a$ y $(s_2, t_1) \in \mathfrak{R}$
  $(s_3, s_4) \in R_a$ y $\exists t_2 \in S'$ tal que $(t_1, t_2) \in R'_a$ y $(s_4, t_2) \in \mathfrak{R}$
- **Caso t_1:**
  $(t_1, t_1) \in R'_a$ y $\exists s_2 \in S$ tal que $(s_3, s_2) \in R_a$ y $(s_2, t_1) \in \mathfrak{R}$
  $(t_1, t_2) \in R'_a$ y $\exists s_1 \in S$ tal que $(s_2, s_1) \in R_a$ y $(s_1, t_2) \in \mathfrak{R}$

Al verificar que se cumplen las tres condiciones, podemos decir que la relación $\mathfrak{R}$ que hemos definido es, en efecto, una bisimulación. Y además, podemos decir, por ejemplo, $(M_1, s_2) \leftrightarroweq (M_2, t_1)$.

Si dos modelos son bisimilares, $(M, s) \leftrightarroweq (M', s')$, entonces se pueden considerar equivalentes sobre el lenguaje de la lógica epistémica. Más formalmente:

Notación 2.3 (Equivalencia elemental). Escribiremos $(M, s) \equiv_{\mathcal{L}_K} (M', s')$ si y solo si $(M, s) \vDash \varphi \Leftrightarrow (M', s') \vDash \varphi$ para todas las fórmulas $\varphi \in \mathcal{L}_K$.

La notación anterior se extenderá de manera natural a otros lenguajes epistémicos y escribiremos $\equiv_{\mathcal{L}}$ cuando corresponda. Si dos modelos son bisimilares, entonces son indistinguibles mediante fórmulas de la lógica epistémica, esto es:

Teorema 2.7. *Fijemos un conjunto de variables atómicas P y un conjunto finito de agentes A. Dados dos modelos, (M, s) y (M', s'), si tenemos que $(M, s) \leftrightarroweq (M', s')$, entonces $(M, s) \equiv_{\mathcal{L}_K} (M', s')$.*

Demostración. Vamos a probar el resultado por inducción sobre φ.
Caso base. Suponemos que $(M, s) \leftrightarroweq (M', s')$. Por la primera condición de la Definición 2.20, tenemos que $(M, s) \vDash p$ si y solo si $(M', s') \vDash p$, $\forall p \in P$.

Hipótesis de inducción: "Para cualquier par de modelos (M, s) y (M', s'), si $(M, s) \leftrightarrows (M', s')$, entonces $(M, s) \vDash \varphi$ si y solo si $(M', s') \vDash \varphi$".
Paso de inducción.
Caso $\neg\varphi$: Supongamos que $(M, s) \vDash \neg\varphi$, es decir, $(M, s) \nvDash \varphi$. Por la hipótesis de inducción tenemos $(M', s') \nvDash \varphi$, que es lo mismo que decir $(M', s') \vDash \neg\varphi$. La implicación opuesta es análoga.
Caso $\varphi_1 \wedge \varphi_2$: Suponemos que $(M, s) \vDash \varphi_1 \wedge \varphi_2$. Por definición esto implica $(M, s) \vDash \varphi_1$ y $(M, s) \vDash \varphi_2$. Por la hipótesis de inducción se tiene $(M', s') \vDash \varphi_1$ y $(M', s') \vDash \varphi_2$. Ahora de nuevo por definición, pero a la inversa, obtenemos $(M', s') \vDash \varphi_1 \wedge \varphi_2$.
Caso $\varphi_1 \vee \varphi_2$: Análogo al anterior.
Caso $\varphi_1 \to \varphi_2$: Análogo al anterior.
Caso $K_a\varphi$: Supongamos que $(M, s) \vDash K_a\varphi$. Tomamos un t' cualquiera tal que $(s', t') \in R'_a$. Por la tercera condición de la Definición 2.20, $\exists t \in S$ tal que $(s, t) \in R_a$ y $(t, t') \in \mathfrak{R}$. Por un lado, utilizando la hipótesis de inducción tenemos que $(M, t) \vDash \varphi$ si y solo si $(M', t') \vDash \varphi$. Y por otro, por definición de $(M, s) \vDash K_a\varphi$ obtenemos que $(M, t) \vDash \varphi$. Por tanto, $(M', s') \vDash \varphi$. Dado que t' era arbitrario, $(M', s') \vDash \varphi, \forall t'$ tal que $(s', t') \in R'_a$. Concluimos que por definición $(M', s') \vDash K_a\varphi$. Solo hemos probado una implicación, para probar la contraria el razonamiento es análogo pero usando la segunda condición de la Definición 2.20. ∎

Este teorema implica que la condición de que dos modelos sean bisimilares es suficiente para que sean elementalmente equivalentes. Pero NO es condición necesaria y lo vamos a comprobar a continuación, dando ejemplos de modelos que son elementalmente equivalentes pero no se puede crear una bisimulación entre ellos.

En nuestro primer ejemplo, vamos a definir dos modelos y la idea es ver que la diferencia entre ellos solo se hace evidente al considerar el conjunto completo de átomos simultáneamente. Como una fórmula epistémica solo contiene un número finito de variables atómicas, esta diferencia no aparece y, de hecho, los dos modelos resultarán elementalmente equivalentes. Sin embargo, será fácil comprobar que no son bisimilares.

Definición 2.21. *Sea un conjunto numerable de átomos P. Suponemos que existe una enumeración de átomos tal que p_n es el n-ésimo átomo en esta enumeración. Definimos, $M^1 = \langle S^1, R^1, V^1 \rangle$, donde:*

- $S^1 = \{s^1\} \cup \mathbb{N}$
- $R^1 = \{s^1\} \times \mathbb{N}$
- $V^1(p_n) = \{n\}$

En este modelo, cada átomo es verdadero en un único estado $n \in \mathbb{N}$. Observése que en el estado destacado s^1 no se verifica ningún átomo. Además la relación del modelo no es reflexiva, ni sobreyectiva, ni euclídea (estamos considerando que hay un único agente en el lenguaje y hablamos, por abuso de notación, de una única relación de accesibilidad entre los estados del dominio del modelo). Ahora definimos el segundo modelo:

Definición 2.22. *Sea un conjunto numerable de átomos P. Suponemos que existe una enumeración de átomos tal que p_n es el n-ésimo átomo en esta enumeración. Definimos, $M^2 = \langle S^2, R^2, V^2 \rangle$, donde:*

- $S^2 = \{s^2, \omega\} \cup \mathbb{N}$
- $R^2 = \{s^2\} \times (\mathbb{N} \cup \{\omega\})$
- $V^2(p_n) = \{n\}$

El segundo modelo es similar al anterior, pero en el estado extra de este modelo, ω, ningún átomo es verdadero.

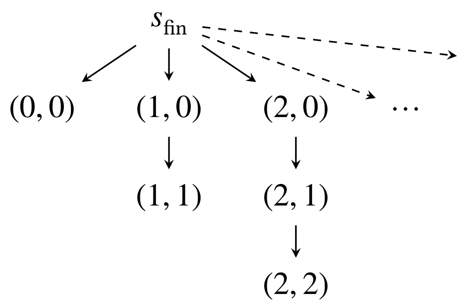
Es evidente que (M^1, s^1) y (M^2, s^2) no son bisimilares, ya que no se satisface la tercera condición de la Definición 2.20: para ω, se tiene $(s^2, \omega) \in R^2$, pero no existe ninguna correlación entre los estados de los dos modelos de tal forma que $\exists t \in S^1$ tal que $(s^1, t) \in R^1$ y que (t, ω) pertenezca a dicha correlación.
Veamos que, sin embargo, (M^1, s^1) y (M^2, s^2) satisfacen las mismas fórmulas. Podemos hacerlo demostrándolo por inducción sobre fórmulas. La prueba es sencilla, el único caso no trivial es el del operador epistémico. Suponemos una fórmula de la forma $K\varphi$ distinguible entre los dos modelos. La única forma de que esto ocurra es que $K\varphi$ sea cierto en (M^1, s^1) y falso en (M^2, s^2). Es decir, que φ sea válida en cada estado $n \in \mathbb{N}$, y falso en ω. Pero φ es finito, así que solo un número finito de átomos, p_i, están en φ. Digamos que p_{n_0} es el átomo con el cardinal más alto que está en φ. Entonces, en el estado $(n_0 + 1)$ no se satisface φ, lo que contradice la hipótesis inicial. Por tanto, todas las fórmulas de la forma $K\varphi$ son indistinguibles entre estos dos modelos. Como conclusión, $(M^1, s^1) \equiv_{\mathcal{L}_K} (M^2, s^2)$.

Incluso si limitamos el número de átomos a un número finito, puede haber un número infinito de fórmulas no equivalentes diferentes, lo cual hace que podamos dar otro tipo de contraejemplo. De hecho, daremos un segundo ejemplo utilizando dos modelos de Kripke en los que en todos los estados se satisfacen los mismos átomos proposicionales: ninguno.

Definición 2.23 (Erizo con espinas finitas). *Definimos $H_{fin} = \langle S, R, V \rangle$ de forma que:*

- $S = \{(n, m) \mid n \in \mathbb{N} \text{ y } m \leq n\} \cup \{s_{fin}\}$
- $R = \{((n, m), (n, m+1)) \mid m < n\} \cup \{(s_{fin}, (n, 0)) \mid n \in \mathbb{N}\}$
- $V(p) = \emptyset$

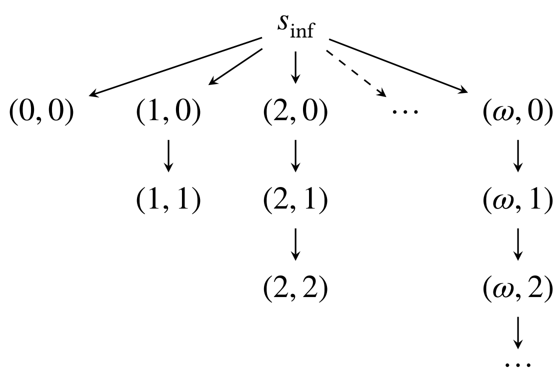
Este modelo consiste en una espina de longitud n para cada $n \in \mathbb{N}$. Todas las espinas son accesibles desde el estado s_{fin}.



Definición 2.24 (Erizo con una espina infinita). *Definimos $H_{inf} = \langle S, R, V \rangle$ donde:*

- $S = \{(n, m) \mid n \in \mathbb{N} \text{ y } m \leq n\} \cup \{(\omega, n) \mid n \in \mathbb{N}\} \cup \{s_{inf}\}$
- $R = \{((n, m), (n, m+1)) \mid m < n \text{ o } n = \omega\} \cup \{(s_{inf}, (n, 0)) \mid n \in \mathbb{N} \cup \{\omega\}\}$
- $V(p) = \emptyset$

Este modelo es bastante parecido al anterior pero tiene una espina infinita formada por los estados (ω, n). Podemos observar que $(H_{\text{fin}}, (n, m)) \leftrightarrows (H_{\text{inf}}, (n, m)), \forall n, m \in \mathbb{N}$, al igual que $(H_{\text{fin/inf}}, (n, m)) \leftrightarrows (H_{\text{inf/inf}}, (n+1, m+1)), \forall n, m \in \mathbb{N}$. Aunque los estados con los que vamos a trabajar son s_{fin} y s_{inf}.



Vamos a dar una fórmula que distinguirá, en los erizos, cada estado relacionado con s_{inf} de los otros. La fórmula para cada estado $(n, 0)$ es $\hat{K}^n\top \wedge K^{n+1}\bot$. Esta fórmula es verdadera en el estado $(n, 0)$ y es falsa en $(m, 0)$, para cada $m \neq n$. La clave es que, si desarrollamos cada conjunción por separado, $\hat{K}^n\top$ se reduce a comprobar si en el estado (n, n) es cierta $p \vee \neg p$, y la otra parte de la conjunción se reduce a que no existe el estado $(n, n+1)$. Además, todas estas fórmulas son falsas en el estado $(\omega, 0)$. Utilizando esto, vamos a probar el siguiente teorema.

Teorema 2.8. *No existe una bisimulación* $(H_{fin}, s_{fin}) \leftrightarroweq (H_{inf}, s_{inf})$.

Demostración. Para cada $n \in \mathbb{N}$ la fórmula $\hat{K}^n\top \wedge K^{n+1}\bot$ es cierta en el estado $(n, 0)$, pero es falsa en el estado $(\omega, 0)$. Por tanto, utilizando el Teorema 2.7, no podemos crear una relación para la bisimulación entre $(\omega, 0)$ y otro estado de H_{fin}. Es decir, no se verifica la tercera condición de la Definición 2.20, y por tanto no podemos crear la bisimulación deseada. ∎

Para concluir el contraejemplo, nos falta probar que los modelos anteriores son elementalmente equivalentes. Para ello debemos probar que no existe ninguna fórmula tal que $(H_{\text{fin}}, s_{\text{fin}}) \vDash \varphi$ y $(H_{\text{inf}}, s_{\text{inf}}) \nvDash \varphi$. Para dicha prueba necesitamos introducir un nuevo concepto que será importante en el futuro.

Definición 2.25 (Profundidad modal). *La profundidad modal de una fórmula se define como la función* $d : \mathcal{L}_K \to \mathbb{N}$ *dada por:*

$$
\begin{aligned}
d(p) &= 0 \\
d(\neg\varphi) &= d(\varphi) \\
d(\varphi \wedge \psi) &= max\,(d(\varphi), d(\psi)) \\
d(\varphi \vee \psi) &= max\,(d(\varphi), d(\psi)) \\
d(\varphi \to \psi) &= max\,(d(\varphi), d(\psi)) \\
d(K_a\varphi) &= 1 + d(\varphi)
\end{aligned}
$$

Utilizando la profundidad modal, obtenemos el siguiente resultado que será clave para demostrar la equivalencia entre los modelos que estamos estudiando. Antes debemos tener en cuenta que (n, m) en H_{fin} es bisimilar a (n, m) en H_{inf}, $\forall n, m \in \mathbb{N}$. Además, para simplificar y realizar la prueba de la manera más clara posible, diremos $(n, m) \vDash \varphi$ para indicar que $(H_{\text{fin}}, (n, m)) \vDash \varphi$ y $(H_{\text{inf}}, (n, m)) \vDash \varphi$. Y $(\omega, n) \vDash \varphi$ para expresar $(H_{\text{inf}}, (\omega, n)) \vDash \varphi$.

Lema 2.5. *Para toda* φ*, tal que* $d(\varphi) = n$ *se tiene que* $(n, 0) \vDash \varphi$ *si y solo si* $(\omega, 0) \vDash \varphi$ *y* $(m, 0) \vDash \varphi$, $\forall m > n$*, siendo* $m, n \in \mathbb{N}$.

Demostración. La demostración la haremos por inducción sobre φ. El caso base y los casos de conjunción y negación son simples de probar. El único caso no trivial es el de las fórmulas de la forma $K\varphi$. Sea $d(\varphi) = n$, la hipótesis de inducción sostiene $(n, 0) \vDash \varphi$ si y solo si $(\omega, 0) \vDash \varphi$ y $(m, 0) \vDash \varphi$ para toda $m > n$. Por otro lado, para todo $k \in ((\mathbb{N} \setminus \{0\} \cup \{\omega\}))$ se tiene que $(k, 0) \vDash K\varphi$ si y solo si $(k, 1) \vDash \varphi$, ya que $(m, 0)$ es bisimilar a $(m + 1, 1)$, $\forall m \in \mathbb{N}$, y $(\omega, 0)$ es bisimilar a $(\omega, 1)$. Con todo esto, tenemos que es cierto que $(n + 1, 1) \vDash \varphi$ si y solo si $(\omega, 1) \vDash \varphi$ y $(m + 1, 1) \vDash \varphi$ para toda $m > n$. Por lo tanto, $(n, 0) \vDash K\varphi$ si y solo si $(\omega, 0) \vDash K\varphi$ y $(m, 0) \vDash K\varphi$, para toda $m > n$. ∎

Ahora sí, pasamos a probar la equivalencia.

Teorema 2.9. $(H_{fin}, s_{fin}) \equiv_{\mathcal{L}_K} (H_{inf}, s_{inf})$

Demostración. Por inducción sobre φ. El caso base, la negación y la conjunción son sencillos de probar. El único caso no trivial es el caso para las fórmulas del tipo $K\varphi$. Es fácil ver que $(H_{\text{inf}}, s_{\text{inf}}) \vDash K\varphi$ implica $(H_{\text{fin}}, s_{\text{fin}}) \vDash K\varphi$. La implicación contraria la demostramos por reducción al absurdo, suponemos que $(H_{\text{fin}}, s_{\text{fin}}) \vDash K\varphi$ y $(H_{\text{inf}}, s_{\text{inf}}) \nvDash K\varphi$. Esto solo ocurre si $(\omega, 0) \nvDash \varphi$, pero entonces, por el Lema 2.5, $(n, 0) \nvDash \varphi$ para n tal que $d(\varphi) = n$. Y esto contradice a el supuesto $(H_{\text{fin}}, s_{\text{fin}}) \vDash K\varphi$. ∎

Con estos dos contraejemplos hemos visto que la noción de bisimilaridad entre modelos de Kripke no captura, en general, el hecho de que dos modelos satisfagan las mismas fórmulas epistémicas. Es una condición suficiente pero no necesaria. Sin embargo, obsérvese que en ambos contraejemplos los correspondientes modelos contenían una cantidad infinita de estados. Esto no es un accidente, como demuestra el siguiente resultado.

Definición 2.26. *Sea un modelo $M = \langle S, R, V\rangle$. Diremos que M es finito si su dominio está formado por un conjunto finito de estados, es decir, si $card(S) \in \mathbb{N}$.*

Teorema 2.10. *Sean dos modelos finitos $M = \langle S, R, V\rangle$ y $M' = \langle S', R', V'\rangle$, y dos estados dados $s \in S$ y $s' \in S'$. Si $(M, s) \equiv_{\mathcal{L}_K} (M', s')$, entonces $(M, s) \leftrightarrows (M', s')$.*

Demostración. Partimos de $(M, s) \equiv_{\mathcal{L}_K} (M', s')$. Ahora vamos a crear una relación entre los dos modelos y probaremos que es una bisimulación. Definimos:

$$\mathfrak{R} = \{(t, t') \mid (M, t) \equiv_{\mathcal{L}_K} (M', t')\}.$$

Veamos que cumplen las tres condiciones de la Definición 2.20.

- La primera es trivial, por la definición de equivalencia.

- Para la segunda, partimos de $(t, t') \in \mathfrak{R}$ y $(t, u) \in R_a$. Ahora suponemos que no existe u' un estado tal que $(t', u') \in R'_a$ y $(u, u') \in \mathfrak{R}$. Esto implica que para cada u' tal que $(t', u') \in R'_a$, existe una fórmula $\varphi_{u'}$ tal que $(M', u') \vDash \varphi_{u'}$ y $(M, u) \nvDash \varphi_{u'}$. Definimos:

  $$\varphi = \bigvee_{(t', u') \in R'_a} \varphi_{u'}$$

  Puesto que hay un número finito de estados, esta fórmula es finita y está bien construida. Pero entonces tenemos que $(M, t) \vDash \hat{K}_a \neg\varphi$, pero $(M', t') \vDash K_a\varphi$, lo cual contradice nuestra hipótesis inicial.
- El razonamiento para verificar que se cumple la tercera condición es análogo al aplicado para probar la segunda. ▮

Observación 2.4. La clave para la demostración anterior es que podemos suponer que la fórmula φ construida es una fórmula finita, algo que no podemos aplicar en el caso general.

Hemos visto que, en el caso general, dos modelos elementalmente equivalentes no son necesariamente bisimilares. Queda pendiente pues encontrar alguna noción entre modelos de Kripke que capture la noción de "tener la misma teoría".

En lo que sigue introduciremos *juegos* como herramientas para comparar modelos. Como veremos, este concepto da una condición necesaria y suficiente para que dos modelos sean equivalentes y será de gran utilidad para el estudio del poder expresivo de los lenguajes epistémicos.

2.4.3 Juegos

La idea de usar juegos para comparar modelos es debido a Ehrenfeucht y Fraïssé. La adaptación a la lógica modal es tan simple que nadie se atribuye su origen. La definición más temprana que se conoce se encuentra en el trabajo de Kees Doets ([Doe87]).

La idea es que dos jugadores, *spoiler* y *duplicador*, juegan con dos modelos, M y M'. Spoiler intentará probar que los modelos son distinguibles y, por su parte, duplicador hará lo posible para demostrar que los modelos son equivalentes. Sin embargo, el juego tiene un número finito de jugadas. Si spoiler no consigue demostrar que los dos modelos son distinguibles, duplicador gana el juego.

Si duplicador tiene una estrategia ganadora para cada posible número de rondas, entonces los dos modelos tienen la misma teoría, es decir, son elementalmente equivalentes. Pero también ocurre lo contrario, si dos modelos son equivalentes, duplicador tendrá una estrategia ganadora para el juego de comparación de modelos, sea cual sea el número de rondas.

La diferencia entre la bisimulación y los juegos es que cuando hay un número infinito de átomos o de estados la bisimulación pierde su eficacia. En cambio, los juegos están configurados para centrarse en una parte finita del modelo y poder tener en cuenta un número finito de átomoso estados.

Veamos ahora las definiciones formales.

Definición 2.27 (El juego $\mathcal{L}_K$). *Consideramos dos modelos $M = \langle S, R, V\rangle$ y $M' = \langle S', R', V'\rangle$ y dos estados $s \in S$ y $s' \in S'$. La n-ésima ronda del juego $\mathcal{L}_K$ entre spoiler y duplicador sobre (M, s) y (M', s') se juega de la siguiente forma:*
Para $n = 0$, spoiler gana si s y s' se diferencian en el valor de sus variables atómicas, en caso contrario, duplicador gana.
Para $n > 0$, spoiler puede elegir entre uno de los siguientes escenarios:

• *Spoiler elige un agente a y un estado t tal que $(s, t) \in R_a$. Duplicador debe responder eligiendo un estado t' tal que $(s', t') \in R'_a$. La jugada se puede resumir como (t, t').*

• *O bien, spoiler elige un agente a y un estado t' tal que $(s', t') \in R'_a$. Duplicador debe responder eligiendo un estado t tal que $(s, t) \in R_a$. La jugada se puede resumir como (t, t').*

Si uno de los dos jugadores no puede realizar la acción asignada arriba (no puede elegir un estado sucesor), ese jugador pierde. Si los dos estados elegidos por los jugadores se diferencian en los valores de las propiedades atómicas, spoiler gana el juego. Si spoiler no ha ganado después de todas las n rondas, duplicador gana el juego.

Debemos identificar el número de rondas n en el juego anterior con la profundidad modal de las fórmulas que se van a estudiar. De esta forma, ocurrirá que spoiler tendrá una estrategia ganadora para la n-ésima ronda del juego si y solo si existe una fórmula φ con $d(\varphi) \leq n$ tal que φ es verdadera en un modelo y falsa en el otro.

Si hacemos restricciones en el lenguaje $\mathcal{L}_K$ teniendo en cuenta la profundidad modal de las fórmulas, distinguiremos cada nivel de profundidad del lenguaje $\mathcal{L}_K$ como un lenguaje propio.

Definición 2.28. *Definimos el lenguaje $\mathcal{L}^0_K$ como el formado por las fórmulas dadas* *por la notación BNF:*

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid (\varphi \vee \varphi) \mid (\varphi \to \varphi)$$

A partir de esto, definimos el lenguaje $\mathcal{L}_K^{n+1}$ de forma inductiva. Estará formado por las fórmulas dadas por la siguiente notación BNF:

$$\varphi ::= \psi \mid K_a\psi \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid (\varphi \vee \varphi) \mid (\varphi \to \varphi),$$

donde $\psi \in \mathcal{L}_K^n$.

Es fácil comprobar que cada nivel del lenguaje se puede representar como $\mathcal{L}_K^n = \{\varphi \in \mathcal{L}_K \mid d(\varphi) \leq n\}$ y, por tanto, se tiene la igualdad $\mathcal{L}_K = \bigcup_{n\in\mathbb{N}} \mathcal{L}_K^n$.

Si consideramos un número finito de variables proposicionales y definimos $[\![\varphi]\!]$ como la clase de todos los modelos donde φ es verdadera, entonces tenemos que el número de clases para un nivel fijo del lenguaje es finito, es decir, $\{[\![\varphi]\!] \mid \varphi \in \mathcal{L}_K^n\}$ es un conjunto finito. Esto lo enunciamos en el siguiente lema.

Lema 2.1. Consideramos un conjunto finito de átomos P. Para cada $n \in \mathbb{N}$, existe un número finito de fórmulas diferentes salvo equivalencia lógica en $\mathcal{L}_K^n(P)$.

Demostración. Vamos a probar este lema por inducción sobre n.

- **Caso base.** Para $n = 0$, el conjunto $\{[\![\varphi]\!] \mid d(\varphi) \leq 0\}$ contiene todas las fórmulas booleanas. Como P es finito, este conjunto es finito salvo equivalencia.
- **Hipótesis de inducción:** "El conjunto $\{[\![\varphi]\!] \mid d(\varphi) \leq n\}$ es finito".
- **Paso de inducción.** El conjunto de fórmulas de profundidad $n + 1$ está formado por todas las fórmulas de la forma $K_a\varphi$, siendo φ de profundidad n, y combinaciones booleanas de fórmulas de este tipo. Ya que el número de agentes es finito y $\{[\![\varphi]\!] \mid d(\varphi) \leq n\}$ también lo es, este conjunto es finito salvo por equivalencia.

Utilizaremos este lema en la prueba del teorema que viene a continuación, en el cual daremos una condición necesaria y suficiente para que duplicador tenga una estrategia ganadora. Antes de enunciar el teorema hay que tener en cuenta que notaremos la equivalencia entre modelos en el lenguaje $\mathcal{L}_K^n$ como $(M, s) \equiv_{\mathcal{L}_K^n} (M', s')$.

Teorema 2.11. *Para todo $n \in \mathbb{N}$, para todo conjunto finito de átomos P y para dos modelos cualesquiera (M, s) y (M', s'), se cumple que: duplicador tiene una estrategia ganadora para la n-ésima ronda del juego $\mathcal{L}_K(P)$ sobre (M, s) y (M', s') si y solo si $(M, s) \equiv_{\mathcal{L}_K^n} (M', s')$.*

Demostración. Vamos a demostrarlo por inducción sobre n.

- **Caso base.** En el lenguaje $\mathcal{L}_K^0$ se prueba el teorema fácilmente por la definición del juego.
- **Hipótesis de inducción:** Para todos los modelos (M, s) y (M', s') y cualquier conjunto finito de átomos P: duplicador tiene una estrategia ganadora para la n-ésima ronda del juego $\mathcal{L}_K(P)$ sobre los dos modelos si y solo si $(M, s) \equiv_{\mathcal{L}_K^n} (M', s')$.
- **Paso de inducción.** Empezamos por probar la implicación de izquierda a derecha. Suponemos pues que duplicador tiene una estrategia ganadora para la $(n+1)$-ésima ronda del juego sobre (M, s) y (M', s'). Para demostrar que $(M, s) \equiv_{\mathcal{L}_K^{n+1}} (M', s')$, volvemos a aplicar inducción, pero ahora sobre $\varphi \in \mathcal{L}_K^{n+1}$.
    - **Casos base.** Supongamos primero que φ es de la forma ψ, donde $\psi \in \mathcal{L}_K^n$. Utilizando la hipótesis de inducción anterior, se tiene que $(M, s) \vDash \psi$ si y solo si $(M', s') \vDash \psi$.
      El segundo caso base corresponde al caso en el que φ es de la forma $K_a\psi$, con $\psi \in \mathcal{L}_K^n$. Suponemos, sin perdida de generalidad, que $(M, s) \vDash K_a\psi$. Tomamos un t' cualquiera tal que $(s', t') \in R'_a$. Como duplicador tiene una estrategia ganadora, tiene una respuesta para cada elección de spoiler. Si spoiler elige t', entonces existe t tal que $(s, t) \in R_a$ y duplicador tiene una estrategia ganadora para la n-ésima ronda del subjuego sobre (M, t) y (M', t'). Por la hipótesis de inducción se tiene que $(M, t) \vDash \psi$ si y solo si $(M', t') \vDash \psi$. Por la suposición inicial $(M, s) \vDash K_a\psi$ y $(s, t) \in R_a$, entonces $(M, t) \vDash \psi$. Por tanto, $(M', t') \vDash \psi$. Como hemos tomado un t' cualquiera, $(M', t') \vDash \psi$ se tiene para todo t' tal que $(s', t') \in R'_a$ y de aquí concluimos que $(M', s') \vDash K_a\psi$.
    - **Hipótesis de inducción.** Dadas dos fórmulas $\varphi, \psi \in \mathcal{L}_K^{n+1}$, tenemos que $(M, s) \vDash \varphi$ si y solo si $(M', s') \vDash \varphi$, y $(M, s) \vDash \psi$ si y solo si $(M', s') \vDash \psi$.
    - **Paso de inducción.** En el caso de la negación,, supongamos que $(M, s) \vDash \neg\varphi$, es decir, $(M, s) \nvDash \varphi$. Por la segunda hipótesis de inducción, tenemos $(M', s') \nvDash \varphi$, que es lo mismo que decir $(M', s') \vDash \neg\varphi$.
      En el caso de la conjunción, supongamos que $(M, s) \vDash \varphi \wedge \psi$. Por definición esto implica $(M, s) \vDash \varphi$ y $(M, s) \vDash \psi$. Por la segunda hipótesis de inducción, se tiene que $(M', s') \vDash \varphi$ y $(M', s') \vDash \psi$. Ahora, de nuevo por definición, pero a la inversa, obtenemos $(M', s') \vDash \varphi \wedge \psi$.
      Los casos de la disyunción y de la implicación son análogos.

  Por lo tanto, $(M, s) \equiv_{\mathcal{L}_K^{n+1}} (M', s')$.

Para demostrar la implicación contraria. Suponemos $(M, s) \equiv_{\mathcal{L}_K^{n+1}} (M', s')$. Ahora vamos a describir la estrategia ganadora de duplicador. Para ello vamos a suponer, sin perdida de generalidad, que spoiler en su primer movimiento elige un estado t tal que $(s,t) \in R_a$. Tenemos que probar que existe t' tal que $(s',t') \in R'_a$ y que $(M,t) \equiv_{\mathcal{L}_K^n} (M',t')$ (y aplicamos entonces la hipótesis de inducción). Vamos a hacerlo por reducción al absurdo. Suponemos que no existe dicho t'. Entonces para cada t' tal que $(s',t') \in R'_a$ spoiler tiene una estrategia ganadora para el subjuego restante. Por la primera hipótesis de inducción, existe una fórmula $\varphi_{t'}$ de profundidad como mucho n para cada t' con $(s',t') \in R'_a$ tal que $(M',t') \vDash \varphi_{t'}$ y $(M,t) \nvDash \varphi_{t'}$. Por el lema anterior, el conjunto $\{[\![\varphi_{t'}]\!] \mid (s',t') \in R'_a\}$ contiene un número finito de fórmulas que no son equivalentes. Necesitamos una función f que escoja una única fórmula de cada clase de equivalencia $[\![\varphi_{t'}]\!]$. Definimos la siguiente fórmula:

$$\varphi = \bigvee_{(s',t') \in R'_a} f([\![\varphi_{t'}]\!])$$

Esta fórmula es finita y su profundidad es a lo sumo n. Tenemos $(M',s') \vDash K_a\varphi$ y $(M,s) \vDash K_a\varphi$. Pero, $d(K_a\varphi) \leq n+1$. Esto está en contradicción con la suposición inicial. Por tanto, duplicador tiene una estrategia ganadora para la ronda $(n+1)$-ésima del juego. ∎

Hemos obtenido que para $\mathcal{L}_K^n$ el juego de comparación de modelos nos da una condición necesaria y suficiente para su equivalencia elemental. Veamos que también podemos aplicar esta condición necesaria y suficiente a todo el lenguaje $\mathcal{L}_K$.

Teorema 2.12. *Dados dos modelos (M,s), (M',s'), se tiene que $(M,s) \equiv_{\mathcal{L}_K} (M',s')$ si y solo si para todo $n \in \mathbb{N}$ duplicador tiene una estrategia ganadora para la n-ésima ronda del juego $\mathcal{L}_K$ sobre (M,s) y (M',s').*

Demostración. Vamos a hacer la prueba de las dos implicaciones por reducción al absurdo.

De izquierda a derecha. Partimos de $(M,s) \equiv_{\mathcal{L}_K} (M',s')$. Suponemos que existe n tal que duplicador no tiene una estrategia ganadora para la n-ésima ronda del juego sobre (M,s) y (M',s'). Esto implica, por el teorema anterior, que existe una fórmula φ, de profundidad a lo sumo n, tal que $(M,s) \vDash \varphi$ y $(M',s') \nvDash \varphi$. Pero esto contradice la equivalencia entre ambos modelos, por tanto, duplicador tiene una estrategia ganadora.

De derecha a izquierda. Partimos de que $\forall n \in \mathbb{N}$ duplicador tiene una estrategia ganadora para la n-ésima ronda del juego sobre (M,s) y (M',s'). Suponemos que

$(M, s) \not\equiv_{\mathcal{L}_K} (M', s')$, es decir, existe φ tal que $(M, s) \vDash \varphi$ y $(M', s') \nvDash \varphi$. Pero por la hipótesis de la que partimos duplicador tiene una estrategia ganadora para la $d(\varphi)$-ésima ronda del juego, por el teorema anterior, esto implica que $(M, s) \vDash \varphi$ si y solo si $(M', s') \vDash \varphi$. Esto está en contradicción con la suposición anterior. ❙

2.4.4 El poder expresivo de $\mathcal{S}5$ con un único agente

En lo que sigue, aplicaremos estos conceptos de comparación de modelos al lenguaje de la lógica epistémica para estudiar su poder expresivo. En primer lugar, vamos a estudiar el poder expresivo del lenguaje de la lógica epistémica con un *único agente*, que denotaremos por $\mathcal{L}_K(\{a\})$. A lo largo de esta subsección, consideramos siempre que estamos trabajando en este lenguaje reducido y que los modelos considerados son modelos epistémicos.

Considerando el juego de comparación de modelos para modelos de $\mathcal{S}5$ para un único agente, obtenemos el siguiente resultado.

Teorema 2.13. *Duplicador tiene una estrategia ganadora para la n-ésima ronda del juego $\mathcal{L}_K$ sobre (M, s) y (M', s') para cada $n \in \mathbb{N}$ si y solo si duplicador tiene una estrategia ganadora para la primera ronda del juego $\mathcal{L}_K$ sobre (M, s) y (M', s').*

Demostración. La implicación de izquierda a derecha es trivial. Para demostrar la implicación de derecha a izquierda, partimos de que duplicador tiene una estrategia ganadora para la primera ronda del juego. Esto implica que s y s' verifican las mismas propiedades atómicas. Además, para cada estado relacionado con s o s' mediante a, duplicador puede responder con el estado equivalente del otro modelo. Pero como el conjunto de estados relacionados con s, o con s', por a es el mismo para todos los estados que están relacionados con s, o con s' por a, la misma estrategia puede ser repetida constantemente, lo que provoca que duplicador tenga una estrategia ganadora para la n-ésima ronda del juego. ❙

De este resultado y utilizando el Teorema 2.11 en ambas direcciones, se obtiene el siguiente corolario.

Corolario 2.2. $(M, s) \equiv_{\mathcal{L}_K} (M', s')$ si y solo si $(M, s) \equiv_{\mathcal{L}_K^1} (M', s')$.

Partiendo de este corolario, vamos a comparar el lenguaje de la lógica epistémica de un único agente con el mismo lenguaje pero restringido por la condición de que las fórmulas tengan profundidad modal menor o igual a 1. Vamos a obtener que $\mathcal{L}_K(\{a\})$ y $\mathcal{L}_K^1(\{a\})$ son igualmente expresivos (sobre la clase $\mathcal{S}5$).

Teorema 2.14. $\mathcal{L}_K(\{a\}) \equiv \mathcal{L}_K^1(\{a\})$.

Demostración. Puesto que $\mathcal{L}_K^1(\{a\})$ es un sublenguaje de $\mathcal{L}_K(\{a\})$, se tiene automáticamente que $\mathcal{L}_K^1(\{a\}) \preceq \mathcal{L}_K(\{a\})$.

Para probar que $\mathcal{L}_K(\{a\}) \preceq \mathcal{L}_K^1(\{a\})$, tomamos una fórmula $\varphi \in \mathcal{L}_K(\{a\})$ y tenemos que definir su traducción al lenguaje $\mathcal{L}_K^1(\{a\})$. Dado un modelo $M = \langle S, \sim, V\rangle \in \mathcal{S}5$ y un estado $s \in S$, definimos $\delta_s = \bigwedge\{p \mid p \in P(\varphi) \text{ y } (M,s) \vDash p\}$, donde $P(\varphi)$ es el conjunto de átomos que forman φ. Denotamos por $\mathcal{S}5(\varphi)$ la clase de estados epistémicos que satisfacen φ. Consideramos la fórmula:

$$\psi = \bigvee_{(M,s)\in\mathcal{S}5(\varphi)} \left(\delta_s \wedge \bigwedge_{s\sim_a t} \hat{K}\delta_t \wedge K \bigvee_{s\sim_a t} \delta_t\right)$$

Puesto que $P(\varphi)$ es finito, existe un número finito de δ_t diferentes. Por tanto, todos los conjuntos de la definición de ψ son finitos y así esta fórmula pertenece a $\mathcal{L}_K^1(\{a\})$. Faltaría probar que $\varphi \equiv \psi$.

Por un lado, suponemos que $(M,s) \vDash \varphi$ donde (M,s) es un estado epistémico. Esto implica que una de las disyunciones de ψ es cierta, por tanto, $(M,s) \vDash \psi$.

Por otro lado, suponemos $(M,s) \vDash \psi$. Esto implica que una de las disyunciones de ψ es válida en (M,s), es decir, existe un (M',s') tal que $(M,s) \vDash \delta_{s'} \wedge \bigwedge_{s'\sim'_a t'} \hat{K}\delta_{t'} \wedge K \bigvee_{s'\sim'_a t'} \delta_{t'}$. Lo que significa que duplicador tiene una estrategia ganadora para la primera ronda del juego $\mathcal{L}_K^1(\{a\})$ sobre (M,s) y (M',s'). Por el Teorema 2.13, duplicador tiene una estrategia ganadora para la n-ésima ronda del juego $\mathcal{L}_K$ sobre (M,s) y (M',s') para cada $n \geq 0$. Luego, por el Corolario 2.2, se tiene que $(M,s) \equiv_{\mathcal{L}_K} (M',s')$. Por tanto, como se tiene que $(M',s') \vDash \varphi$, se obtiene que $(M,s) \vDash \varphi$.

Este resultado de equivalencia entre fórmulas es muy conocida en la lógica modal. Fue probado por Hughes y Cresswell en *"A New Introduction to Modal Logic"* ([HC96]), cuya base viene de *"Ein erweiterter klassenkalkün"* de Wajsberg (véase [Waj33]). Por otro lado, fue probado de forma independiente por Meyer y van der Hoek en *"Epistemic Logic for AI and Computer Science"* ([MH95]).

2.4.5 El poder expresivo de $\mathcal{S}5$ multiagente

Cuando consideramos más de un agente en el lenguaje, los resultados que hemos visto no son válidos, ya que con varios agentes existe un número infinito de fórmulas no equivalentes entre sí y que no son equivalentes a ninguna fórmula de profundidad modal 1.

Daremos un contraejemplo muy similar al de los erizos utilizado anteriormente. Pero lo tendremos que modificar porque, en el caso de dos modelos de $S5$, los estados son bisimilares si la valoración $V(p)$ de ambos modelos es vacía.

Proposición 2.5. Dados dos modelos $M = \langle S, \sim, V\rangle$ y $M' = \langle S', \sim', R'\rangle$ en $S5$, si $V = V' = \emptyset$, entonces $(M, s) \leftrightarrows (M', s'), \forall s \in S$ y $s' \in S'$.

Demostración. Vamos a probar que la relación universal $S \times S'$ es una bisimulación. Al ser la valoración de ambos modelos vacía, cualquier par de estados que estén en la bisimulación tendrán las mismas propiedades atómicas. Para probar la segunda condición, tomamos dos estados $s \in S$ y $s' \in S'$. Suponemos que $s \sim_a t$. Ya que las relaciones son de $S5$, tenemos que existe t' tal que $s' \sim'_a t'$. Como la relación de bisimulación es la relación universal, tenemos (t, t'). La tercera condición se prueba de forma análoga a la segunda. ∎

Vamos a definir los modelos que utilizaremos a continuación para obtener el contraejemplo deseado. Obsérvese que vamos a definir los modelos con dos agentes a y b y una única proposición atómica p en el lenguaje subyacente.

Definición 2.29 (Espina $S5$ finita). *Para cada $n \in \mathbb{N}$, definimos $Sp(n) = \langle S, R, V\rangle$, donde:*

- $S = \{m \mid m \in \mathbb{N} \text{ y } m \leq (n+1)\}$
- $R(a) = \{(s, s) \mid s \in S\} \cup \{(m, k) \mid min(m, k) \bmod 2 = 0 \text{ y } |m - k| = 1\}$
- $R(b) = \{(s, s) \mid s \in S\} \cup \{(m, k) \mid min(m, k) \bmod 2 = 1 \text{ y } |m - k| = 1\}$
- $V(p) = \{n + 1\}$

Este modelo es como la espina n-ésima del modelo de la Definición 2.23, pero en este caso las relaciones entre estados son de equivalencia. La relación para el agente a relaciona cada estado par con su sucesor, y la relación para el agente b lo hace con los estados impares. Cabe destacar que el único estado donde p es válida es el estado $(n+1)$-ésimo. Gráficamente:

$$\begin{array}{ccccccccccccc} 0 & \xleftrightarrow{a} & 1 & \xleftrightarrow{b} & 2 & \xleftrightarrow{a} & 3 & \xleftrightarrow{b} & \cdots & \longleftrightarrow & n & \longleftrightarrow & n+1 \\ & & & & & & & & & & & & p \end{array}$$

Definición 2.30 (Espina $S5$ infinita). *Definimos $Sp(\omega) = \langle S, R, V\rangle$, donde:*

- $S = \mathbb{N}$
- $R(a) = \{(s,s) \mid s \in S\} \cup \{(m,k) \mid min(m,k) \; mod \, 2 = 0 \; y \, |m-k| = 1\}$
- $R(b) = \{(s,s) \mid s \in S\} \cup \{(m,k) \mid min(m,k) \; mod \, 2 = 1 \; y \, |m-k| = 1\}$
- $V(p) = \emptyset$

En este segundo modelo tenemos la misma espina pero con un número infinito de estados. Es similar a la espina ω del erizo infinito visto en la Definición 2.24, pero con relaciones de equivalencia entre estados. Obsérvese que en este modelo la variable p no es verdadera en ningún estado.

$$0 \xleftrightarrow{a} 1 \xleftrightarrow{b} 2 \xleftrightarrow{a} 3 \xleftrightarrow{b} 4 \xleftrightarrow{a} 5 \xleftrightarrow{b} \cdots$$

En primer lugar vamos a probar que los $S5$-modelos introducidos no son bisimilares.

Teorema 2.15. *No existe $n \in \mathbb{N}$ tal que:*

$$(Sp(n), 0) \leftrightarrows (Sp(\omega), 0).$$

Demostración. Para cada $n \in \mathbb{N}$, podemos definir una fórmula que distingue los dos estados. La definiremos inductivamente de la siguiente manera:

$$\begin{aligned} \varphi_0 &= \hat{K}_a p \\ \varphi_1 &= \hat{K}_a \hat{K}_b p \\ \varphi_{n+2} &= \hat{K}_a \hat{K}_b \varphi_n \end{aligned}$$

Es claro que $(\mathrm{Sp}(n), 0) \vDash \varphi_n$ y $(\mathrm{Sp}(\omega), 0) \nvDash \varphi_n$ para cada $n \in \mathbb{N}$. Por tanto, no puede existir una relación de bisimulación para $(\mathrm{Sp}(n), 0)$ y $(\mathrm{Sp}(\omega), 0)$.

Ahora vamos a utilizar el juego de comparación de modelos para demostrar la equivalencia entre estos dos modelos hasta nivel n. Hay que tener en cuenta que la clave está en el estado $(n+1)$-ésimo, que es donde se encuentra la diferencia entre los dos modelos.

Teorema 2.16. *Para cada $n \in \mathbb{N}$ se tiene que:*

$$(Sp(n), 0) \equiv_{\mathcal{L}^n_K} (Sp(\omega), 0).$$

Demostración. Vamos a demostrar que duplicador tiene una estrategia ganadora para la n-ésima ronda del juego $\mathcal{L}_K$ sobre $(\mathrm{Sp}(n), 0)$ y $(\mathrm{Sp}(\omega), 0)$. La única opción de que spoiler gane el juego es que pueda realizar un movimiento hacia el $(n+1)$-ésimo estado, pero necesita $(n+1)$ rondas debido a la estructura de los modelos. Por tanto, spoiler no tiene una estrategia ganadora para la ronda n-ésima del juego. Para finalizar aplicamos el Teorema 2.11 y obtenemos $(\mathrm{Sp}(n), 0) \equiv_{\mathcal{L}_K^n} (\mathrm{Sp}(\omega), 0)$.

Como consecuencia de los dos últimos teoremas, obtenemos que cada fórmula φ_n definida en la prueba del primer teorema no es equivalente a una fórmula epistémica de profundidad modal $\leq n$. En particular, en el caso multiagente, no es cierto que toda fórmula sea equivalente a una fórmula del lenguaje $\mathcal{L}_K^1$.

Cerramos esta sección con un resultado que se sigue del Teorema 2.16 y será de utilidad en el siguiente capítulo.

Teorema 2.17. *No existe $\varphi \in \mathcal{L}_K$ tal que $(Sp(\omega), 0) \notin [\![\varphi]\!]$ y $\{(Sp(n), 0) \mid n \in \mathbb{N}\} \subseteq [\![\varphi]\!]$.*

3 Lógica epistémica con conocimiento común

En este capítulo vamos a añadir la noción de conocimiento común al lenguaje epistémico $\mathcal{S}5$. A este nuevo lenguaje lo llamaremos $\mathcal{S}5\mathcal{C}$. En primer lugar, desarrollaremos el lenguaje y la semántica que se forman al incluir este concepto. En la segunda sección de este capítulo veremos los axiomas que forman el sistema lógico **S5C**. Y, con estos conceptos, al igual que hicimos con $\mathcal{S}5$ en el capítulo anterior, estudiaremos más en profundidad las propiedades de completitud del sistema **S5C** y de expresividad en este nuevo lenguaje.

3.1 Lenguaje y semántica

En este capítulo será de vital importancia el operador lógico E_B que vimos en la Definición 2.4. Lo primero que hay que tener en cuenta sobre este operador es que no satisface la introspección positiva. Es decir, aunque el operador modal K_a satisface en **S5** el axioma $K_a\varphi \rightarrow K_aK_a\varphi$, si consideramos el operador E_B, entonces $E_B\varphi$ no necesariamente implica $E_BE_B\varphi$. Vimos ejemplos de este hecho en el Ejemplo 2.2 sobre números consecutivos en el capítulo anterior. De hecho, en general cada fórmula $E_B^m\varphi$ puede ser diferente para cada $m \in \mathbb{N}$. Cuando se verifica $E_B^m\varphi$, para todo $m \in \mathbb{N}$, aparece el operador lógico del conocimiento común, que definimos a continuación.

Definición 3.1. *Dado un subconjunto, B, del conjunto total de agentes, A, definimos el operador lógico:*

$$C_B\varphi = \bigwedge_{n=0}^{\infty} E_B^n\varphi$$

La noción de conocimiento común es muy fuerte y, en general, se obtiene solo para fórmulas débiles. Nóteste que para que no podamos aplicarla, basta con que un único agente dentro de B considere como una posibilidad que algún agente de B considere como una posibilidad … que φ no es cierto.

Añadiendo este operador al lenguaje $\mathcal{L}_K$ obtenemos un nuevo lenguaje al que llamaremos $\mathcal{L}_{KC}$.

Definición 3.2 (El lenguaje $\mathcal{L}_{KC}$). *Sean P un conjunto de variables atómicas y A un conjunto finito de agentes. El lenguaje de la lógica epistémica multiagente con conocimiento común basada en P y A consta de los siguientes símbolos:*

1. *Variables: los elementos de P.*
2. *Conectivas proposicionales: $\neg$, $\wedge$, $\vee$, $\rightarrow$*
3. *Operadores modales: K_a, para cada agente $a \in A$.*
4. *Operadores modales: C_B, para cada $B \subseteq A$.*

Definición 3.3 (Fórmulas de $\mathcal{L}_{KC}$). *Sean P un conjunto de variables atómicas y A un conjunto finito de agentes. Las fórmulas de la lógica epistémica multiagente con conocimiento común está generado por la siguiente notación BNF:*

$$\varphi \ ::= \ p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid (\varphi \vee \varphi) \mid (\varphi \rightarrow \varphi) \mid K_a\varphi \mid C_B\varphi$$

donde $a \in A$ y $B \subseteq A$.

Para grupos pequeños de agentes, a veces escribiremos C_a, C_{ab}, …en lugar de $C_{\{a\}}$, $C_{\{a,b\}}$, …

Ejemplo 3.1 (Generales bizantinos). Supongamos que dos generales aliados, a y b están situados en la cima de dos montañas, y su enemigo está en el valle que hay entre ellas. El enemigo puede defenderse de cada uno de los generales por separado, pero si atacan los dos a la vez la victoria está asegurada.

El general a envía un mensajero al general b con el siguiente mensaje: "Te propongo que ataquemos el primer día del siguiente mes a las 8 en punto" (llamaremos m al mensaje a partir de ahora). No está garantizado que el mensajero consiga llevar con éxito el mensaje. Si suponemos que el mensajero completa el trayecto y consigue entregar el mensaje al general b, se tendría que K_am y K_bK_am, pero no sería buena idea atacar. Para poder atacar el general a debe tener la certeza de que el general b atacará junto a él. Entonces, b envía al mensajero de vuelta con el general a. En el caso en el que el mensajero llegue a entregar la confirmación, tendremos $K_aK_bK_am$. Pero, aún así, no sería buena idea comenzar el ataque, porque el general b no sabe si su confirmación ha llegado, es decir, $K_b K_a K_b m$ no es el caso. O dicho de otro modo, el conocimiento común no se ha establecido.

En general, para cada $n \geq 0$, se puede probar por inducción que, en este caso, el conocimiento común nunca se establecerá. Antes vamos a definir la notación $(K_a K_b)^n$ como la abreviación de $2n$ operadores modales K_a y K_b alternados.

• Cuando n es impar, el mensajero ha realizado con éxito $2n + 1$ viajes. Tenemos como cierto que $K_b(K_a K_b)^n m$, pero $(K_a K_b)^{n+1} m$ no.

• Cuando n es par, el mensajero ha realizado con éxito $2n+2$ viajes. Tenemos cierto que $(K_a K_b)^{n+1} m$, pero $K_b(K_a K_b)^{n+1} m$ no.

Por tanto, el conocimiento común nunca será cierto en este caso usando un mensajero. Además, nunca se podrá realizar dicho ataque porque es necesario que m sea de conocimiento común para realizarlo con éxito.

Hablemos ahora de la semántica para el lenguaje $\mathcal{L}_{KC}$. Vamos a utilizar los mismos modelos de Kripke de la Definición 2.5. Dentro de estos modelos, el conjunto de estados no sufre ninguna variación en su definición. Y utilizamos la misma función de validación, V. Pero será necesario definir la relación de accesibilidad para los nuevos operadores modales C_B a partir de las relaciones $R_a (a \in A)$.

Definición 3.4. *Sea S un conjunto de estados y sea R_b $(b \in B)$ un conjunto de relaciones binarias entre los estados de S. Recordemos que cada R_b es un conjunto de la forma $\{(x, y) \mid R_b xy\}$. Definimos:*

- $R_{E_B} = \bigcup_{b \in B} R_b$
- *Se llama clausura reflexiva transitiva de la relación R, y se denota R^*, a la relación más pequeña tal que:*
  1. $R \subseteq R^*$.
  2. *Para todo x, $R^* xx$.*
  3. *Para todo x, y, z, si $(R^* xy \wedge R^* yz)$, entonces $R^* xz$.*

Definiremos la semántica para el operador de conocimiento común C_B a partir de la clausura reflexiva transitiva de la relación R_{E_B}.

Definición 3.5 (Relación de validez). *Consideramos un conjunto P de variables atómicas, un conjunto finito de agentes A y un modelo epistémico $M = \langle S, \sim, V \rangle$. La definición de la relación binaria de validez, $(M, s) \vDash \varphi$, con $\varphi \in \mathcal{L}_{KC}$, es una ampliación de la que vimos en la Definición 2.6.*

$M, s \vDash p$	*sii* $s \in V(p)$
$M, s \vDash (\varphi \wedge \psi)$	*sii* $M, s \vDash \varphi$ *y* $M, s \vDash \psi$
$M, s \vDash (\varphi \vee \psi)$	*sii* $M, s \vDash \varphi$ *o* $M, s \vDash \psi$
$M, s \vDash (\varphi \to \psi)$	*sii no se tiene* $M, s \vDash \varphi$ *o* $M, s \vDash \psi$
$M, s \vDash \neg\varphi$	*sii no se tiene* $M, s \vDash \varphi$
$M, s \vDash K_a\varphi$	*sii para todo* t *tal que* $R_a st$ *se tiene que* $M, t \vDash \varphi$
$M, s \vDash \hat{K}_a\varphi$	*sii existe* t *tal que* $R_a st$ *y* $M, t \vDash \varphi$
$M, s \vDash E_B\varphi$	*sii para todo* t *tal que* $R_{E_B} st$ *se tiene que* $M, t \vDash \varphi$
$M, s \vDash C_B\varphi$	*sii para todo* t *tal que* $R^*_{E_B} st$ *se tiene que* $M, t \vDash \varphi$

Por abuso de notación, para $R^*_{E_B}$, escribiremos R_{C_B}, o incluso, R_B. Además, nótese que si R_a es una relación de equivalencia entonces $R_a = R^*_a$.

Observación 3.1. El operador E_B puede expresarse en el lenguaje $\mathcal{L}_K$ como vimos en el capítulo anterior. Es inmediato comprobar que la definición de la semántica para E_B que aquí se da concide con la obtenida al desarrollar la definición de E_B en el lenguaje $\mathcal{L}_K$.

Definición 3.6. *La clase de modelos epistémicos* $\mathcal{S}5\mathcal{C}$*, es* $\mathcal{S}5$ *añadiéndole la definición de validez para* $C_B\varphi$ *dada en la Definición 3.5.*

A continuación, vamos a ver la semántica definida para $\mathcal{S}5\mathcal{C}$ puesta en práctica. Para ello, retomaremos el ejemplo de los números consecutivos visto en el capítulo anterior y tomaremos conclusiones sobre el conocimiento común que tenemos en este caso.

Ejemplo 3.2 (Números consecutivos con conocimiento común). Recordemos que los números de Ángel y Beatriz eran 3 y 2, respectivamente. Beatriz no tiene un 4, pero no todos saben que Beatriz no tiene un 4, es decir, $M, \langle 3,2\rangle \vDash \neg b_4 \wedge \neg E_{ab}\neg b_4$, porque $\neg K_b \neg b_4$. También sabemos que todo el mundo sabe que Ángel no tiene un 5, pero no todo el mundo sabe que todo el mundo sabe que Ángel no tiene un 5, es decir, $M, \langle 3,2\rangle \vDash E_{ab}\neg a_5 \wedge \neg E_{ab}E_{ab}\neg a_5$ ya que $\langle 3,2\rangle \sim_b \langle 3,4\rangle$ y $\langle 3,4\rangle \sim_a \langle 5,4\rangle$ y $M, \langle 5,4\rangle \vDash a_5$. Para el 6 se tiene $M, \langle 3,2\rangle \vDash E_{ab}E_{ab}\neg a_6 \wedge \neg E_{ab}E_{ab}E_{ab}\neg a_6$.

Podemos continuar con este razonamiento infinitamente, hasta poder decir que no es de conocimiento común que Ángel no tiene el número 1235 en su cabeza. Pero sí se tiene que es de conocimiento común que Beatriz no tiene el 1235, porque ambos agentes solo contemplan estados en los cuales el número de Beatriz es par. Por tanto, tenemos:

$$M \vDash (a_3 \wedge b_2) \to (\neg C_{ab}\neg a_{1235} \wedge C_{ab}\neg b_{1235})$$

3.2 Axiomatización

El siguiente sistema lógico, que extiende el sistema **S5** del capítulo anterior, se denotará por **S5C** y, como veremos, constituye una axiomatización completa de las $\mathcal{L}_{KC}$-fórmulas válidas en los modelos epistémicos $\mathcal{S5C}$.

Definición 3.7 (S5C). *Sea A un conjunto finito de agentes y sean los operadores modales K_a ($a \in A$) y C_B ($B \subseteq A$). El sistema **S5C** viene dado por:*

- *instanciaciones de la tautologías proposicionales* — *(Prop)*
- $K_a(\varphi \to \psi) \to (K_a\varphi \to K_a\psi)$ — *(Distribución de K_a sobre $\to$)*
- $K_a\varphi \to \varphi$ — *(Axioma de verdad)*
- $K_a\varphi \to K_aK_a\varphi$ — *(Introspección positiva)*
- $\neg K_a\varphi \to K_a\neg K_a\varphi$ — *(Introspección negativa)*
- $\varphi, \varphi \to \psi \vDash \psi$ — *(Modus Ponens)*
- $\varphi \vDash K_a\varphi$ — *(Necesitación para K_a)*
- $C_B(\varphi \to \psi) \to (C_B\varphi \to C_B\psi)$ — *(Distribución de C_B sobre $\to$)*
- $C_B\varphi \to (\varphi \wedge E_BC_B\varphi)$ — *(Mixto)*
- $C_B(\varphi \to E_B\varphi) \to (\varphi \to C_B\varphi)$ — *(Inducción para C_B)*
- $\varphi \vDash C_B\varphi$ — *(Necesitación para C_B)*

Los siete primeros axiomas definen el sistema **S5**. A partir del octavo, se encuentran los axiomas que añadimos para completar **S5C**. El axioma mixto implica, en particular, que el conocimiento común es verídico $C_B\varphi \to \varphi$. También implica que $C_B\varphi \to E_B^k\varphi$ para cada $k \in \mathbb{N}$. Con el axioma de distribución y la regla de necesitación verificamos que C_B es un operador modal normal. El axioma de inducción típicamente se utilizará para demostrar que una cierta propiedad φ es conocimiento común. Es rutina comprobar que todos los axiomas considerados son válidos para la clase de los $\mathcal{S5C}$-modelos (adecuación del sistema lógico).

La noción de prueba o derivación en **S5C** se traslada a este contexto de manera inmediata. A partir de estos axiomas, es posible obtener las siguientes propiedades derivadas de **S5C**.

Proposición 3.1. Consideramos $a \in B \subseteq A$. De la axiomatización **S5C** se obtiene:

1. $C_B\varphi \leftrightarrow C_BC_B\varphi$
2. $\neg C_B\varphi \leftrightarrow C_B\neg C_B\varphi$
3. $C_B\varphi \leftrightarrow K_aC_B\varphi$
4. $\neg C_B\varphi \leftrightarrow K_a\neg C_B\varphi$

5. $C_B\varphi \to K_{a_1}K_{a_2}K_{a_3}\dots K_{a_n}\varphi$, donde cada $a_i \in B$.
6. $C_B\varphi \to C_{B'}\varphi$ siempre que $B' \subseteq B$.

Demostración. Para ilustrar el uso del sistema lógico, incluimos un esquema de la prueba del último apartado, esto es, **S5C** $\vdash C_B\varphi \to C_{B'}\varphi$.
Antes de comenzar la prueba, es necesario notar que se verifica claramente $E_B\varphi \to E_{B'}\varphi$ $(*)$. Ahora sí, presentamos la demostración.

1. $C_B\varphi \to E_BC_B\varphi$	$mix, Prop, MP$
2. $E_BC_B\varphi \to E_{B'}C_B\varphi$	$(*)$
3. $C_B\varphi \to E_{B'}C_B\varphi$	$1, 2, HS$
4. $C_{B'}(C_B\varphi \to E_{B'}C_B\varphi)$	$3, Nec$
5. $C_{B'}(C_B\varphi \to E_{B'}C_B\varphi) \to (C_B\varphi \to C_{B'}C_B\varphi)$	$inducción\ para\ C_{B'}$
6. $C_B\varphi \to C_{B'}C_B\varphi$	$MP, 4, 5$
7. $C_B\varphi \to \varphi$	$mix, Prop, MP$
8. $C_{B'}(C_B\varphi \to \varphi)$	$7, Nec$
9. $C_{B'}C_B\varphi \to C_{B'}\varphi$	$8, mix, distribución, MP$
10. $C_B\varphi \to C_{B'}\varphi$	$6, 9, HS$

∎

Obsérvese que el primer y el segundo apartado de esta proposición nos da introspección positiva y negativa, respectivamente, para el operador C_B. La última propiedad nos asegura que el conocimiento común se conserva por subgrupos de agentes.

3.3 Completitud de S5C

En esta sección probaremos la completitud del sistema **S5C** para la clase de los modelos epistémicos $\mathcal{S5C}$, esto es, toda fórmula válida en la clase $\mathcal{S5C}$ es demostrable en el sistema de axiomas **S5C** introducido más arriba. La prueba que presentaremos está basada en notas de Rineke Verbrugge en [MH95] y es similar a la realizada en [KP81].

La prueba es similar a la que dimos en el capítulo anterior para **S5**, pero ahora hay una complicación a la hora de construir el modelo canónico para la lógica: en general $\mathcal{S5C}$ no es compacto.

Proposición 3.2. La clase $\mathcal{S5C}$ no es compacta

Demostración. Tomemos dos agentes $a, b \in A$ diferentes y $p \in P$. Consideramos el conjunto de fórmulas

$$\Lambda = \{E_B^n p \mid n \in \mathbb{N}\} \cup \{\neg C_B p\}$$

Es fácil comprobar que cada subconjunto finito de Λ es $\mathcal{S}5\mathcal{C}$-satisfacible, pero el conjunto completo Λ no lo es. ❙

Esto nos impide construir un modelo canónico para el lenguaje completo, tal y como hicimos con $\mathcal{S}5$. Por este motivo, construiremos un modelo canónico para un fragmento finito del lenguaje, dependiendo de la fórmula en la que estamos interesados. El fragmento del lenguaje que consideraremos será la *clausura de la fórmula*, definida como sigue.

Definición 3.8. *Sea la función $cl : \mathcal{L}_{KC} \to \wp(\mathcal{L}_{KC})$ tal que para cada fórmula $\varphi \in \mathcal{L}_{KC}$, $cl(\varphi)$ es el menor conjunto tal que:*

1. *$\varphi \in cl(\varphi)$.*
2. *Si $\psi \in cl(\varphi)$, entonces $Sub(\psi) \subseteq cl(\varphi)$, donde $Sub(\varphi)$ es el conjunto de subfórmulas de ψ.*
3. *Si $\psi \in cl(\varphi)$ y ψ no es una negación, entonces $\neg\psi \in cl(\varphi)$.*
4. *Si $C_B\psi \in cl(\varphi)$, entonces $\{K_a C_B \psi \mid a \in B\} \subseteq cl(\varphi)$.*

Necesitamos demostrar que la clausura de φ es finita para poder construir un modelo canónico sobre ella. Esto lo probaremos con el siguiente lema.

Lema 3.1. *La clausura de φ, $cl(\varphi)$, es finita para cualquier fórmula $\varphi \in \mathcal{L}_{KC}$.*

Demostración. Vamos a probarlo por inducción sobre φ.

- **Caso base.** Si φ es una variable proposicional p, entonces la clausura de φ es $\{p, \neg p\}$ y, por lo tanto, es finita.
- **Hipótesis de inducción.** Suponemos que $cl(\varphi)$ y $cl(\psi)$ son finitos.
- **Paso de inducción.** Distinguimos los siguientes casos:
  - **Negación:** La clausura de $\neg\varphi$ es el conjunto $\{\neg\varphi\} \cup cl(\varphi)$. Como $cl(\varphi)$ es finita por la hipótesis de inducción, el conjunto $\{\neg\varphi\} \cup cl(\varphi)$ es finito.
  - **Conjunción:** La clausura de $(\varphi \wedge \psi)$ es el conjunto $\{(\varphi \wedge \psi), \neg(\varphi \wedge \psi)\} \cup cl(\varphi) \cup cl(\psi)$. Como $cl(\varphi)$ y $cl(\psi)$ son finitos por la hipótesis de inducción, este conjunto es finito. Para el resto de conectivas boolenas, la prueba es similar.
  - **Caso para $K_a\varphi$:** La clausura de $K_a\varphi$ es el conjunto $\{K_a\varphi, \neg K_a\varphi\} \cup cl(\varphi)$. Como $cl(\varphi)$ es finito, el conjunto $\{K_a\varphi, \neg K_a\varphi\} \cup cl(\varphi)$ es finito.

- **Caso para $C_B\varphi$.** La clausura de $C_B\varphi$ es la siguiente $cl(\varphi) \cup \{C_B\varphi, \neg C_B\varphi\} \cup \{K_aC_B\varphi, \neg K_aC_B\varphi \mid a \in B\}$. Como $cl(\varphi)$ es finita por la hipótesis de inducción, este conjunto es finito.

Consideramos ahora conjuntos consistentes maximales en la clausura de una fórmula, no es necesario que sean maximales en todo el lenguaje.

Definición 3.9. *Sea $\Phi \subseteq \mathcal{L}_{KC}$ la clausura de alguna fórmula. Diremos que $\Gamma \subseteq \mathcal{L}_{KC}$ es consistente maximal en Φ si:*

1. $\Gamma \subseteq \Phi$.
2. *Γ es **S5C**-consistente, es decir,* $\Gamma \nvdash_{S5C} \bot$.
3. *Γ es maximal en Φ, es decir, no existe ningún $\Gamma' \subseteq \Phi$ tal que $\Gamma \subset \Gamma'$ y $\Gamma' \nvdash_{S5C} \bot$.*

Estos conjuntos consistentes maximales en Φ serán los estados del modelo canónico para Φ. Como dichos conjuntos son finitos, la conjunción de todos los elementos del conjunto maximal es también una fórmula finita y escribiremos $\underline{\Gamma} = \bigwedge \Gamma$. Ahora definimos el modelo canónico para Φ.

Definición 3.10 (Modelo canónico para Φ). *Sea $\Phi \subseteq \mathcal{L}_{KC}$ la clausura de alguna fórmula. El modelo canónico $M^c = \langle S^c, \sim^c, V^c \rangle$ para Φ está definido como sigue:*

- $S^c = \{\Gamma \mid \Gamma$ *es consistente maximal en* $\Phi\}$
- $\Gamma \sim_a^c \Delta$ *si y solo si* $\{K_a\varphi \mid K_a\varphi \in \Gamma\} = \{K_a\varphi \mid K_a\varphi \in \Delta\}$
- $V_p^c = \{\Gamma \in S^c \mid p \in \Gamma\}$

Puesto que Φ es finito, demostrar el Lema de Lindenbaum en este caso es casi trivial.

Lema 3.2 (Lema de Lindenbaum). *Sea $\Phi \subseteq \mathcal{L}_{KC}$ la clausura de alguna fórmula. Cada subconjunto **S5C**-consistente de Φ puede extenderse a un conjunto consistente maximal en Φ.*

Demostración. Sea $\Delta \subseteq \Phi$ un conjunto **S5C**-consistente de fórmulas, sea $|\Phi| = n$ su cardinal y sea φ_k la k-ésima fórmula en una enumeración de las fórmulas de Φ, $1 \leq k \leq n$. Consideramos la siguiente sucesión de conjuntos de fórmulas:

$$\begin{aligned} \Gamma_0 &= \Delta \\ \Gamma_{k+1} &= \begin{cases} \Gamma_k \cup \{\varphi_{k+1}\} & \text{si } \Gamma_k \cup \{\varphi_{k+1}\} \text{ es } \textbf{S5C}\text{-consistente} \\ \Gamma_k & \text{en otro caso} \end{cases} \end{aligned}$$

Es evidente que $\Delta \subseteq \Gamma_n$. Vamos a probar que Γ_n es consistente maximal en Φ.

Para ver que es **S5C**-consistente, basta con observar que, por inducción sobre k, cada Γ_k es **S5C**-consistente. En particular, lo es Γ_n.

Para ver que es maximal, tomamos una fórmula arbitraria $\varphi_k \in \Phi$ tal que $\varphi_k \notin \Gamma_n$. Por lo tanto, $\Gamma_{k-1} \cup \{\varphi_k\}$ es **S5C**-inconsistente y también lo es $\Gamma_n \cup \{\varphi_k\}$. Ya que φ_k era arbitrario, no existe $\Gamma' \subseteq \Phi$ tal que $\Gamma_n \subset \Gamma'$ y Γ' es **S5C**-consistente.

La semántica del operador de conocimiento común C_B usa la clausura reflexiva transitiva de la unión de las relaciones de accesibilidad de miembros de B. Una forma de mirar la clausura reflexiva transitiva es verla como todos los caminos etiquetados por los agentes en B. La noción de caminos nos será de utilidad en futuras demostraciones.

Definición 3.11 (Caminos).
Sea $B \subseteq A$. Un B-camino desde Γ es una sucesión $\Gamma_0, \dots, \Gamma_n$ de conjuntos consistentes maximales en Φ tal que $\forall k \ (0 \leq k < n)$ existe un agente $a \in B$ tal que $\Gamma_k \sim_a^c \Gamma_{k+1}$ y $\Gamma_0 = \Gamma$.

Un φ-camino es una sucesión $\Gamma_0, \dots, \Gamma_n$ de conjuntos consistentes maximales en Φ tal que $\forall k \ (0 \leq k < n)$ se tiene que $\varphi \in \Gamma_k$. Diremos que la longitud de un camino $\Gamma_0, \dots, \Gamma_n$ es n.

La mayoría de propiedades que usamos para demostrar la completitud de **S5C** son variantes de las que usamos para la completitud de **S5** pero con la salvedad de que actúan sobre la clausura Φ. Aparte de esto, tenemos que añadir una nueva propiedad sobre el operador de conocimiento común.

Lema 3.3. *Sea Φ la clausura de alguna fórmula. Sea $M^c = \langle S^c, \sim^c, V^c \rangle$ el modelo canónico para Φ. Si Γ y Δ son conjuntos consistentes maximales en Φ, entonces:*

1. *Γ es deduciblemente cerrado en Φ, es decir, para todas las fórmulas $\varphi \in \Phi$, si $\vdash_{S5C} \underline{\Gamma} \rightarrow \varphi$, entonces $\varphi \in \Gamma$.*
2. *Si $\neg\varphi \in \Phi$, entonces $\varphi \in \Gamma$ si y solo si $\neg\varphi \notin \Gamma$*
3. *Si $(\varphi \wedge \psi) \in \Phi$, entonces $(\varphi \wedge \psi) \in \Gamma$ si y solo si $\varphi \in \Gamma$ y $\psi \in \Gamma$; y análogamente para el resto de conectivas boolenas.*
4. *Si $\underline{\Gamma} \wedge \hat{K}_a \underline{\Delta}$ es **S5C**-consistente, entonces $\Gamma \sim_a^c \Delta$*
5. *Si $K_a\psi \in \Phi$, entonces $\{K_a\varphi \mid K_a\varphi \in \Gamma\} \vdash_{S5C} \psi$ si y solo si $\{K_a\varphi \mid K_a\varphi \in \Gamma\} \vdash_{S5C} K_a\psi$*
6. *Si $C_B\varphi \in \Phi$, entonces $C_B\varphi \in \Gamma$ si y solo si cada B-camino desde Γ es un φ-camino.*

Demostración. 1. Suponemos $\varphi \in \Phi$. Suponemos también que $\Gamma \vdash \varphi$. Ya que Γ

es consistente, entonces $\Gamma \cup \{\varphi\}$ también lo es. Y ya que Γ es maximal en Φ, se tiene que $\varphi \in \Gamma$.

2. Suponemos $\neg\varphi \in \Phi$ y, entonces, también $\varphi \in \Phi$. Para la implicación de izquierda a derecha, suponemos que $\varphi \in \Gamma$. Por consistencia se tiene que $\neg\varphi \notin \Gamma$. Para la implicación contraria, suponemos que $\neg\varphi \notin \Gamma$. Por ser Γ maximal, $\Gamma \cup \{\neg\varphi\}$ es inconsistente. Por tanto, $\Gamma \vdash \varphi$. De 1, se sigue que $\varphi \in \Gamma$.
3. Suponemos $(\varphi \wedge \psi) \in \Phi$. Para la implicación de izquierda a derecha, suponemos $(\varphi \wedge \psi) \in \Gamma$. Entonces, $\Gamma \vdash \varphi$ y $\Gamma \vdash \psi$. Ya que Φ es cerrado bajo subfórmulas, $\varphi \in \Phi$ y $\psi \in \Phi$. Por lo tanto, aplicando 1, tenemos $\varphi \in \Gamma$ y $\psi \in \Gamma$. Para la implicación de derecha a izquierda, suponemos $\varphi \in \Gamma$ y $\psi \in \Gamma$. Por lo tanto, $\Gamma \vdash (\varphi \wedge \psi)$. Por lo tanto, por el primer apartado de este lema, $(\varphi \wedge \psi) \in \Gamma$.
4. Supongamos que $\underline{\Gamma} \wedge \hat{K}_a\underline{\Delta}$ es consistente. Vamos a realizar la prueba por reducción al absurdo, para ello suponemos que no se tiene $\Gamma \sim_a^c \Delta$. Por lo tanto, existe φ tal que, o bien, $K_a\varphi \in \Gamma$ y $K_a\varphi \notin \Delta$, o bien, $K_a\varphi \notin \Gamma$ y $K_a\varphi \in \Delta$.
   Para el primer caso, mediante el apartado 2 de este lema y utilizando que Φ es cerrado bajo negaciones simples, se tiene $\neg K_a\varphi \in \Delta$. Sin embargo, por la introspección positiva $\Gamma \vdash K_aK_a\varphi$. Sabemos que $K_aK_a \wedge \hat{K}_a\neg K_a\varphi$ es inconsistente. Pero, $\vdash \hat{K}_a\underline{\Delta} \rightarrow \hat{K}_a\neg K_a\varphi$. Por lo tanto, $\underline{\Gamma} \wedge \hat{K}_a\underline{\Delta}$ es inconsistente, contradiciendo nuestra suposición inicial.
   Para el segundo caso, por el apartado 2 de este lema y utilizando el hecho de que Φ es cerrado bajo negaciones simples, se tiene $\neg K_a\varphi \in \Gamma$. Sin embargo, tenemos que $\vdash \hat{K}_a\underline{\Delta} \rightarrow \hat{K}_aK_a\varphi$ y $\vdash \hat{K}_aK_a\varphi \rightarrow K_a\varphi$. Por lo tanto, $\underline{\Gamma} \wedge \hat{K}_a\underline{\Delta}$ es inconsistente, esto contradice nuestra suposición inicial.
5. De derecha a izquierda la demostración es trivial usando el axioma de la verdad. De izquierda a derecha, partimos de $\{K_a\varphi \mid K_a\varphi \in \Gamma\} \vdash \psi$. Por lo tanto, $\vdash \bigwedge\{K_a\varphi \mid K_a\varphi \in \Gamma\} \rightarrow \psi$. Por el axioma de necesidad y distribución se tiene que $\vdash \bigwedge\{K_aK_a\varphi \mid K_a\varphi \in \Gamma\} \rightarrow K_a\varphi$. Por el axioma de introspección positiva $\{K_a\varphi \mid K_a\varphi \in \Gamma\} \vdash \bigwedge\{K_aK_a\varphi \mid K_a\varphi \in \Gamma\}$. Por lo tanto, $\{K_a\varphi \mid K_a\varphi \in \Gamma\} \vdash K_a\varphi$.
6. Vamos a probar la implicación de izquierda a derecha por inducción sobre la longitud del B-camino, pero demostraremos algo más fuerte. Vamos a probar que si $C_B\varphi \in \Gamma$, entonces cada B-camino es un φ-camino y un $C_B\varphi$-camino. Suponemos que $C_B\varphi \in \Gamma$.

   - **Caso base.** Suponemos que la longitud del B-camino es 0, es decir, $\Gamma = \Gamma_0 = \Gamma_n$. Como $\vdash C_B\varphi \rightarrow \varphi$, tanto $C_B\varphi$ como φ pertenecen a Φ, y además, como Γ es deductivamente cerrado en Φ, se tiene que $C_B\varphi \in \Phi$ y $\varphi \in \Gamma$.
   - **Hipótesis de inducción.** "Si $C_B\varphi \in \Gamma$, entonces cada B-camino de longitud n es un φ-camino y un C_B-camino".

- **Paso de inducción.** Tomamos un B-camino de longitud $n+1$ desde Γ. Por la hipótesis de inducción, $C_B\varphi \in \Gamma_n$. Sea a un agente de B tal que $\Gamma_n \sim_a^c \Gamma_{n+1}$. Puesto que $\vdash C_B\varphi \to E_BC_B\varphi$ y $\vdash E_BC_B\varphi \to K_aC_B\varphi$, por la definición de $\sim_a^c$ y, ya que $C_B\varphi \in \Phi$ implica que $K_aC_B\varphi \in \Phi$, obtenemos $K_aC_B\varphi \in \Gamma_{n+1}$. Por un razonamiento similar al del caso base se tiene que $C_B\varphi \in \Gamma_{n+1}$ y $\varphi \in \Gamma_{n+1}$.

Para probar la implicación de derecha a izquierda, suponemos que cada B-camino desde Γ es un φ-camino. Sea $S_{B,\varphi}$ el conjunto de todos los conjuntos Δ consistentes maximales en Φ tal que cada B-camino en Δ es un φ-camino. Definimos la fórmula:

$$\chi = \bigvee_{\Delta \in S_{B,\varphi}} \underline{\Delta}$$

Se tiene que
(a) $\vdash \underline{\Gamma} \to \chi$, porque $\underline{\Gamma}$ es una de las disyunciones de χ.
(b) $\vdash \chi \to \varphi$. Nótese que $\varphi \in \Delta$ si Δ está en $S_{B,\varphi}$. Por lo tanto, φ es una conjunción de cada disyunción de χ y $\vdash \chi \to \varphi$.
(c) $\vdash \chi \to E_B\chi$. Para ello vamos a suponer, buscando una contradicción, que $\chi \wedge \neg E_B\chi$ es consistente. Ya que χ es una disyunción, debería existir una disyunción $\underline{\Delta}$ tal que $\underline{\Delta} \wedge \neg E_B\chi$ es consistente. Mediante un razonamiento similar, debería existir un agente a en B tal que $\underline{\Delta} \wedge \hat{K}_a\neg\chi$ es consistente. Se sabe que $\vdash \bigvee\{\underline{\Delta} \mid \Delta \in S^c\}$, por tanto $\underline{\Delta} \wedge \hat{K}_a \bigvee_{\Theta \in (S^c \setminus S_{B,\varphi})} \underline{\Theta}$ es consistente. Entonces, por la definición de $\hat{K}$, $\underline{\Delta} \wedge \bigvee_{\Theta \in (S^c \setminus S_{B,\varphi})} \hat{K}_a\underline{\Theta}$ es consistente. Por lo tanto, debe existir una $\Theta \notin S_{B,\varphi}$, el cual sea consistente maximal en Φ, tal que $\underline{\Delta} \wedge \hat{K}_a\underline{\Theta}$ es consistente. Usando el apartado 4 de este lema, se tiene que $\Delta \sim_a^c \Theta$. Ya que Θ no está en $S_{B,\varphi}$, debe existir un B-camino desde Θ, el cual no es un φ-camino. Pero esto implica que existe un B-camino desde Δ, el cual no es un φ-camino. Esto contradice que $\Delta \in S_{B,\varphi}$, ya que $\underline{\Delta}$ es una de las disyunciones de χ. Por lo tanto, $\vdash \chi \to E_B\chi$.
Usando (a), (b) y (c) se termina la prueba. Partiendo de $\vdash \chi \to E_B\chi$, por la regla de necesitación se tiene $\vdash C_B(\chi \to E_B\chi)$. Por el axioma de inducción, se tiene que $\vdash \chi \to C_B\chi$. Utilizando (a), obtenemos $\vdash \underline{\Gamma} \to C_B\chi$. Para finalizar, de $\vdash \chi \to \varphi$ y usando necesitación y distribución para $C_B\varphi$ se obtiene $\vdash \underline{\Gamma} \to C_B\varphi$. Por lo tanto, $C_B\varphi \in \Gamma$.

Esto concluye la prueba. ∎

Probadas estas propiedades necesarias para las siguientes demostraciones, pasaremos a demostrar el Lema de la verdad para $\mathcal{S5C}$.

Lema 3.4 (Lema de la verdad). *Sea $\Phi \subseteq \mathcal{L}_{KC}$ la clausura de alguna fórmula. Sea $M^c = \langle S^c, \sim^c, V^c \rangle$ el modelo canónico para Φ. Para todo $\Gamma \in S^c$ y toda fórmula $\varphi \in \Phi$:*

$$\varphi \in \Gamma \textit{ si y solo si } (M^c, \Gamma) \models \varphi$$

Demostración. Suponemos que $\varphi \in \Phi$. Vamos a realizar la prueba por inducción sobre φ.

- **Caso base.** Suponemos que φ es una variable proposicional p. Por definición de V^c se tiene que $p \in \Gamma$ si y solo si $\Gamma \in V^c_p$, lo cual, por la semántica de los modelos de Kripke, es equivalente a $(M^c, \Gamma) \models p$.
- **Hipótesis de inducción.** "Para cada conjunto Γ consistente maximal en se tiene que $\varphi \in \Gamma$ si y solo si $(M^c, \Gamma) \models \varphi$".
- **Paso de inducción.** Para probar los casos de la negación, operadores booleanos y el caso del operador modal K_a se sigue exactamente el mismo razonamiento que en el Lema 2.3. Para el caso de $C_B\varphi$, suponemos que $C_B\varphi \in \Gamma$. De el apartado 6 del lema anterior se sigue que $C_B\varphi \in \Gamma$ si y solo si cada B-camino en Γ es un φ-camino. Por la semántica esto es equivalente a decir que $(M^c, \Gamma) \models C_B\varphi$.

Esto concluye la prueba.

Para tener todos los ingredientes de la prueba del teorema de completitud, nos queda demostrar que el modelo canónico es, después de todo, un $\mathcal{S5C}$ modelo.

Lema 3.5. *Sea Φ la clausura de alguna fórmula. El modelo canónico para Φ es reflexivo, transitivo y Euclídeo.*

Demostración. La prueba de este lema se basa en la definición de $\sim^c_a$ y se prueba de la misma forma que el Lema 2.4 del capítulo anterior.

Para finalizar la sección, podemos ya dar la prueba del teorema de completitud para la lógica epistémica multiagente con conocimiento común **S5C**.

Teorema 3.1 (Completitud). *Para cada fórmula $\varphi \in \mathcal{L}_{KC}$:*

$$\mathcal{S5C} \models \varphi \textit{ implica } \vdash_{\mathbf{S5C}} \varphi.$$

Demostración. Vamos a probarlo por reducción al absurdo. Partimos de $\mathcal{S5C} \models \varphi$ y suponemos que $\nvdash_{\mathbf{S5C}} \varphi$. Por lo tanto, $\{\neg\varphi\}$ es un conjunto **S5C**-consistente. Por el Lema de Lindenbaum, $\{\neg\varphi\}$ es un subconjunto de algún conjunto Γ consistente maximal en $cl(\varphi)$. Sea M^c el modelo canónico para $cl(\varphi)$. Aplicando el Lema de la verdad, $(M^c, \Gamma) \models \neg\varphi$. Puesto que el modelo canónico es un $\mathcal{S}5C$-modelo, se tendría que $\mathcal{S}5C \not\models \varphi$; lo cual contradice nuestra hipótesis de partida. ∎

Observación 3.2. Puesto que la clase $\mathcal{S}5C$ no es compacta, no podemos esperar una versión fuerte del teorema de completitud como en el teorema 2.6. Solo podríamos asegurar completitud fuerte para conjuntos de premisas Γ finitos.

3.4 Expresividad

En esta sección vamos a estudiar la expresividad del lenguaje $\mathcal{L}_{KC}$. Veremos que la noción de conocimiento común no puede ser expresada por $\mathcal{S}5$ cuando en el lenguaje actúan más de un agente.

Empezaremos comparando los lenguajes $\mathcal{L}_K$ y $\mathcal{L}_{KC}$ si solo hay un único agente. Demostraremos que en este caso, y al contrario del caso general, ambos lenguajes son equivalentes sobre $\mathcal{S}5$-modelos.

Teorema 3.2. $\mathcal{L}_K(\{a\}) \equiv \mathcal{L}_{KC}(\{a\})$.

Demostración. Para probar $\mathcal{L}_K(\{a\}) \preceq \mathcal{L}_{KC}(\{a\})$, es suficiente observar que $\mathcal{L}_K(\{a\})$ es un sublenguaje de $\mathcal{L}_{KC}(\{a\})$. Para probar $\mathcal{L}_K(\{a\}) \succeq \mathcal{L}_{KC}(\{a\})$, tomamos una fórmula $\varphi \in \mathcal{L}_{KC}(\{a\})$ y definimos su traducción al lenguaje $\mathcal{L}_K(\{a\})$. Si reemplazamos cada operador C_a en φ por el operador K_a, se tiene una nueva fórmula $\varphi' \in \mathcal{L}_K(\{a\})$. Puesto que $\sim_a = \sim_{\{a\}}$, en $\mathcal{S}5C$ se tiene que $K_a\varphi \leftrightarrow C_a\varphi$ y, por tanto, $\varphi \equiv \varphi'$. ∎

Sin embargo, en el caso multiagente se tiene que

Teorema 3.3. $\mathcal{L}_K \prec \mathcal{L}_{KC}$.

Demostración. Es evidente que $\mathcal{L}_K \preceq \mathcal{L}_{KC}$, ya que $\mathcal{L}_K$ es un sublenguaje de $\mathcal{L}_{KC}$. Falta probar que $\mathcal{L}_K \not\equiv \mathcal{L}_{KC}$ sobre $\mathcal{S}5$-modelos. Para ello necesitamos encontrar $\varphi \in \mathcal{L}_{KC}$ tal que no exista $\psi \in \mathcal{L}_K$ con $\varphi \equiv \psi$ sobre $\mathcal{S}5$-modelos. Con tal fin, consideramos la $\mathcal{L}_{KC}$-fórmula

$$C_{ab}\neg p.$$

Es fácil comprobar que para todo $n \in \mathbb{N}$ se tiene que $(\mathcal{S}p(n), 0) \not\models C_{ab}\neg p$ pero, sin embargo, $(\mathcal{S}p(\omega), 0) \models C_{ab}\neg p$. Utilizando el Teorema 2.17, se obtiene que $C_{ab}\neg p$ no es equivalente a ninguna fórmula en $\mathcal{L}_K$. ∎

El lenguaje $\mathcal{L}_{KC}$, al igual que $\mathcal{L}_K$, no distingue dos modelos si estos son bisimilares. Esto lo vemos en el siguiente teorema.

Teorema 3.4. *Para todos los modelos (M, s) y (M', s'), si $(M, s) \leftrightarrows (M', s')$, entonces $(M, s) \equiv_{\mathcal{L}_{KC}} (M', s')$.*

Demostración. El único caso que no se ha demostrado ya en el paso de inducción de la prueba del Teorema 2.7 es el correspondiente al nuevo operador modal C_B.

Suponemos que $(M, s) \vDash C_B\varphi$. Tomamos un t' arbitrario tal que $(s', t') \in R'^{*}_B$. Aplicando repetidamente la tercera condición de la definición de Bisimulación, encontramos un t tal que $(s, t) \in R^*_B$ y $(t, t') \in \mathfrak{R}$. Por lo tanto, por la hipótesis de inducción, $(M, t) \vDash \varphi$ si y solo si $(M', t') \vDash \varphi$. Puesto que $(M, s) \vDash C_B\varphi$, por la semántica de C_B, $(M, t) \vDash \varphi$; luego $(M', t') \vDash \varphi$. Puesto que t' era arbitrario, $(M', t') \vDash \varphi$ para todo t' tal que $(s', t') \in R'^{*}_B$. Por lo tanto, $(M', s') \vDash C_B\varphi$.

La implicación contraria es análoga pero aplicando la segunda condición de la definición de Bisimulación. ∎

El recíproco del teorema anterior no se cumple. En su lugar, utilizaremos el concepto de juego para dar una condición necesaria y suficiente para la equivalencia entre modelos en el lenguaje $\mathcal{L}_{KC}$. Pero antes, debemos redefinir el concepto de juego añadiendo dos movimientos adicionales al jugador *spoiler*, con sus correspondientes respuestas del jugador *duplicador*. Esta extensión puede ser encontrada en [BMS99]. Por simplicidad, usaremos también la notación $R(B)^*$ para denotar la relación R^*_B correspondiente al operador de conocimiento común C_B.

Definición 3.12 (El juego $\mathcal{L}_{KC}$). *Dados dos modelos $M = \langle S, R, V \rangle$ y $M' = \langle S', R', V' \rangle$ y dados dos estados $s \in S$ y $s' \in S'$, la n-ésima ronda del juego $\mathcal{L}_{KC}$ entre spoiler y duplicador sobre (M, s) y (M', s') es la siguiente: si $n = 0$, spoiler gana si s y s' difieren en sus propiedades atómicas para alguna variable del conjunto de átomos P; en caso contrario, duplicador gana. Si $n \neq 0$, puede ocurrir uno de los siguientes movimientos:*

- *Spoiler elige un agente a y un estado t tal que $(s, t) \in R_a$. Duplicador responderá eligiendo un t' tal que $(s', t') \in R'_a$. La salida de este movimiento será (t, t'). El juego continua con sobre los modelos (M, t) y (M', t').*
- *Spoiler elige un agente a y un estado t' tal que $(s', t') \in R'_a$. Duplicador responderá eligiendo un t tal que $(s, t) \in R_a$. La salida de este movimiento será (t, t'). El juego continua con sobre los modelos (M, t) y (M', t').*
- *Spoiler elige un grupo de agentes B y un estado t tal que $(s, t) \in R(B)^*$. Duplicador responderá eligiendo un estado t' tal que $(s', t') \in R'(B)^*$. La salida de este*

movimiento será (t, t'). Y el juego continuará con los nuevos estados.

- *Spoiler elige un grupo de agentes B y un estado t' tal que $(s', t') \in R'(B)^*$. Duplicador responderá eligiendo un estado t tal que $(s, t) \in R(B)^*$. La salida de este movimiento será (t, t'). Y el juego continuará con los nuevos estados.*

Si un jugador no puede realizar su movimiento, ese jugador pierde. Si los estados de salida son diferentes en sus propiedades atómicas para P, spoiler gana el juego. Si spoiler no ha ganado después de n rondas, duplicador gana el juego.

Hay que destacar que las longitudes de los B-caminos seleccionados por spoiler y duplicador en los dos últimos casos de la definición pueden ser diferentes, ya que utilizamos R^* en el tercer y cuarto movimientos de la definición del juego (el cierre reflexivo transitivo de la realación R) y eso hace que ambos jugadores puedan utilizar distintos caminos para relacionar los dos estados elegidos.

Vamos a demostrar que duplicador tiene una estrategia ganadora si y solo si los modelos satisfacen las mismas fórmulas con la misma profundidad que la ronda del juego en la que se encuentren. Por esto, debemos expandir la definición de profundidad modal para utilizarla en $\mathcal{L}_{KC}$.

Definición 3.13. *La profundidad modal de una fórmula viene dada por la función $d : \mathcal{L}_{KC} \to \mathbb{N}$ definida como sigue*

$$\begin{aligned}
d(p) &= 0 \\
d(\neg\varphi) &= d(\varphi) \\
d(\varphi \wedge \psi) &= max\,(d(\varphi), d(\psi)) \\
d(\varphi \vee \psi) &= max\,(d(\varphi), d(\psi)) \\
d(\varphi \to \psi) &= max\,(d(\varphi), d(\psi)) \\
d(K_a\varphi) &= 1 + d(\varphi) \\
d(C_B\varphi) &= 1 + d(\varphi)
\end{aligned}$$

Vamos a distinguir dentro del lenguaje $\mathcal{L}_{KC}$ diferentes sublenguajes, diferenciándolos por la profundidad máxima de las fórmulas que lo forman.

Definición 3.14. *El lenguaje $\mathcal{L}^n_{KC}$, formado por fórmulas de profundidad n, se define inductivamente como sigue. El lenguaje $\mathcal{L}^0_{KC}$ viene dado por la siguiente notación BNF:*

$$\varphi \ ::= \ p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \varphi \to \varphi$$

El lenguaje $\mathcal{L}^{n+1}_{KC}$ viene dado por:

$$\varphi \ ::= \ \psi \mid K_a\psi \mid C_B\psi \mid \neg\varphi \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \varphi \to \varphi$$

donde $\psi \in \mathcal{L}^n_{KC}$

Al igual que hicimos con $\mathcal{L}^n_K$, debemos probar que solo pueden ser expresadas un número finito de fórmulas (salvo equivalencia). Si no se cumpliera esta condición, no podríamos probar el teorema 3.5 más abajo.

Lema 3.1. Dado un conjunto finito de átomos P, para cada n, existe un número finito de fórmulas diferentes (salvo por equivalencia lógica) en $\mathcal{L}^n_{KC}$.

Demostración. El lema se prueba por inducción sobre n y es similar al lema correspondiente del capítulo anterior. ∎

Para concluir esta sección, demostraremos la condición necesaria y suficiente para la equivalencia elemental entre $S5C$ modelos, gracias a la definición de juegos para el lenguaje $\mathcal{L}^n_{KC}$.

Teorema 3.5. *Para todo $n \in \mathbb{N}$, dados dos modelos cualesquiera $M = \langle S, R, V \rangle$ y $M' = \langle S', R', V' \rangle$ y dado un conjunto de átomos, P, cualquiera, duplicador tiene una estrategia ganadora para la n-ésima ronda del juego $\mathcal{L}_{KC}$ sobre (M, s) y (M', s') si y solo si $(M, s) \equiv_{\mathcal{L}^n_{KC}} (M', s')$.*

Demostración. La prueba es por inducción en n y es, en parte, idéntica a la del Teorema 2.11. La prueba de izquierda a derecha es por inducción sobre la longitud de una fórmula. La única diferencia con la prueba del capítulo anterior es que aquí añadimos un nuevo caso para fórmulas de la forma $C_B\psi$.

Suponemos, sin pérdida de generalidad, que $(M, s) \models C_B\psi$. Tomamos un t' arbitrario tal que $(s', t') \in R'(B)^*$. Ya que duplicador tiene una estrategia ganadora, podemos decir que tiene una respuesta para cada movimiento que spoiler pueda hacer. Así que si spoiler elige t' y un grupo de agentes B, entonces existe t tal que $(s, t) \in R(B)^*$ y duplicador tiene una estrategia ganadora para el subjuego sobre t y t'. Por la hipótesis de inducción se tiene que $(M, t) \models \psi$ si y solo si $(M', t') \models \psi$. Ya que t' era arbitrario, se tiene $(M', t') \models \psi$ para todo t' tal que $(s', t') \in R'(B)^*$. Por lo tanto, por semántica, $(M', s') \models C_B\psi$.

Para la implicación de derecha a izquierda, partimos de $(M, s) \equiv_{\mathcal{L}^{n+1}_{KC}} (M', s')$. Ahora debemos describir la estrategia ganadora de duplicador. Con la definición del juego $\mathcal{L}_{KC}$, si spoiler elige uno de los dos primeros movimientos posibles la estrategia ganadora de duplicador es la misma que en el capítulo anterior.
Podemos suponer, sin perdida de generalidad, que spoiler elige el tercer movimiento posible de la definición. Suponemos, por tanto, que spoiler elige un conjunto de agentes B y un estado t tal que $(s, t) \in R(B)^*$. Tenemos que probar que existe t' tal que

$(s', t') \in R'(B)^*$ y $(M, t) \equiv_{\mathcal{L}^n_{KC}} (M', t')$. Probado eso, por la hipótesis de inducción, duplicador tiene una estrategia ganadora para el subjuego restante.
Vamos a probarlo por reducción al absurdo. Suponemos que no existe dicho t. Esto significa que para cada t' tal que $(s', t') \in R'(B)^*$ spoiler tiene una estrategia ganadora para el subjuego restante. Por la hipótesis de inducción existe una fórmula $\varphi_{t'}$ de profundidad al menos n para cada t' tal que $(s', t') \in R'(B)^*$, donde $(M', t') \vDash \varphi_{t'}$ y $(M, t) \nvDash \varphi_{t'}$. Por el lema anterior, el conjunto $\{\varphi_{t'} \mid (s')(t') \in R'(B)^*\}$ contiene un número finito de fórmulas no equivalentes diferentes. Sea una función f tal que a cada clase de equivalencia $[t']$ le asigna un elemento de la clase. Consideramos la fórmula:

$$\varphi = \bigvee_{(s',t') \in R'(B)^*} f([t'])$$

Dicha fórmula es finita. Además, su profundidad es al menos n. Se tiene que $(M', s') \vDash C_B\varphi$, pero $(M, s) \nvDash C_B\varphi$. Sin embargo, $d(C_B\varphi) \leq (n+1)$. Esto contradice la hipótesis inicial, por lo tanto, duplicador tiene una estrategia ganadora para la $(n+1)$-ésima ronda del juego. ∎

4 Lógica de anuncios públicos

En líneas muy generales, la *lógica epistémica dinámica* engloba el estudio de una familia de lógicas modales diseñadas para describir formalmente acciones que transforman modelos epistémicos. Dichas lógicas permiten expresar propiedades del tipo: "La fórmula F es verdadera tras la ejecución de la acción epistémica A". En este capítulo estudiaremos una de estas lógicas dinámicas, la lógica de los anuncios públicos. Dicha lógica se obtiene a partir de $\mathcal{S}5$ o $\mathcal{S}5C$ añadiendo un nuevo operador modal $[\varphi]\psi$ con el significado "tras anunciar públicamente la fórmula φ, se tiene que la fórmula ψ es el caso".

En la primera sección, introduciremos la noción de anuncios públicos, y hablaremos en detalle del lenguaje y la semántica asociada a este concepto. Después, definiremos de manera formal el sistema de axiomas de la lógica de anuncios públicos **PA** y de la lógica de anuncios públicos con conocimiento común **PAC**. Más adelante, al igual que hicimos en capítulos anteriores, trataremos las propiedades de completitud y expresividad de dichas lógicas.

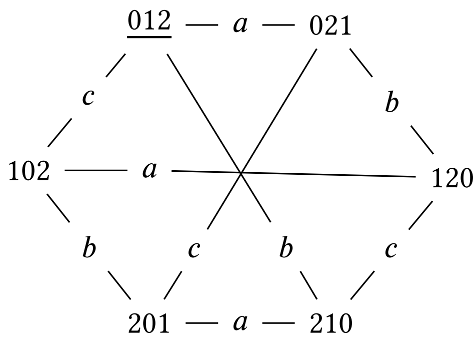
La lógica epistémica multiagente con anuncios públicos sin conocimiento común fue formulada y axiomatizada por Plaza en [Pla89]. La lógica de anuncios públicos con conocimiento común fue axiomatizada por Baltag, Moss y Solecki en [BMS99].

4.1 Sintaxis y semántica

Antes de definir de manera formal el lenguaje de la lógica de anuncios públicos, daremos un ejemplo donde veremos cómo actúa en la práctica un *anuncio público*.

Ejemplo 4.1 (Cartas rusas). Consideramos tres agentes, Ángel, Beatriz y Celia. Cada uno de ellos extrae una carta de un mazo de tres cartas: 0, 1 y 2. Todo esto es comúnmente conocido por los tres agentes. En este ejemplo, tomaremos el caso donde Ángel ha extraído la carta 0 (lo cual notaremos mediante la variable atómica 0_a); Beatriz, el 1 (denotado por 1_b) y Celia, el 2 (que escribiremos como 2_c). Los agentes solo pueden ver su carta. Utilizaremos, como notación de los estados que se nos plantean, un número de tres cifras, donde la primera es la carta de Ángel, la segunda es la de Beatriz y la tercera, la de Celia. Nuestro ejemplo lo expresamos como 012.

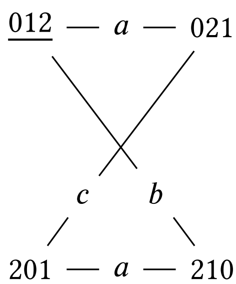
Ángel, con la información que tiene, no puede distinguir los casos 012 y 021. Para Beatriz los estados posibles son 012 y 210. Y para Celia, 012 y 102. De forma general, existen 6 estados posibles relacionados por los tres agentes según la carta que tengan en su mano, y los podemos representar de la siguiente manera:



Antes de empezar con los anuncios públicos, cabe recalcar que suponemos que dichos anuncios son ciertos y completamente públicos, es decir, todos los agentes conocen la fórmula publicada al mismo tiempo y la dan por cierta. El primer anuncio público lo hace Ángel, que dice:

Ángel: "no tengo la carta con el 1".

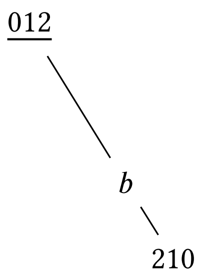
Este anuncio se escribe como la fórmula $\neg 1_a$. Esto hace que en el esquema anterior se actualicen los casos posibles y por tanto nos queda:



Gracias al anuncio público de Ángel, Celia tiene suficiente información para que se pueda afirmar que conoce la carta de Ángel, es decir, $K_c 0_a$. Pero Beatriz sigue sin saberlo, que puede escribirse como $\neg(K_b 0_a \vee K_b 1_a \vee K_b 2_a)$. Si ahora Beatriz hace un segundo anuncio público y dice:

Beatriz: "Sigo sin saber la carta de Ángel".

(esto es, un anuncio público de la fórmula $\neg(K_b 0_a \vee K_b 1_a \vee K_b 2_a)$), este anuncio equivale a una nueva restricción en el espacio de mundos posibles. Antes del segundo anuncio, teníamos dos estados (021 y 201) que Beatriz no relacionaba con ningún otro, y por tanto, se podría decir que, en estos estados, conocía las cartas de los otros dos agentes; y otros dos estados (012 y 210) que Beatriz relacionaba entre ellos y tenía como estados posibles. Después del anuncio público de Beatriz, se eliminan los estados que Beatriz no relaciona con ningún otro. Nos quedaría:



Con estos dos anuncios públicos, tanto Ángel como Celia conocen las cartas de los tres agentes.

Una vez introducidos de manera informal los anuncios públicos con el ejemplo anterior, definimos los lenguajes lógicos en los que vamos a trabajar a partir de ahora.

Definición 4.1. *Dados un conjunto finito de agentes A y un conjunto numerable de átomos P, las fórmulas del lenguaje de la lógica de anuncios públicos $\mathcal{L}_{K[]}$ están definidas inductivamente por la notación BNF:*

$$\varphi \ ::= \ p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid [\varphi]\varphi,$$

donde $a \in A$ y $p \in P$. Si queremos resaltar qué conjuntos A y P estamos considerando, escribiremos también $\mathcal{L}_{K[]}(A, P)$. Cuando añadimos el operador de conocimiento común C_B, obtenemos el lenguaje $\mathcal{L}_{KC[]}$, cuyas fórmulas vienen dadas por:

$$\varphi \ ::= \ p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid C_B\varphi \mid [\varphi]\varphi,$$

donde $a \in A$, $B \subseteq A$ y $p \in P$.

La nueva estructura del lenguaje que hemos introducido, $[\varphi]\psi$, puede leerse como "tras anunciar φ se tiene ψ" o, alternativamente, "tras actualizar el conjunto de los mundos posibles con φ, se tiene que ψ". Por economía del lenguaje, en este último capítulo hemos optado por considerar como primitivas booleanas solamente las conectivas $\{\neg, \wedge\}$ y considerar $\vee, \rightarrow, \leftrightarrow$ como abreviaturas de sus formulaciones equivalentes en términos de $\neg$ y $\wedge$.

Por otra parte, usaremos la notación $\langle\varphi\rangle\psi$ para denotar $\neg[\varphi]\neg\psi$, esto es, el operador modal dual asociado.

Ejemplo 4.2 (Cartas rusas). Siguiendo con el ejemplo anterior, consideramos de nuevo los dos anuncios públicos "Ángel no tiene el 1" y "Beatriz no conoce la carta de Ángel". Si el tercer anuncio público es la distribución de las cartas, se obtiene la proposición "Beatriz sabe la carta de Ángel". La cadena de anuncios públicos, junto con la fórmula que se obtiene, se escribe:

$$[\neg 1_a][\neg(K_b 0_a \vee K_b 1_a \vee K_b 2_a)][0_a \wedge 1_b \wedge 2_c](K_b 0_a \vee K_b 1_a \vee K_b 2_a)$$

Pasamos a definir la semántica de este nuevo lenguaje. Interpretaremos las fórmulas en modelos epistémicos como en los dos capítulos anteriores. La principal diferencia es que ahora tenemos que dotar de significado al operador de anuncios públicos [_]_. Como hemos visto anteriormente, el anuncio de una fórmula es una restricción del conjunto de todos los estados posibles. También podemos entenderlo como una transformación del estado epistémico.

Definición 4.2 (Relación de validez). *Consideremos un modelo epistémico $M = \langle S, \sim, V\rangle$ para un conjunto de agentes A y de átomos P. La relación binaria de validez , $\models$, se define inductivamente de la siguiente forma:*

$M, s \models p$	*sii* $s \in V(p)$
$M, s \models (\varphi \wedge \psi)$	*sii* $M, s \models \varphi$ *y* $M, s \models \psi$
$M, s \models \neg\varphi$	*sii* $M, s \not\models \varphi$
$M, s \models K_a\varphi$	*sii para todo* $t \in S$ *con* $s \sim_a t$ *se tiene* $M, t \models \varphi$
$M, s \models C_B\varphi$	*sii para todo* $t \in S$ *con* $s \sim_B t$ *se tiene* $M, t \models \varphi$
$M, s \models [\varphi]\psi$	*sii* $M, s \models \varphi$ *implica* $M|_\varphi, s \models \psi$

donde $\sim_B \equiv (\bigcup_{a\in B} \sim_a)^*$ *y, con la notación* $[\![\varphi]\!]_M = \{s \in \mathcal{D}(M) \mid M, s \models \varphi\}$*, definimos*

$M|_{\varphi} = \langle S', \sim', V' \rangle$ *como el modelo epistémico dado por:*

$$\begin{aligned} S' &= [\![\varphi]\!]_M \\ \sim'_a &= \sim_a \cap ([\![\varphi]\!]_M \times [\![\varphi]\!]_M) \\ V'(p) &= V(p) \cap [\![\varphi]\!]_M \end{aligned}$$

Observación 4.1.

1. Nótese que, con la definición dada, si $M, s \not\models \varphi$, entonces se tiene que trivialmente $M, s \models [\varphi]\psi$, sea cual sea la fórmula ψ. Por ejemplo, la fórmula $[\bot]p$ es una fórmula válida.
2. Nótese que $M, s \models \langle\varphi\rangle\psi$ sii $M, s \models \varphi$ y $M|_{\varphi}, s \models \psi$. De hecho, $\langle\varphi\rangle\psi \leftrightarrow \varphi \wedge [\varphi]\psi$. En particular, la fórmula $\langle\bot\rangle(p \vee \neg p)$ es, por ejemplo, insatisfacible.

Observación 4.2. En cierto modo de manera contraintuitiva, el anuncio público de una fórmula verdadera φ no siempre nos garantiza que φ sea verdadera tras el anuncio. Por ejemplo, supongamos que tenemos dos agentes, Ana y Bruno, y Ana le anuncia a Bruno la siguiente proposición: "No sabes que la compañía United Agents está haciendo las cosas bien". Esta afirmación implica que "United Agents lo está haciendo bien" y que "Bruno no sabe que United Agents lo está haciendo bien". Por lo tanto, en el momento que se anuncia la proposición queda anulada; pues ahora el único mundo posible que sobrevive es aquel en el que "United Agents lo está haciendo bien" y por tanto Bruno ya sabe esta proposición. Esto es, se tiene que $\langle p \wedge \neg K_b p\rangle\neg(p \wedge K_b p)$ o, equivalentemente, $(p \wedge \neg K_b p) \wedge [p \wedge \neg K_b p]\neg(p \wedge K_b p)$.

Ejemplo 4.3 (Cartas rusas). Notamos el modelo epistémico del ejemplo de las tres cartas como $(Hexa, 012)$. Después de que Ángel anunciara que él no tiene el 1, Celia sabe que Ángel tiene el 0. Formalmente se escribe como: $Hexa, 012 \models [\neg 1_a]K_c 0_a$. Vamos a probarlo.

Se tiene que $Hexa, 012 \models [\neg 1_a]K_c 0_a$ si y solo si "la precondición $Hexa, 012 \models \neg 1_a$ implica $Hexa|_{\neg 1_a}, 012 \models K_c 0_a$". Veamos si $Hexa, 012 \models \neg 1_a$ es cierto o no. Esto se tiene si y solo si $Hexa, 012 \not\models 1_a$, y este es el caso si y solo si $012 \notin V_{1_a} = \{102, 120\}$. Por tanto, el antecedente es cierto.

Nos faltaría probar $Hexa|_{\neg 1_a}, 012 \models K_c 0_a$. Esto, por la definición anterior, es equivalente a: $\forall s \in \mathcal{D}(Hexa|_{\neg 1_a}) : 012 \sim_c s$ implica $Hexa|_{\neg 1_a}, s \models 0_a$. Solo el estado 012 es accesible mediante c desde 012 en el conjunto $\mathcal{D}(Hexa|_{\neg 1_a}) = \{012, 021, 210, 201\}$. Luego la condición de cumple si $Hexa|_{\neg 1_a}, 012 \models 0_a$ y esto es cierto porque $012 \in V_{0_a} = \{012, 021\}$.

Las definiciones de satisfacibilidad y validez de fórmulas vistas en los capítulos anteriores se extienden de manera inmediata al presente lenguaje. Escribiremos PA $\models \varphi$ para indicar que $\varphi \in \mathcal{L}_{K[]}$ es válida para la clase de los modelos epistémicos; y escribiremos PAC $\models \varphi$ para indicar que $\varphi \in \mathcal{L}_{KC[]}$ es válida para la clase de los modelos epistémicos. El conjunto de todas las fórmulas del lenguaje $\mathcal{L}_{K[]}$ válidas en la clase de los modelos epistémicos se denotará también por PA; mientras que si consideramos el lenguaje $\mathcal{L}_{KC[]}$, escribiremos PAC. En la siguiente sección, daremos axiomatizaciones de estos conjuntos.

4.2 Axiomatización

En esta sección presentaremos las axiomatizaciones para los conjuntos de fórmulas válidas PA y PAC. La primera será la lógica de anuncios públicos sin operadores de conocimiento común y la denotaremos **PA**. El segundo sistema lógico, en el que sí incluiremos dichos operadores de conocimiento común, será denotado por **PAC**.

Definición 4.3 (PA). *Dado un conjunto finito de agentes A y un conjunto de variables atómicas P, la axiomatización* **PA***, sobre el lenguaje $\mathcal{L}_{K[]}$, viene dada, además de por las instancias de las tautologías de la lógica proposicional, por:*

- $K_a(\varphi \to \psi) \to (K_a\varphi \to K_a\psi)$ *(Distribución de K_a sobre $\to$)*
- $K_a\varphi \to \varphi$ *(Axioma de verdad)*
- $K_a\varphi \to K_aK_a\varphi$ *(Introspección positiva)*
- $\neg K_a\varphi \to K_a\neg K_a\varphi$ *(Introspección negativa)*
- $\varphi, \varphi \to \psi \models \psi$ *(Modus Ponens)*
- $\varphi \models K_a\varphi$ *(Necesitación para K_a)*
- $[\varphi]p \leftrightarrow (\varphi \to p)$ *(Permanencia atómica)*
- $[\varphi]\neg\psi \leftrightarrow (\varphi \to \neg[\varphi]\psi)$ *(Anuncio y negación)*
- $[\varphi](\psi \wedge \chi) \leftrightarrow ([\varphi]\psi \wedge [\varphi]\chi)$ *(Anuncio y conjunción)*
- $[\varphi]K_a\psi \leftrightarrow (\varphi \to K_a[\varphi]\psi)$ *(Anuncio y conocimiento)*
- $[\varphi][\psi]\chi \leftrightarrow [\varphi \wedge [\varphi]\psi]\chi$ *(Composición de anuncios)*

donde $a \in A$ y $p \in P$.

Ejemplo 4.4. Vamos a probar en **PA** la fórmula $[p]K_ap$. Cuando digamos 'proposicional' querrá decir que aplicaremos varias tautologías y la regla del modus ponens.

1. $p \to p$	tautología
2. $[p]p \leftrightarrow (p \to p)$	permanencia atómica
3. $[p]p$	1,2,proposicional
4. $K_a[p]p$	3,necesitación
5. $p \to K_a[p]p$	4,proposicional
6. $[p]K_a p \leftrightarrow (p \to K_a[p]p)$	anuncio y conocimiento
7. $[p]K_a p$	5,6,proposicional

Proposición 4.1. Sean φ, ψ, χ tres fórmulas dentro de $\mathcal{L}_{K[]}$. Algunas propiedades de **PA** son:

1. Sustitución de igualdades.
   Si $\vdash \psi \leftrightarrow \chi$, entonces $\vdash \varphi(p/\psi) \leftrightarrow \varphi(p/\chi)$
2. Funcionalidad parcial.
   $\vdash (\varphi \to [\varphi]\psi) \leftrightarrow [\varphi]\psi$
3. Anuncios públicos e implicación.
   $\vdash [\varphi](\psi \to \chi) \leftrightarrow ([\varphi]\psi \to [\varphi]\chi)$

La axiomatización **PAC** consiste en añadir axiomas adicionales a la axiomatización **PA**. Veamos la definición de **PAC**.

Definición 4.4 (PAC). *Dados un conjunto finito de agentes A y un conjunto de variables atómicas P, la axiomatización* ***PAC****, sobre el lenguaje $\mathcal{L}_{KC[]}$, viene dada por las instancias de las tautologías de la lógica proposicional y el siguiente conjunto de axiomas:*

• $K_a(\varphi \to \psi) \to (K_a\varphi \to K_a\psi)$	*(Distribución de K_a sobre $\to$)*
• $K_a\varphi \to \varphi$	*(Axioma de verdad)*
• $K_a\varphi \to K_aK_a\varphi$	*(Introspección positiva)*
• $\neg K_a\varphi \to K_a\neg K_a\varphi$	*(Introspección negativa)*
• $\varphi, \varphi \to \psi \vDash \psi$	*(Modus Ponens)*
• $\varphi \vDash K_a\varphi$	*(Necesitación para K_a)*
• $[\varphi]p \leftrightarrow (\varphi \to p)$	*(Permanencia atómica)*
• $[\varphi]\neg\psi \leftrightarrow (\varphi \to \neg[\varphi]\psi)$	*(Anuncio y negación)*
• $[\varphi](\psi \wedge \chi) \leftrightarrow ([\varphi]\psi \wedge [\varphi]\chi)$	*(Anuncio y conjunción)*
• $[\varphi]K_a\psi \leftrightarrow (\varphi \to K_a[\varphi]\psi)$	*(Anuncio y conocimiento)*
• $[\varphi][\psi]\chi \leftrightarrow [\varphi \wedge [\varphi]\psi]\chi$	*(Composición de anuncios)*
• $C_B(\varphi \to \psi) \to (C_B\varphi \to C_B\psi)$	*(Distribución de C_B sobre $\to$)*
• $C_B\varphi \to (\varphi \wedge E_BC_B\varphi)$	*(Axioma mixto de conocimiento común)*

- $C_B(\varphi \to E_B\varphi) \to (\varphi \to C_B\varphi)$ *(Inducción de conocimiento común)*
- $\varphi \vDash C_B\varphi$ *(Necesitación para C_B)*
- $\varphi \vDash [\psi]\varphi$ *(Necesitación para $[\psi]$)*
- $\chi \to [\varphi]\psi, (\chi \wedge \varphi) \to E_B\chi \ \vDash\ \chi \to [\varphi]C_B\psi$ *(Anuncios y conocimiento común)*

donde $a \in A$, $B \subseteq A$ y $p \in P$.

Ejemplo 4.5. Como un ejemplo para ilustrar el uso de este sistema, probaremos que el hecho de que se anuncie alguna proposición atómica la hace de conocimiento común, es decir, probaremos $\vdash_{\mathbf{PAC}} [\neg p]C_A\neg p$.

1. $\neg p \to \neg(\neg p \to p)$	tautología
2. $[\neg p]p \leftrightarrow (\neg p \to p)$	permanencia atómica
3. $\neg p \to \neg[\neg p]p$	1,2,proposicional
4. $[\neg p]\neg p \leftrightarrow (\neg p \to \neg[\neg p]p)$	anuncio y negación
5. $[\neg p]\neg p$	3,4,proposicional
6. $\top \to [\neg p]\neg p$	5,proposicional
7. $\top$	tautología
8. $K_a\top$	7,necesitación
9. $\top \wedge \neg p \to K_a\top$	8,proposicional
10. $\top \wedge \neg p \to E_A\top$	9 para todo $a \in A$,proposicional
11. $\top \to [\neg p]C_A\neg p$	10,6,anuncio y conocimiento común
12. $[\neg p]C_A\neg p$	11,proposicional

Las propiedades que vimos para la axiomatización **PA** en la Proposición 4.1, también se verifican para el sistema de axiomas **PAC**.

Proposición 4.2. Sean φ, ψ, χ tres fórmulas dentro de $\mathcal{L}_{KC[]}$. Algunas propiedades de **PAC** son:

1. Sustitución de igualdades.
   Si $\vdash \psi \leftrightarrow \chi$, entonces $\vdash \varphi(p/\psi) \leftrightarrow \varphi(p/\chi)$
2. Funcionalidad parcial.
   $\vdash (\varphi \to [\varphi]\psi) \leftrightarrow [\varphi]\psi$
3. Anuncios públicos e implicación.
   $\vdash [\varphi](\psi \to \chi) \leftrightarrow ([\varphi]\psi \to [\varphi]\chi)$

4.3 Completitud

Es fácil comprobar que las axiomatizaciones anteriores son adecuadas (si una propiedad es demostrable en ellas, entonces es válida en la clase de los modelos epistémicos). En esta sección probaremos que las dos axiomatizaciones que acabamos de introducir son también completas y, por tanto, capturan el conjunto de las fórmulas válidas en los modelos epistémicos con anuncios públicos, sin y con conocimiento común, respectivamente.

4.3.1 Completitud de **PA**

La prueba de completitud de la lógica de anuncios públicos sin conocimiento común es diferente a las realizadas hasta ahora, tanto para $\mathcal{S}5$ como para $\mathcal{S}5\mathcal{C}$. Veremos que toda fórmula de **PA** puede traducirse a una fórmula equivalente de **S5** y aplicaremos completitud para este último sistema. Cuando observamos el sistema de axiomas **PA**, podemos observar que, en cierto sentido, podemos permutar el operador de anuncios públicos y el resto de operadores del lenguaje. Con la siguiente definición ponemos esto de manifiesto.

Definición 4.5 (Traslación). *La traslación $t : \mathcal{L}_{K[]} \to \mathcal{L}_K$ se define como:*

$$\begin{array}{ll}
t(p) & = p \\
t(\neg\varphi) & = \neg t(\varphi) \\
t(\varphi \wedge \psi) & = t(\varphi) \wedge t(\psi) \\
t(K_a\varphi) & = K_a t(\varphi) \\
t([\varphi]p) & = t(\varphi \to p) \\
t([\varphi]\neg\psi) & = t(\varphi \to \neg[\varphi]\psi) \\
t([\varphi](\psi \wedge \chi)) & = t([\varphi]\psi \wedge [\varphi]\chi) \\
t([\varphi]K_a\psi) & = t(\varphi \to K_a[\varphi]\psi) \\
t([\varphi][\psi]\chi) & = t([\varphi \wedge [\varphi]\psi]\chi)
\end{array}$$

Para la prueba de completitud de **PA** nos será de gran utilidad el hecho de que cada fórmula en $\mathcal{L}_{K[]}$ es equivalente a su traslación. Al igual que fue Plaza quien introdujo la lógica de anuncios públicos, también introdujo en [Pla89] esta técnica para probar la completitud mediante la traslación. La prueba de que son equivalentes la haremos por inducción, pero para ello necesitamos ordenar de alguna manera las fórmulas. Anteriormente lo hacíamos mediante subfórmulas, pero ese método no nos vale en este caso. Para ello definiremos la siguiente medida de complejidad.

Definición 4.6 (Complejidad). *La complejidad $c : \mathcal{L}_{K[]} \to \mathbb{N}$ se define como:*

$$
\begin{aligned}
c(p) &= 1\\
c(\neg\varphi) &= 1 + c(\varphi)\\
c(\varphi \wedge \psi) &= 1 + max(c(\varphi), c(\psi))\\
c(K_a\varphi) &= 1 + c(\varphi)\\
c([\varphi]\psi) &= (4 + c(\varphi)) \cdot c(\psi)
\end{aligned}
$$

El número 4 que aparece en la definición no es arbitrario. Es el menor número natural que permite que se den las siguientes propiedades.

Lema 4.1. *Para todo φ, ψ y χ:*

1. $c(\psi) \geq c(\varphi)$ *si* $\varphi \in Sub(\psi)$
2. $c([\varphi]p) > c(\varphi \to p)$
3. $c([\varphi]\neg\psi) > c(\varphi \to \neg[\varphi]\psi)$
4. $c([\varphi](\psi \wedge \chi)) > c([\varphi]\psi \wedge [\varphi]\chi)$
5. $c([\varphi]K_a\psi) > c(\varphi \to K_a[\varphi]\psi)$
6. $c([\varphi][\psi]\chi) > c([\varphi \wedge [\varphi]\psi]\chi)$

Demostración.

1. Vamos a probarlo por inducción sobre ψ:
   - **Caso base.** Si ψ es una fórmula atómica, su complejidad es 1 y el conjunto de sus subfórmulas es ella misma.
   - **Hipótesis de inducción.** $c(\psi) \geq c(\varphi)$ si $\varphi \in Sub(\psi)$ y $c(\chi) \geq c(\varphi)$ si $\varphi \in Sub(\chi)$.
   - **Paso de inducción.** Se nos presenta los siguientes casos:
     - **Negación.** Suponemos que φ es una subfórmula de $\neg\psi$. Entonces φ, o bien es $\neg\psi$, o bien, es una subfórmula de ψ. En el caso $\varphi = \neg\psi$, es evidente que la complejidad es igual. Por otro lado, si φ es una subfórmula de ψ, se tiene por la hipótesis de inducción que $c(\varphi)$ es como máximo $c(\psi)$. Por tanto, $c(\neg\psi) = 1 + c(\psi) \geq c(\psi)$.
     - **Conjunción.** Suponemos que φ es una subfórmula de $\psi \wedge \chi$. Entonces, o bien, φ es la misma $\psi \wedge \chi$, o bien, es una subfórmula de ψ, o bien, φ es una subfórmula de χ. Para los dos últimos casos se tiene por la hipótesis de inducción. En el caso en que $\varphi = \psi \wedge \chi$, se tiene claramente la igualdad.
     - **Operador epistémico.** Para este caso, la prueba es completamente análoga a la del caso de la conjunción.

- **Anuncio público.** Suponemos que φ es una subfórmula de $[\psi]\chi$. Al igual que antes, o bien tenemos que $\varphi = [\psi]\chi$, o bien es una subfórmula de χ o ψ. En el primer caso, se tiene la igualdad. En el segundo, $c([\psi]\chi) = (4 + c(\psi)) \cdot c(\chi)$, como $c(\chi) \geq c(\varphi)$ y $c(\psi) \geq c(\varphi)$, se tiene que: $c([\psi]\chi) \geq c(\varphi)$.

2. Por un lado:
$$\begin{aligned} c([\varphi]p) &= (4 + c(\varphi)) \cdot 1 \\ &= 4 + c(\varphi) \end{aligned}$$
Por otro lado:
$$\begin{aligned} c(\varphi \to p) &= c(\neg(\varphi \wedge \neg p)) \\ &= 1 + c(\varphi \wedge \neg p) \\ &= 2 + \max(c(\varphi), 2) \end{aligned}$$
Sea cual sea el valor de $c(\varphi)$, se verifica $4 + c(\varphi) > 2 + \max(c(\varphi), 2)$.
3. Por un lado:
$$\begin{aligned} c([\varphi]\neg\psi) &= (4 + c(\varphi)) \cdot (1 + c(\psi)) \\ &= 4 + c(\varphi) + 4 \cdot c(\psi) + c(\varphi) \cdot c(\psi) \end{aligned}$$
Por otro lado:
$$\begin{aligned} c(\varphi \to \neg[\varphi]\psi) &= c(\neg(\varphi \wedge \neg\neg[\varphi]\psi)) \\ &= 1 + c(\varphi \wedge \neg\neg[\varphi]\psi) \\ &= 2 + \max(c(\varphi), 2 + ((4 + c(\varphi)) \cdot c(\psi))) \\ &= 2 + \max(c(\varphi), 2 + 4 \cdot c(\psi) + c(\varphi) \cdot c(\psi)) \end{aligned}$$
Se cumple, tanto $4 + c(\varphi) + 4 \cdot c(\psi) + c(\varphi) \cdot c(\psi) > 2 + c(\varphi)$, como, $4 + c(\varphi) + 4 \cdot c(\psi) + c(\varphi) \cdot c(\psi) > 4 + 4 \cdot c(\psi) + c(\varphi) \cdot c(\psi)$. Por tanto, se cumple la propiedad.
4. Suponemos, sin perdida de generalidad, que $c(\psi) \geq c(\chi)$. Entonces, por un lado:
$$\begin{aligned} c([\varphi](\psi \wedge \chi) &= (4 + c(\varphi)) \cdot (1 + \max(c(\psi), c(\chi))) \\ &= 4 + c(\varphi) + 4 \cdot \max(c(\psi), c(\chi)) + c(\varphi) \cdot \max(c(\psi), c(\chi)) \\ &= 4 + c(\varphi) + 4 \cdot c(\psi) + c(\varphi) \cdot c(\psi) \end{aligned}$$
Por otro lado:
$$\begin{aligned} c([\varphi]\psi \wedge [\varphi]\chi) &= 1 + \max((4 + c(\varphi)) \cdot c(\psi), (4 + c(\varphi)) \cdot c(\chi)) \\ &= 1 + ((4 + c(\varphi)) \cdot c(\psi) \\ &= 1 + 4 \cdot c(\psi) + c(\varphi) \cdot c(\psi) \end{aligned}$$
La segunda ecuación es claramente más pequeña que la primera.

5. Este caso es análogo al segundo caso, el caso de la negación.
6. Por un lado:

$$\begin{aligned} c([\varphi][\psi]\chi) &= (4 + c(\varphi)) \cdot (4 + c(\psi)) \cdot c(\chi) \\ &= (16 + 4 \cdot c(\varphi) + 4 \cdot c(\psi) + c(\varphi) \cdot c(\psi)) \cdot c(\chi) \end{aligned}$$

Por otro lado:

$$\begin{aligned} c([\varphi \wedge [\varphi]\psi]\chi) &= (4 + (1 + \max(c(\varphi), (4 + c(\varphi)) \cdot c(\psi)))) \cdot c(\chi) \\ &= (5 + ((4 + c(\varphi)) \cdot c(\psi))) \cdot c(\chi) \\ &= (5 + 4 \cdot c(\psi) + c(\varphi) \cdot c(\psi)) \cdot c(\chi) \end{aligned}$$

La segunda expresión es más pequeña que la primera.

Esto concluye la prueba.

Con la ayuda de estas propiedades podemos probar que las fórmulas de $\mathcal{L}_{K[]}$ son equivalentes a su traslación.

Lema 4.2. *Para toda fórmula $\varphi \in \mathcal{L}_{K[]}$, se tiene que:*

$$\vdash_{PA} \varphi \leftrightarrow t(\varphi).$$

Además, para toda φ se tiene que $t(\varphi) \in \mathcal{L}_K$.

Demostración. Vamos a demostrarlo por inducción sobre $c(\varphi)$:

- **Caso base.** En el caso donde $c(\varphi) = 1$, φ es una variable atómica y entonces $t(\varphi) = p$. Por tanto, es trivial que $\vdash_{\mathbf{PA}} p \leftrightarrow p$ y $p \in \mathcal{L}_K$.
- **Hipótesis de inducción.** Para toda φ tal que $c(\varphi) \leq n$: $\vdash_{\mathbf{PA}} \varphi \leftrightarrow t(\varphi)$, y $t(\varphi) \in \mathcal{L}_K$.
- **Paso de inducción.** Para los casos de negación, conjunción y el operador epistémico K_a la prueba es sencilla, usando el apartado 1 de Lema 4.1 y la hipótesis de inducción. Ahora, para probar este lema para el operador de anuncios públicos, distinguiremos los siguientes casos.
  - **El caso para** $[\varphi]p$. La primera parte se sigue del axioma de 'permanencia atómica', el apartado 2 del Lema 4.1 y la hipótesis de inducción. Por otra parte, falta ver que $t([\varphi]p) \in \mathcal{L}_K$. Tenemos que $t([\varphi]p) = t(\varphi \to p) = t(\neg(\varphi \wedge \neg p)) = \neg(t(\varphi) \wedge \neg p)$, y esta expresión pertenece a $\mathcal{L}_K$ si $t(\varphi) \in \mathcal{L}_K$, que se verifica por la hipótesis de inducción.
    Las siguientes demostraciones siguen la misma estructura pero usando diferentes elementos del Lema 4.1 y diferentes axiomas de la axiomatización **PA**.

- **El caso para** $[\varphi]\neg\psi$**.** Utilizamos el axioma de 'anuncio y negación' y el apartado 3 del Lema 4.1.
- **El caso para** $[\varphi](\psi \land \chi)$**.** Utilizamos el axioma de 'anuncio y conjunción' y el apartado 4 del Lema 4.1.
- **El caso para** $[\varphi]K_a\psi$**.** Utilizamos el axioma de 'anuncio y conocimiento' y el apartado 5 del Lema 4.1.
- **El caso para** $[\varphi][\psi]\chi$**.** Utilizamos el axioma de 'composición de anuncios' y el apartado 6 del Lema 4.1.

Esto concluye la prueba. ∎

Ahora probaremos la completitud partiendo del Lema 4.2 y el Teorema de Completitud de $\mathcal{S}5$.

Teorema 4.1 (Completitud de PA). *Para cada $\varphi \in \mathcal{L}_{K[]}$ se tiene que*

$$PA \models \varphi \textit{ implica } \vdash_{PA} \varphi.$$

Demostración. Empezamos por suponer que PA $\models \varphi$. Utilizando el Lema 4.2, se tiene $\mathcal{S}5 \models t(\varphi)$. La fórmula $t(\varphi)$ no contiene operadores de anuncios públicos, por lo tanto, gracias a la completitud de **S5** se tiene que $\vdash_{\mathbf{S5}} t(\varphi)$. Puesto que **S5** es un subsistema de **PA**, también se tiene que $\vdash_{\mathbf{PA}} t(\varphi)$. Utilizando de nuevo el Lema 4.2 se sigue que $\vdash_{\mathbf{PA}} \varphi$. ∎

Observación 4.3. Usando ideas similares a las de la prueba anterior, es fácil comprobar que **PA** es también fuertemente completa y que la clase de los modelos epistémicos en el lenguaje $\mathcal{L}_{K[]}$ es compacta. Basta aplicar la traducción $t(\varphi)$ y los correspondientes resultados para $\mathcal{S}5$.

4.3.2 Completitud de **PAC**

En esta sección probaremos la completitud para la lógica de anuncios públicos con conocimiento común. Para ello, combinaremos los procedimientos de las pruebas de completitud de $\mathcal{S}5C$ y PA. De nuevo consideraremos el modelo canónico en un trozo finito del lenguaje. En primer lugar, definiremos la clausura de una fórmula en este contexto.

Definición 4.7. *Sea la función $cl : \mathcal{L}_{KC[]} \to \wp(\mathcal{L}_{KC[]})$ tal que para cada fórmula $\varphi \in \mathcal{L}_{KC[]}$, $cl(\varphi)$ es el menor conjunto de fórmulas tal que:*

1. $\varphi \in cl(\varphi)$.

2. *Si $\psi \in cl(\varphi)$, entonces $Sub(\psi) \subseteq cl(\varphi)$.*
3. *Si $\psi \in cl(\varphi)$ y ψ no es una negación, entonces $\neg\psi \in cl(\varphi)$.*
4. *Si $C_B\psi \in cl(\varphi)$, entonces $\{K_a C_B \psi \mid a \in B\} \subseteq cl(\varphi)$.*
5. *Si $[\psi]p \in cl(\varphi)$, entonces $(\psi \to p) \in cl(\varphi)$.*
6. *Si $[\psi]\neg\chi \in cl(\varphi)$, entonces $(\psi \to \neg[\psi]\chi) \in cl(\varphi)$.*
7. *Si $[\psi](\chi \wedge \xi) \in cl(\varphi)$, entonces $([\psi]\chi \wedge [\psi]\xi) \in cl(\varphi)$.*
8. *Si $[\psi]K_a\chi \in cl(\varphi)$, entonces $(\psi \to K_a[\psi]\xi) \in cl(\varphi)$.*
9. *Si $[\psi]C_B\chi \in cl(\varphi)$, entonces $[\psi]\chi \in cl(\varphi)$ y $\{K_a[\psi]C_B\chi \mid a \in B\} \subseteq cl(\varphi)$.*
10. *Si $[\psi][\chi]\xi \in cl(\varphi)$, entonces $[\psi \wedge [\psi]\chi]\xi \in cl(\varphi)$.*

Como en el capítulo anterior, es fácil comprobar que:

Lema 4.3. *$cl(\varphi)$ es finito para todas las fórmulas $\varphi \in \mathcal{L}_{KC[]}$.*

Las definiciones de conjuntos maximales consistentes en Φ y el modelo canónico para Φ sobre este lenguaje serán iguales a las dadas en el capítulo anterior.

Definición 4.8. *Sea $\Phi \subseteq \mathcal{L}_{KC[]}$ la clausura de alguna fórmula. Diremos que $\Gamma \subseteq \mathcal{L}_{KC[]}$ es consistente maximal en Φ si y solo si:*

1. *$\Gamma \subseteq \Phi$.*
2. *Γ es **PAC**-consistente, es decir, $\Gamma \nvdash_{PAC} \bot$.*
3. *Γ es maximal en Φ, es decir, no existe ningún $\Gamma' \subseteq \Phi$ tal que $\Gamma \subset \Gamma'$ y $\Gamma' \nvdash_{PAC} \bot$.*

Definición 4.9 (Modelo canónico para Φ). *Sea $\Phi \subseteq \mathcal{L}_{KC[]}$ la clausura de alguna fórmula. El modelo canónico $M^c = \langle S^c, \sim^c, V^c \rangle$ para Φ está definido como:*

- *$S^c = \{\Gamma \mid \Gamma$ es consistente maximal en $\Phi\}$*
- *$\Gamma \sim_a^c \Delta$ si y solo si $\{K_a\varphi \mid K_a\varphi \in \Gamma\} = \{K_a\varphi \mid K_a\varphi \in \Delta\}$*
- *$V_p^c = \{\Gamma \in S^c \mid p \in \Gamma\}$*

Para la demostración de completitud del sistema **PAC**, necesitaremos el Lema de Lindenbaum. La prueba de este lema en este lenguaje es exactamente la misma a la dada para el lenguaje $S5C$ y la omitimos.

Lema 4.4 (Lema de Lindenbaum). *Sea $\Phi \subseteq \mathcal{L}_{KC[]}$ la clausura de alguna fórmula. Cada subconjunto **PAC**-consistente de Φ puede extenderse a un conjunto consistente maximal en Φ.*

Nuestro próximo objetivo es demostrar el correspondiente Lema de la Verdad en este contexto. En dicha prueba, hemos de abordar en esencia solamente un nuevo caso: las fórmulas de la forma $[\varphi]C_B\psi$. Para verificar el Lema en este nuevo caso, nos será útil definir un nuevo concepto, los B-φ-caminos.

Definición 4.10 (B-φ-camino). *Sea $B \subseteq A$ un subconjunto de agentes del lenguaje. Un B-φ-camino desde Γ es un B-camino que también es un φ-camino.*

Los modelos canónicos y los conjuntos consistentes maximales cumplen una nueva propiedad para las fórmulas de la forma $[\varphi]C_B\psi$. Dicha propiedad aparece en el siguiente Lema, el cual es una ampliación del Lema 3.3.

Lema 4.5. *Sea Φ la clausura de alguna fórmula. Sea $M^c = \langle S^c, \sim^c, V^c \rangle$ el modelo canónico para Φ. Si Γ y Δ son conjuntos consistentes maximales en Φ, entonces:*

1. *Γ es deductivamente cerrado en Φ, es decir, para todas las fórmulas $\varphi \in \Phi$, si $\vdash_{PAC} \underline{\Gamma} \to \varphi$, entonces $\varphi \in \Gamma$.*
2. *Si $\neg\varphi \in \Phi$, entonces $\varphi \in \Gamma$ si y solo si $\neg\varphi \notin \Gamma$.*
3. *Si $(\varphi \wedge \psi) \in \Phi$, entonces $(\varphi \wedge \psi) \in \Gamma$ si y solo si $\varphi \in \Gamma$ y $\psi \in \Gamma$*
4. *Si $\underline{\Gamma} \wedge \hat{K}_a\underline{\Delta}$ es **PAC**-consistente, entonces $\Gamma \sim_a^c \Delta$.*
5. *Si $K_a\psi \in \Phi$, entonces $\{K_a\varphi \mid K_a\varphi \in \Gamma\} \vdash_{PAC} \psi$ si y solo si $\{K_a\varphi \mid K_a\varphi \in \Gamma\}_{PAC} \vdash K_a\psi$.*
6. *Si $C_B\varphi \in \Phi$, entonces $C_B\varphi \in \Gamma$ si y solo si cada B-camino desde Γ es un φ-camino.*
7. *Si $[\varphi]C_B\psi \in \Phi$, entonces $[\varphi]C_B\psi \in \Gamma$ si y solo si cada B-φ-camino desde Γ es un $[\varphi]\psi$-camino*

Demostración. Los apartados desde el 1 hasta el 6 se demuestran de manera similar a los del Lema 3.3. Veamos la prueba del último apartado.

7. Probaremos la implicación de izquierda a derecha mediante una inducción sobre la longitud del camino. De hecho, probaremos por inducción una propiedad más fuerte: si $[\varphi]C_B\psi \in \Gamma$, entonces cada B-φ-camino es un $[\varphi]\psi$-camino y un $[\varphi]C_B\psi$-camino. Suponemos $[\varphi]C_B\psi \in \Gamma$ e iniciamos la inducción:
   - **Caso base.** Suponemos que la longitud de el B-φ-camino es 0. Por lo tanto, tenemos que probar que si $[\varphi]C_B\psi \in \Gamma_0$, entonces $[\varphi]\psi \in \Gamma_0$. Se tiene que $\vdash_{\mathbf{PAC}} C_B\psi \to \psi$ y entonces, por necesitación y distribución de $[\varphi]$, se tiene que $\vdash_{\mathbf{PAC}} [\varphi]C_B\psi \to [\varphi]\psi$. Utilizando el apartado 9 de la definición de clausura y que Γ_0 es deductivamente cerrado en la clausura, obtenemos $[\varphi]\psi \in \Gamma_0$.
   - **Hipótesis de inducción.** Cada B-φ-camino de longitud n es un $[\varphi]\psi$-camino y un $[\varphi]C_B\psi$-camino.
   - **Paso de inducción.** Suponemos que el B-φ-camino es de longitud $n+1$, es decir, que existe un B-φ-camino $\Gamma_0, \Gamma_1, \ldots, \Gamma_n, \Gamma_{n+1}$. Suponemos que $\Gamma_n \sim_a^c \Gamma_{n+1}$. Por la hipótesis de inducción podemos suponer que $[\varphi]C_B\psi \in$

Γ_n. Por el axioma mixto tenemos, $\vdash_{\mathbf{PAC}} C_B\psi \to K_a C_B\psi$. Utilizando necesitación y el axioma de distribución de $[\varphi]$, tenemos $\vdash_{\mathbf{PAC}} [\varphi]C_B\psi \to [\varphi]K_a C_B\psi$. Por el axioma de anuncio y conocimiento, obtenemos $\vdash_{\mathbf{PAC}} [\varphi]K_a C_B\psi \to (\varphi \to K_a[\varphi]C_B\psi)$. Ya que $\varphi \in \Gamma_n$, se tiene $\Gamma_n \vdash K_a[\varphi]C_B\psi$. Por lo tanto, por la definición de $\sim_a^c$, $K_a[\varphi]C_B\psi \in \Gamma_{n+1}$ y $[\varphi]C_B\psi \in \Gamma_{n+1}$. Usando el razonamiento del caso base, $[\varphi]\psi \in \Gamma_{n+1}$.

Para la implicación de derecha a izquierda, en primer lugar suponemos que cada B-φ-camino desde Γ es un $[\varphi]\psi$-camino. Sea $S_{B,\varphi,[\varphi]\psi}$ el conjunto de los conjuntos consistentes maximales, Δ, tales que cada B-φ-camino desde Δ es un $[\varphi]\psi$-camino. Ahora definimos la fórmula:

$$\chi = \bigvee_{\Delta \in S_{B,\varphi,[\varphi]\psi}} \underline{\Delta}$$

Probaremos en **PAC** lo siguiente:

a) $\vdash \underline{\Gamma} \to \chi$
Se prueba fácilmente, ya que $\Gamma \in S_{B,\varphi,[\varphi]\psi}$ y $\underline{\Gamma}$ es una de las disyunciones de χ.

b) $\vdash \chi \to [\varphi]\psi$
Puesto que cada B-φ-camino es un $[\varphi]\psi$-camino, para cada $\Delta \in S_{B,\varphi,[\varphi]\psi}$, se tiene que $[\varphi]\psi \in \Delta$. Por lo tanto, $[\varphi]\psi$ es una de las conjunciones en cada disyunción de χ. Esto hace que se verifique $\chi \to [\varphi]\psi$.

c) $\vdash \chi \wedge \varphi \to E_B\chi$
Buscando una contradicción, suponemos $\chi \wedge \varphi \wedge \neg E_B\chi$ es **PAC**-consistente. Como χ es una disyunción, debe existir un $\underline{\Theta}$ tal que $\underline{\Theta} \wedge \varphi \wedge \neg E_B\chi$ es **PAC**-consistente. Puesto que $\underline{\Theta} \wedge \varphi$ es **PAC**-consistente y Θ es consistente maximal en Φ y $\varphi \in \Phi$, se tiene que $\varphi \in \Theta$. Por lo tanto, $\underline{\Theta} \wedge \neg E_B\chi$ es **PAC**-consistente. Entonces, debería existir un agente a tal que $\underline{\Theta} \wedge \hat{K}_a\neg\chi$ es **PAC**-consistente. Puesto que $\vdash \bigvee\{\underline{\Gamma} \mid \Gamma \in S^c\}$, debería existir un Ξ en el complemento de $S_{B,\varphi,[\varphi]\psi}$ tal que $\underline{\Theta} \wedge \hat{K}_a\underline{\Xi}$ es **PAC**-consistente. Por el apartado 4 del Lema 4.5, esto es equivalente a $\Theta \sim_a^c \Xi$. Pero como existe un B-φ-camino desde Ξ, el cual no es un $[\varphi]\psi$-camino y $a \in B$, también existe un B-φ-camino desde Θ, el cual no es un $[\varphi]\psi$-camino. Esto está en contradicción con $\Theta \in S_{B,\varphi,[\varphi]\psi}$. Por lo tanto, $\vdash \chi \wedge \varphi \to E_B\chi$.

Con esto probado, aplicando la regla de anuncio y conocimiento común a (b) y a (c) se sigue que $\vdash_{\mathbf{PAC}} \chi \to [\varphi]C_B\psi$. Y por (a) se tiene que $\vdash_{\mathbf{PAC}} \underline{\Gamma} \to [\varphi]C_B\psi$. Por lo tanto, $[\varphi]C_B\psi \in \Gamma$.

Esto concluye la prueba. ∎

Para probar el Lema de la verdad necesitaremos el concepto de medida de complejidad, que nos servirá para aplicar la hipótesis de inducción sobre fórmulas que no son subfórmulas de las anteriormente consideradas. La definición de medida de complejidad que daremos a continuación será simplemente una ampliación de la definida para probar la completitud de **PA**.

Definición 4.11 (Complejidad). *La complejidad $c : \mathcal{L}_{KC[]} \to \mathbb{N}$ se define como:*

$$
\begin{aligned}
c(p) &= 1\\
c(\neg\varphi) &= 1 + c(\varphi)\\
c(\varphi \wedge \psi) &= 1 + max(c(\varphi), c(\psi))\\
c(K_a\varphi) &= 1 + c(\varphi)\\
c(C_B\varphi) &= 1 + c(\varphi)\\
c([\varphi]\psi) &= (4 + c(\varphi)) \cdot c(\psi)
\end{aligned}
$$

Lema 4.6. *Para todo φ, ψ y χ:*

1. $c(\psi) \geq c(\varphi)$ *si* $\varphi \in Sub(\psi)$
2. $c([\varphi]p) > c(\varphi \to p)$
3. $c([\varphi]\neg\psi) > c(\varphi \to \neg[\varphi]\psi)$
4. $c([\varphi](\psi \wedge \chi)) > c([\varphi]\psi \wedge [\varphi]\chi)$
5. $c([\varphi]K_a\psi) > c(\varphi \to K_a[\varphi]\psi)$
6. $c([\varphi]C_B\psi) > c([\varphi]\psi)$
7. $c([\varphi][\psi]\chi) > c([\varphi \wedge [\varphi]\psi]\chi)$

Demostración. El único apartado de este lema que no está probado es el 6. Los demás están demostrados en el Lema 4.1. Por un lado:

$$
\begin{aligned}
c([\varphi]C_B\psi) &= (4 + c(\varphi)) \cdot (1 + c(\psi))\\
&= 4 + c(\varphi) + 4 \cdot c(\psi) + c(\varphi) \cdot c(\psi)
\end{aligned}
$$

Por otro lado:

$$
\begin{aligned}
c([\varphi]\psi) &= (4 + c(\varphi)) \cdot c(\psi)\\
&= 4 \cdot c(\psi) + c(\varphi) \cdot c(\psi)
\end{aligned}
$$

Se cumple claramente la desigualdad. ∎

Ahora podemos probar el Lema de la verdad.

Lema 4.7 (Lema de la verdad). *Sea Φ la clausura de alguna fórmula. Sea $M^c = \langle S^c, \sim^c, V^c\rangle$ el modelo canónico para Φ. Para todo $\Gamma \in S^c$ y toda fórmula $\varphi \in \Phi$:*

$$\varphi \in \Gamma \text{ si y solo si } (M^c, \Gamma) \vDash \varphi.$$

Demostración. Suponemos que $\varphi \in \Phi$. Vamos a realizar la prueba por inducción sobre $c(\varphi)$.

- **Caso base.** Suponemos que φ es una variable proposicional p. Por definición de V^c se tiene que $p \in \Gamma$ si y solo si $\Gamma \in V^c_p$, lo cual por la semántica es equivalente a decir que $(M^c, \Gamma) \vDash p$.
- **Hipótesis de inducción.** Para cada φ tal que $c(\varphi) \leq n$: $\varphi \in \Gamma$ si y solo si $(M^c, \Gamma) \vDash \varphi$
- **Paso de inducción.** Suponemos $c(\varphi) = n + 1$. Los casos de negación, conjunción, operadores individuales epistémicos y operadores de conocimiento común están probados en el Lema de la verdad del capítulo anterior. Cuando φ es de la forma $[\psi]\chi$ distinguimos los siguientes casos:
  - **El caso para $[\psi]p$.** Suponemos $[\psi]p \in \Gamma$. Dado que $[\psi]p \in \Phi$, $[\psi]p \in \Gamma$ es equivalente a $(\psi \to p) \in \Gamma$ por el axioma de permanencia atómica. Por el apartado 2 del Lema 4.6, podemos aplicar la hipótesis de inducción. Por lo tanto, esto es equivalente a $(M^c, \Gamma) \vDash \psi \to p$. Y, por semántica, se tiene $(M^c, \Gamma) \vDash [\psi]p$.
  - **El caso para $[\psi]\neg\chi$.** Suponemos $[\psi]\neg\chi \in \Gamma$. Dado que $[\psi]\neg\chi \in \Phi$, $[\psi]\neg\chi \in \Gamma$ es equivalente a $(\psi \to \neg[\psi]\chi) \in \Gamma$ por el axioma de anuncio y negación. Por el apartado 3 del Lema 4.6, podemos aplicar la hipótesis de inducción. Por lo tanto, esto es equivalente a $(M^c, \Gamma) \vDash \psi \to \neg[\psi]\chi$. Y por semántica, se tiene $(M^c, \Gamma) \vDash [\psi]\neg\chi$.
  - **El caso para $[\psi](\chi \wedge \xi)$.** Suponemos $[\psi](\chi \wedge \xi) \in \Gamma$. Dado que $[\psi](\chi \wedge \xi) \in \Phi$, $[\psi](\chi \wedge \xi) \in \Gamma$ es equivalente a $([\psi]\chi \wedge [\psi]\xi) \in \Gamma$ por el axioma de anuncio y conjunción. Por el apartado 4 del Lema 4.6, podemos aplicar la hipótesis de inducción. Por lo tanto, esto es equivalente a $(M^c, \Gamma) \vDash [\psi]\chi \wedge [\psi]\xi$. Y, por semántica, se tiene $(M^c, \Gamma) \vDash [\psi](\chi \wedge \xi)$.
  - **El caso para $[\psi]K_a\chi$.** Suponemos $[\psi]K_a\chi \in \Gamma$. Dado que $[\psi]K_a\chi \in \Phi$, $[\psi]K_a\chi \in \Gamma$ es equivalente a $(\psi \to K_a[\psi]\chi) \in \Gamma$ por el axioma de anuncio y conocimiento. Por el apartado 4 del Lema 4.6, podemos aplicar la hipótesis de inducción. Por lo tanto, esto es equivalente a $(M^c, \Gamma) \vDash \psi \to K_a[\psi]\chi$. Y, por semántica, se tiene $(M^c, \Gamma) \vDash [\psi]K_a\chi$.
  - **El caso para $[\psi]C_B\chi$.** Suponemos $[\psi]C_B\chi \in \Gamma$. Dado que $[\psi]C_B\chi \in \Phi$, $[\psi]C_B\chi \in \Gamma$ si y solo si cada B-ψ-camino desde Γ es un $[\psi]\chi$-camino por el apartado 6 del Lema 4.5. Por el apartado 6 del Lema 4.6, podemos aplicar la hipótesis de inducción. Por lo tanto, esto es equivalente a que cada B-ψ-camino desde Γ es un camino donde $[\psi]\chi$ es cierto. Y, por semántica, se tiene $(M^c, \Gamma) \vDash [\psi]C_B\chi$.
  - **El caso para $[\psi][\chi]\xi$.** Suponemos $[\psi][\chi]\xi \in \Gamma$. Dado que $[\psi][\chi]\xi \in \Phi$,

$[\psi][\chi]\xi \in \Gamma$ es equivalente a $[\psi \wedge [\psi]\chi]\xi \in \Gamma$ por el axioma de composición de anuncios. Por el apartado 7 del Lema 4.6, podemos aplicar la hipótesis de inducción. Por lo tanto, esto es equivalente a $(M^c, \Gamma) \vDash [\psi \wedge [\psi]\chi]\xi$. Y, por semántica, se tiene $(M^c, \Gamma) \vDash [\psi][\chi]\xi$.

Esto concluye la prueba. ∎

Una vez más, hemos de comprobar que el modelo canónico es, después de todo, un modelo epistémico.

Lema 4.8. *Sea Φ la clausura de alguna fórmula. El modelo canónico para Φ es reflexivo, transitivo y Euclídeo.*

Demostración. La prueba de este lema se basa en la definición de $\sim_a^c$ y se prueba de la misma forma que el Lema 2.4. ∎

Con todo esto, nos disponemos a probar el Teorema de Completitud para el lenguaje de anuncios públicos con conocimiento común.

Teorema 4.2 (Completitud). *Para cada $\varphi \in \mathcal{L}_{KC[]}$, se tiene que:*

$$PAC \vDash \varphi \text{ implica } \vdash_{\mathbf{PAC}} \varphi.$$

Demostración. Probaremos el contra-recíproco. Suponemos $\nvdash_{\mathbf{PAC}} \varphi$. Por lo tanto, $\{\neg\varphi\}$ es un conjunto **PAC**-consistente. Por el Lema de Lindenbaum, existe un conjunto Γ el cual es consistente maximal en $\Phi = cl(\neg\varphi)$ y tal que $\neg\varphi \in \Gamma$. Sea M^c el modelo canónico para $cl(\neg\varphi)$. Usando el Lema de la verdad, $(M^c, \Gamma) \vDash \neg\varphi$; y por el lema 4.8, M^c es un modelo epistémico. Por lo tanto, PAC $\nvDash \varphi$. ∎

Observación 4.4. El mismo ejemplo que consideramos en el capítulo anterior muestra que la clase de modelos epistémicos en el lenguaje $\mathcal{L}_{KC[]}$ no es compacta. En consecuencia, no podemos esperar un resultados de completitud fuerte para cualquier conjunto de premisas Γ.

4.4 Expresividad

En esta sección estudiamos cómo varía el poder expresivo de nuestro lenguaje epistémico si añadimos al lenguaje los operadores de anuncios públicos. Como veremos, la situación es completamente distinta en presencia de operadores de conocimiento común o en ausencia de ellos.

4.4.1 PA

La prueba de completitud del sistema **PA** de la sección anterior nos será de gran utilidad para comparar el poder expresivo de los lenguajes $\mathcal{L}_{K[]}$ y $\mathcal{L}_K$ sobre modelos epistémicos. Utilizaremos la función de traslación dada en la Definición 4.5, la cual traduce una fórmula con operadores de anuncios públicos, φ, a una fórmula equivalente del lenguaje epistémico sin operador de anuncios, $t(\varphi)$. Usando dicha traslación, probaremos que los lenguajes $\mathcal{L}_{K[]}$ y $\mathcal{L}_K$ son equivalentes (sobre modelos epistémicos).

Teorema 4.3. $\mathcal{L}_K \equiv \mathcal{L}_{K[]}$.

Demostración. Puesto que $\mathcal{L}_K$ es un sublenguaje de $\mathcal{L}_{K[]}$, se tiene que $\mathcal{L}_K \preceq \mathcal{L}_{K[]}$. Falta probar $\mathcal{L}_{K[]} \preceq \mathcal{L}_K$. Dada $\varphi \in \mathcal{L}_{K[]}$, por la definición de traslación y utilizando el Lema 4.2, se tiene que $\vdash_{\mathbf{PA}} \varphi \leftrightarrow t(\varphi)$. Sea (M, s) un modelo epistémico cualquiera. Por la adecuación de **PA**, se tiene que $M, s \vDash \varphi \leftrightarrow t(\varphi)$. Por tanto, $M, s \vDash \varphi$ si y solo si $M, s \vDash t(\varphi)$, y $t(\varphi)$ no contiene ninguna ocurrencia del operador de anuncios públicos. ∎

El resultado anterior es bastante sorprendente en tanto en cuanto nos dice que el operador de anuncios públicos no añade capacidad expresiva al lenguaje de la lógica epistémica miltiagente aunque, a primera vista, dicho operador parece expresar un concepto que no está presente en el lenguaje. Sin embargo, cabe destacar que esta traducción al lenguaje $\mathcal{L}_K$ no es trivial y viene con un precio: la longitud de la fórmula traducida $t(\varphi)$ puede ser exponencialmente más grande que la longitud de la fórmula original φ.

4.4.2 PAC

En esta sección veremos que añadir operadores de anuncios público sí incrementa la potencia expresiva del lenguaje de la lógica multiagente en presencia de operadores de *conocimiento común*, esto es, el lenguaje $\mathcal{L}_{KC}$ es menos expresivo que $\mathcal{L}_{KC[]}$.

Anteriormente vimos que $\mathcal{S}5C$ puede distinguir modelos de espinas, por lo que necesitamos buscar otro tipo de modelos que $\mathcal{S}5C$ no pueda distinguir y PAC sí. Para distinguir modelos, el juego $\mathcal{L}_{KC}$ utilizaba que spoiler podía dar un salto hasta el final de un modelo finito. Tenemos que encontrar modelos, en los cuales, este salto no sea útil. Con tal fin, definimos los *modelos horquilla*.

Definición 4.12 (Modelos Horquilla(n)). *Definimos $Horquilla(n) = \langle S, \sim, V \rangle$, donde:*

- $S = \{s_m \mid m \leq (n + (n \; mod \, 2))\} \cup \{t_m \mid m \leq (n + (n \; mod \, 2))\} \cup \{u, v\}$
- $s_m \sim_a s_k$ *si y solo si* $min(m,k) \; mod \, 2 = 0$ *y* $|m - k| = 1$
  $t_m \sim_a t_k$ *si y solo si* $min(m,k) \; mod \, 2 = 0$ *y* $|m - k| = 1$
  $u \sim_a v$
- $s_m \sim_b s_k$ *si y solo si* $min(m,k) \; mod \, 2 = 1$ *y* $|m - k| = 1$
  $t_m \sim_b t_k$ *si y solo si* $min(m,k) \; mod \, 2 = 1$ *y* $|m - k| = 1$
  $u \sim_b s_{(n+(n \; mod \, 2))}$
  $v \sim_b t_{(n+(n \; mod \, 2))}$
- $V(p) = \{u\}$

Este modelo está formado por dos modelos espinas que se unen en el estado donde p es válido. Como consecuencia de esta definición, observamos que los modelos $Horquilla(n)$ y $Horquilla(n+1)$ son idénticos si n es impar. La representación del modelo es la siguiente:

$$\begin{tikzcd}
s_0 \arrow[r, leftrightarrow, "a"] & s_1 \arrow[r, leftrightarrow, "b"] & s_2 \arrow[r, leftrightarrow, "a"] & s_3 \arrow[r, leftrightarrow] & \cdots \arrow[r, leftrightarrow] & s_{n-1} \arrow[r, leftrightarrow, "a"] & s_n \arrow[r, leftrightarrow, "b"] & \overset{p}{u} \arrow[d, leftrightarrow, "a"] \\
t_0 \arrow[r, leftrightarrow, "a"] & t_1 \arrow[r, leftrightarrow, "b"] & t_2 \arrow[r, leftrightarrow, "a"] & t_3 \arrow[r, leftrightarrow] & \cdots \arrow[r, leftrightarrow] & t_{n-1} \arrow[r, leftrightarrow, "a"] & t_n \arrow[r, leftrightarrow, "b"] & v
\end{tikzcd}$$

Aplicando el juego $\mathcal{L}_{KC}$ sobre dos estados del modelo horquilla, comprobamos que este lenguaje no puede distinguir estos dos estados. Consideramos $(Horquilla(n), s_0)$ y $(Horquilla(n), t_0)$. En primer lugar, de los cuatro movimientos que puede realizar spoiler definidos en la Definición 3.12, el tercero y el cuarto no le son de utilidad, ya que si toma un grupo de un único agente, los movimientos se reducen a los dos primeros. Y si el grupo que toma es $\{a, b\}$, duplicador puede repetir el movimiento, ya que se trata del mismo modelo. En segundo lugar, el estado donde p es válido no es alcanzable utilizando los dos primeros movimientos de la definición, ya que serían necesarias $(n+1)$ rondas y solo tenemos a nuestra disposición n. Por tanto, se tiene que:

Lema 4.1. Para todo $n \in \mathbb{N}$, se tiene:

$$(Horquilla(n), s_0) \equiv_{\mathcal{L}^n_{KC}} (Horquilla(n), t_0)$$

Como corolario obtenemos el siguiente resultado:

Corolario 4.1. No existe ninguna fórmula $\varphi \in \mathcal{L}_{KC}$ tal que para todo $n \in \mathbb{N}$, se verifique $(Horquilla(n), s_0) \vDash \varphi$ y $(Horquilla(n), t_0) \nvDash \varphi$.

Demostración. Por reducción al absurdo, suponemos que sí existe dicha fórmula. Consideramos que tiene profundidad m. Por tanto, utilizando el lema anterior tenemos $(Horquilla(m), s_0) \vDash \varphi$ y $(Horquilla(m), t_0) \vDash \varphi$, lo cual contradice nuestra hipótesis. ▮

Si ahora consideramos el anuncio público de la fórmula

$$\neg p \to K_a \neg p,$$

este anuncio divide el modelo en dos: una primera parte donde se encuentra el estado donde p es válido y una segunda parte que no está conectada a dicho estado. Esto es, la actualización del modelo tras el anuncio público de la fórmula anterior quedaría de la siguiente forma:

$$s_0 \xleftrightarrow{a} s_1 \xleftrightarrow{b} s_2 \xleftrightarrow{a} s_3 \longleftrightarrow \cdots \longleftrightarrow s_{n-1} \xleftrightarrow{a} s_n \xleftrightarrow{b} \overset{p}{u}$$

$$t_0 \xleftrightarrow{a} t_1 \xleftrightarrow{b} t_2 \xleftrightarrow{a} t_3 \longleftrightarrow \cdots \longleftrightarrow t_{n-1} \xleftrightarrow{a} t_n$$

En consecuencia, la fórmula $[\neg p \to K_a \neg p]C_{ab}\neg p$ es falsa en $(Horquilla(n), s_0)$, y es cierta en $(Horquilla(n), t_0)$.

Este contraejemplo nos sirve para demostrar el siguiente teorema.

Teorema 4.4. $\mathcal{L}_{KC} \prec \mathcal{L}_{KC[]}$.

Demostración. Tenemos que $\mathcal{L}_{KC} \preceq \mathcal{L}_{KC[]}$, por ser el primero un sublenguaje del segundo.
Para ver que no se verifica $\mathcal{L}_{KC} \equiv \mathcal{L}_{KC[]}$, consideramos la fórmula $[\neg p \to K_a \neg p]C_{ab}\neg p$. Como hemos visto anteriormente, esta fórmula es cierta para $(Horquilla(n), t_0)$ y falsa para $(Horquilla(n), s_0)$. Por el Corolario 4.1, se tiene que no existe dicha fórmula en $\mathcal{L}_{KC}$. Por lo tanto, $\mathcal{L}_{KC} \prec \mathcal{L}_{KC[]}$. ▮

Por tanto, el lenguaje de la lógica con anuncios públicos y conocimiento común es realmente más expresivo que el lenguaje de la lógica epistémica con conocimiento común. El lenguaje de la lógica con anuncios públicos y conocimiento común no puede distinguir, sin embargo, modelos bisimilares. Esto se debe a que los anuncios públicos preservan la bisimilaridad de dos modelos. Luego, si dos modelos eran bisimilares antes de un anuncio público lo seguirán siendo tras él.

Cerramos esta sección mostrando cómo podemos adaptar los juegos spoiler-duplicador vistos en los capítulos anteriores al presente contexto.

Definición 4.13 (El juego $\mathcal{L}_{KC[]}$). *Dados dos modelos $M = \langle S, R, V\rangle$ y $M' = \langle S', R', V'\rangle$ y dados dos estados $s \in S$ y $s' \in S'$, la n-ésima ronda del juego $\mathcal{L}_{KC[]}$ entre spoiler y duplicador sobre (M, s) y (M', s') es la siguiente: si $n = 0$ spoiler gana si s y s' difieren en sus propiedades atómicas para P; si no, duplicador gana. Si $n \neq 0$, puede ocurrir uno de los siguientes movimientos:*

- *Spoiler elige un agente a y un estado t tal que $(s, t) \in R_a$. Duplicador responderá eligiendo un t' tal que $(s', t') \in R'_a$. La salida de este movimiento será (t, t'). El juego continúa sobre los modelos (M, t) y (M', t').*
- *Spoiler elige un agente a y un estado t' tal que $(s', t') \in R'_a$. Duplicador responderá eligiendo un t tal que $(s, t) \in R_a$. La salida de este movimiento será (t, t'). El juego continúa sobre los modelos (M, t) y (M', t').*
- *Spoiler elige un grupo B y un estado t tal que $(s, t) \in R(B)^*$. Duplicador responderá eligiendo un estado t' tal que $(s', t') \in R'(B)^*$. La salida de este movimiento será (t, t'). Y el juego continuará con los nuevos estados.*
- *Spoiler elige un grupo B y un estado t' tal que $(s', t') \in R'(B)^*$. Duplicador responderá eligiendo un estado t tal que $(s, t) \in R(B)^*$. La salida de este movimiento será (t, t'). Y el juego continuará con los nuevos estados.*
- *Spoiler toma un número $r < n$, y dos conjuntos, $T \subseteq S$ y $T' \subseteq S'$, donde $s \in T$, y $s' \in T'$.*
  En primer lugar, duplicador elige dos estados $t \in (T \cup T')$ y $t' \in (\overline{T} \cup \overline{T'})$. Entonces, spoiler y duplicador juega la r-ésima ronda para estos estados. Si duplicador gana este subjuego, gana la n-ésima ronda del juego principal. En otro caso, el juego continúa sobre los modelos restringidos $(M|_T, s)$ y $(M'|_{T'}, s')$ con $(n - r) - 1$ rondas.

Si un jugador no puede realizar su movimiento, ese jugador pierde. Si los estados de salida son diferentes en sus propiedades atómicas para P, spoiler gana el juego. Si spoiler no ha ganado después de n rondas, duplicador gana el juego.

En esta nueva definición spoiler tiene la capacidad de restringir ambos modelos. Esta restricción la entenderemos como el anuncio de una fórmula que es cierta en los estados que son restringidos. La profundidad de esta fórmula será a lo sumo r, que es el número de rondas que jugarán en el subjuego. Al igual que hicimos anteriormente, demostraremos que duplicador tiene una estrategia ganadora para el juego, si y solo si los modelos satisfacen las mismas fórmulas de profundidad menor o igual al número de rondas del juego. Para probar esto, en primer lugar, definiremos de nuevo la profundidad modal, ampliando el concepto.

Definición 4.14. *La profundidad modal de una fórmula viene dada por la función* $d : \mathcal{L}_{KC[]} \to \mathbb{N}$ *tal que:*

$$\begin{aligned}
d(p) &= 0 \\
d(\neg\varphi) &= d(\varphi) \\
d(\varphi \wedge \psi) &= max\,(d(\varphi), d(\psi)) \\
d(K_a\varphi) &= 1 + d(\varphi) \\
d(C_B\varphi) &= 1 + d(\varphi) \\
d([\varphi]\psi) &= 1 + d(\varphi) + d(\psi)
\end{aligned}$$

Definimos $\mathcal{L}^n_{KC[]}$ para distinguir las fórmulas según su profundidad.

Definición 4.15. *Definimos el lenguaje* $\mathcal{L}^0_{KC[]}$ *como el formado por las fórmulas dadas por la notación BNF:*

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi)$$

A partir de esto, definimos el lenguaje $\mathcal{L}^{n+1}_{KC[]}$ *de forma inductiva. Está formado por las fórmulas dadas por la siguiente notación BNF:*

$$\varphi ::= \psi \mid K_a\psi \mid C_B\varphi \mid [\chi]\xi \mid \neg\varphi \mid (\varphi \wedge \varphi),$$

donde $\psi \in \mathcal{L}^n_K$, *y existen* m *y* k *tales que* $\chi \in \mathcal{L}^m_{KC[]}$, $\xi \in \mathcal{L}^k_{KC[]}$ *y* $m + k = n$.

Necesitamos poder decir que cada lenguaje de cada profundidad expresa un número finito de fórmulas.

Lema 4.2. Consideremos un conjunto finito de átomos P. Para cada n, existe un número finito de fórmulas diferentes, salvo por equivalencia lógica, en $\mathcal{L}^n_{KC[]}$.

Demostración. La prueba es análoga a la realizada para demostrar el Lema 3.1.

Ahora probaremos el teorema deseado:

Teorema 4.5. *Para todo $n \in \mathbb{N}$, todo modelos $M = \langle S, R, V \rangle$ y $M' = \langle S', R', V' \rangle$ y todo conjunto finito de átomos P, se tiene que: duplicador tiene una estrategia ganadora para la n-ésima ronda del juego $\mathcal{L}_{KC[]}$ sobre (M, s) y (M', s') si y solo si se cumple que $(M, s) \equiv_{\mathcal{L}^n_{KC[]}} (M', s')$.*

Demostración. La prueba seguirá la misma estructura que las del mismo teorema para otros lenguajes. En primer lugar, aplicamos inducción sobre n. Para la implicación de izquierda a derecha volvemos a aplicar otra inducción, ahora sobre fórmulas de $\mathcal{L}^n_{KC[]}$. La única diferencia con la prueba del Teorema 3.5 es el caso para fórmulas de la forma $[\varphi]\psi$.
Suponemos, sin perdida de generalidad, que $(M, s) \vDash [\varphi]\psi$. Por un lado, suponemos que φ es falsa en (M, s). Partimos de que duplicador tiene una estratega ganadora para la ronda $(n + 1)$-ésima del juego. Entonces, duplicador debe tener una estrategia ganadora para la ronda n-ésima. Puesto que $d(\varphi) < d([\varphi]\psi)$, aplicamos la hipótesis de inducción, por lo que obtenemos $(M', s') \nvDash \varphi$ y, por tanto, $(M', s') \vDash [\varphi]\psi$. Por otra parte, suponemos que spoiler un $[\varphi]$-movimiento y, en su movimiento, restringe los modelos a los casos donde φ es cierta, de manera que ahora φ es válida en ambos. Además elige $r = d(\varphi)$. Por la hipótesis de inducción, spoiler ganará el subjuego, aprovechando que en los estados dentro de la restricción, φ se satisface y en los estados que quedan fuera de la restricción, φ no se cumple. Pero, como duplicador tiene una estrategia ganadora para la $(n + 1)$-ésima ronda, también la tendrá para la $(n - d(\varphi))$-ésima ronda. Entonces, por la hipótesis de inducción, si $(M|_\varphi, s) \vDash \psi$, se cumple $(M'|_\varphi, s') \vDash \psi$. Por lo tanto, $(M, s) \vDash [\varphi]\psi$.

Para la implicación de derecha a izquierda, partimos de $(M, s) \equiv_{\mathcal{L}^n_{KC[]}} (M', s')$, y tenemos que dar la estrategia ganadora de duplicador. Vamos a darla para el nuevo movimiento introducido en este capítulo. Los demás están probados en capítulos anteriores. Suponemos que spoiler realiza el movimiento de restricción y elige dos conjuntos de estados, T y T', y el número $r < (n + 1)$. Por la hipótesis de inducción, duplicador tiene una estrategia ganadora para el subjuego si y solo si existen $t \in (T \cup T')$ y $t' \in \bar{t} \in (\overline{T} \cup \overline{T'})$ tales verifican las mismas fórmulas de profundidad a lo sumo r. Si esta estrategia es posible, duplicador toma esta estrategia. En otro caso, spoiler ganará el subjuego. Por lo tanto, volvemos al juego en la $(n - r)$-ésima ronda sobre los estados $(M|_T, s)$ y $(M'|_{T'}, s')$. Por la hipótesis de inducción duplicador tiene una estrategia ganadora si y solo si $(M|_T, s) \equiv_{\mathcal{L}^{n-r}_{KC[]}} (M'|_{T'}, s')$. Para probar esto, tomamos un estado $t \in (T \cup T')$. Tenemos que para cada estado $\bar{t} \in (\overline{T} \cup \overline{T'})$, existe una fórmula, $\varphi_{t\bar{t}}$, de profundidad a lo sumo r, la cual es válida en el estado t, pero falsa en el estado $\bar{t}$. Por el Lema 4.2, el conjunto $\{\varphi_{t\bar{t}} \mid \bar{t} \in (\overline{T} \cup \overline{T'})\}$ tiene un número finito de fórmulas que no son equivalentes. Definimos una función f, que a cada clase de equivalencia $[\bar{t}]$ le asigna un elemento de la clase. Con esto definimos la siguiente fórmula:

$$\varphi_t = \bigvee_{\bar{t}\in(\overline{T}\cup\overline{T'})} f([\bar{t}])$$

Esta fórmula tiene profundidad menor o igual que r. Mediante un razonamiento similar definimos el conjunto $\{\varphi_t \mid t \in (T \cup T')\}$, que, al igual que antes, es un conjunto finito, salvo equivalencia, por el lema anterior. Sea la función g, que lleva cada clase de equivalencia, $[t]$, a un elemento de ella. Definimos otra fórmula:

$$\varphi = \bigvee_{t\in(T\cup T')} g([t])$$

La profundidad de esta fórmula también es a lo sumo r. La fórmula φ es cierta para cada estado en $T \cup T'$ y falsa para cada estado en $\overline{T} \cup \overline{T'}$. Buscando la contradicción, suponemos $(M|_T, s) \not\equiv_{\mathcal{L}^{n-r}_{KC[]}} (M'|_{T'}, s')$. Esto implica que existe una fórmula ψ, de profundidad a lo sumo $(n-r)$, tal que $(M|_T, s) \vDash \psi$ y $(M'|_{T'}, s') \not\vDash \psi$. Pero entonces $(M, s) \vDash [\varphi]\psi$ y $(M', s') \not\vDash [\varphi]\psi$. Como, $d([\varphi]\psi) = n + 1$, lo anterior contradice nuestra hipótesis de partida. Por lo tanto, $(M|_T, s) \equiv_{\mathcal{L}^{n-r}_{KC[]}} (M'|_{T'}, s')$, es decir, duplicador tiene una estrategia ganadora para el juego que se inicia, después de que spoiler gane el subjuego, por la hipótesis de inducción. ∎

Bibliografía

[BMS99] A. Baltag, L.S. Moss y S. Solecki. *The logic of public announcements, common knowledge, and private suspicions.* CCWI, Amsterdam, 1999.

[BRV01] P. Blackburn, M. Rijke e Y. Venema. *Modal Logic (Cambridge Tracts in Theoretical Computer Science).* Cambridge: Cambridge University Press., 2001.

[Car42] R. Carnap. *Introduction to Semantics.* Cambridge, MA: Harvard, 1942.

[Car47] R. Carnap. *Meaning and Necessity.* Chicago: University of Chicago Press, 1947.

[DHK93] Hans van Ditmarsch, Wiebe van der Hoek y Barteld Kooi. *Dynamic Epistemic Logic.* Springer, 1993.

[Doe87] H. C. Doets. *Completeness and Definability, applications of the Ehrenfeucht game in second-order and intensional logic.* Ph.D. thesis, University of Amsterdam, 1987.

[Fag+95] R. Fagin y col. *Reasoning about knowledge.* Springer, 1995.

[Fig] Diego Figueira. *Bisimulaciones en Neighbourhood Semantics.* URL: `http://dc.sigedep.exactas.uba.ar/media/academic/grade/thesis/dfigueira.pdf`.

[Gar18] James Garson. *Modal Logic.* Ed. por Edward N. Zalta. 2018. URL: `https://plato.stanford.edu/entries/logic-modal/`.

[HS12] R. Hakli y Negri S. "Does the deduction theorem fail for modal logic?" En: *Synthese* 187 (2012), págs. 849-867. DOI: `https://doi.org/10.1007/s11229-011-9905-9`.

[HB34] D. Hilbert y P. Bernays. *Grundlagen der Mathematik I.* Springer, 1934.

[Hin57] J. Hintikka. "Quantifiers in Deontic Logic". En: *Societas Scientiarum Fennica, Commentationes humanarum litterarum* 23 (1957), pág. 4.

[HC96] G.E. Hughes y M.J. Cresswell. *A New Introduction to Modal Logic.* Routledge, London, 1996.

[Kan57a] S. Kanger. "A Note on Quantification and Modalities". En: *Theoria* 23 (1957), págs. 131-4.

[Kan57b] S. Kanger. *Provability in Logic.* Dissertation, University of Stockholm, 1957.

[KP81] D. Kozen y R. Parikh. "An elementary proof of the completeness of PDL." En: *Theoretical Computer Science* 14 (1981), págs. 113-118.

[Kri59] S. Kripke. "A completeness theorem in modal logic". En: *Journal of Symbolic Logic* 24 (1959), págs. 1-14. DOI: `http://www-cs-faculty.stanford.edu/~uno/abcde.html`.

[Lew12] C. I. Lewis. "Implication and the Algebra of Logic". En: *Mind* 12 (1912), págs. 522-31.

[Lew18] C. I. Lewis. *A Survey of Symbolic Logic.* Berkeley: University of California Press, 1918.

[LL32] C. I. Lewis y C. Langford. *Symbolic Logic.* New York: The Century Company, 1932.

[MH95] J.C.C. Meyer y W. van der Hoek. *Epistemic Logic for AI and Computer Science (Cambridge Tracts in Theoretical Computer Science).* Cambridge: Cambridge University Press., 1995.

[Pla89] J.A. Plaza. "Logic of public communications". En: *Proceeding of the 4th Internacional Symposium on Methodologies for Intelligent Systems* (1989). Ed. por M.L. Emrich y col., págs. 201-216.

[Smo84] C. Smorynski. *Modal logic and self-reference.* In D. Gabbay, y F. Guenthner (Eds.), Handbook of philosophical logic (Vol. II, pp. 441–495)., 1984.

[Waj33] M. Wajsberg. "Ein erweiterter klassenkalkül". En: *Monatshefte für Mathematik* 40 (1933), págs. 113-126.