

# COMPLEX CONCEPT LATTICES FOR SIMULATING HUMAN PREDICTION IN SPORT\*

ARANDA-CORRAL, Gonzalo A, BORREGO-DÍAZ, Joaquín,

GALÁN-PÁEZ Juan

**Abstract** In order to address the study of complex systems, the detection of patterns in their dynamics could play a key role in understanding their evolution. In particular, global patterns are required to detect emergent concepts and trends, some of them of a qualitative nature. Formal concept analysis (FCA) is a theory whose goal is to discover and extract knowledge from qualitative data (organized in concept lattices). In complex environments, such as sport competitions, the large amount of information currently available turns concept lattices into complex networks. The authors analyze how to apply FCA reasoning in order to increase confidence in sports predictions by means of detecting regularities from data through the management of intuitive and natural attributes extracted from publicly available information. The complexity of concept lattices -considered as networks with complex topological structure- is analyzed. It is applied to building a knowledge based system for confidence-based reasoning, which simulates how humans tend to avoid the complexity of concept networks by means of bounded reasoning skills.

**Key words** Bounded rationality, complex networks, formal concept analysis, sport forecasting.

## 1 Introduction

Sports data analysis, and the study of the complex systems (CS) behind sport competitions, have experienced rapid growth in the recent years. The increasing interest in models of prediction both for betting markets and for optimizing player/team performance, can be explained by the amazing growth of business in the field. Much evidence about the predictability in sport betting encourages many projects and research<sup>[1-4]</sup>. This paper concerns human reasoning with basic qualitative data for predicting results. The understanding and simulation of forecasting ability allow estimation of global trends for betting markets and eventually to improve betting among the general public. Sport fans make predictions by reasoning with features of qualitative nature (or those that they make qualitative by estimating thresholds).

---

ARANDA-CORRAL Gonzalo A

*Universidad de Huelva. Department of Information Technology; Crta. Palos de La Frontera s/n. 21819 Palos de La Frontera. Spain. Email: gonzalo.aranda@dti.uhu.es.*

BORREGO-DÍAZ Joaquín · GALÁN-PÁEZ Juan

*Universidad de Sevilla. Department of Computer Science and Artificial Intelligence; Avda. Reina Mercedes s/n. 41012 Sevilla. Spain. Email: jborrego@us.es; juangalan@us.es.*

This work is supported by the TIN2009-09492 project of the Spanish Ministry of Science and Innovation, and Excellence project TIC-6064 of Junta de Andalucía co-financed by FEDER funds.

Sport forecasting represents one of the most increasing research topics in this field, because of its economic impact in betting markets as well as for its potential application to problems with similar behavior (other markets)<sup>[1]</sup>. Because of its economic impact, a considerable number of predictive models have been developed by companies and sport franchises (cf. Chapter 10 in [3]). Particularly interesting is the estimation of betting distribution in crowd betting markets and services (for example, the Spanish national soccer betting service Quiniela, <http://www.loteriasyapuestas.es/quiniela>). In this case, the aim is to simulate and detect betting trends based on users' reasoning about the short-term evolution of the competition system. Like other CS, the competition in sport is an open system influenced by an astonishing amount of factors and subsystems. However, humans usually make a quick selection of a few (subjective) features from a limited amount of information (often sport magazines and TV sport news) as well as from their own experience. Therefore, for predicting popular bets it's more precise to adopt rational strategies for reasoning instead of pure statistical methods. Two basic tasks are involved in the process: attribute selection and how to apply them to the reasoning. Actually, the second task attempts to find relationships between the attributes and the results. Formal concept analysis is the theory used in this paper for discovering these relationships.

Formal concept analysis (FCA)<sup>[5]</sup> is a mathematical theory for data analysis using formal contexts and concept lattices as key tools. Domains can be formally modeled according to the extent and the intent of each formal concept. In FCA, the basic data structure is the formal context (with a qualitative nature) which represents a set of objects and their properties and is useful both in detecting and describing regularities and structures of concepts. It also provides a sound formalism for reasoning with such structures, mainly stem basis and association rules. Therefore, it is interesting to consider its application for reasoning with temporal qualitative data in order to find temporal trends<sup>[6]</sup>.

In this paper, the scope of the FCA application is the challenge of sports betting, specifically, the forecasting of soccer league results. Soccer leagues represent a challenging system with a huge amount of knowledge, available through the WWW and several commercial products. Their behavior is exhaustively analyzed every week by journalists, betting companies, and supporters. This analysis can involve an exhaustive study of previous matches<sup>[7-8]</sup> and the performance of players. With the aim of processing the information for predicting the evolution of the system, three dimensions have been considered in existing systems: 1) Those which analyze information on teams (endogenous) versus those which analyze results (exogenous); 2) Those which exploit quantitative data versus those which exploit qualitative knowledge, and finally, 3) Statistic-based ones versus other methods. It is usual to find hybrid models in literature, but rarely pure qualitative and exogenous reasoning systems, although their use is considered for experiments (for example, frugal methods<sup>[9]</sup> and based on recognition heuristic<sup>[10]</sup>) or as part of hybrid systems (see, e.g., [11]). There are two reasons that may justify this point.

On one hand, when transforming a large quantitative dataset to a qualitative problem one is faced with the selection of an acceptable threshold and the discovery of better relationships between the properties (see, e.g., [12]). On the other hand, a qualitative dataset must be achieved with some amount of information based on confidence, trust, or probability of these data sets.

This paper describes the research for selecting and computing attribute sets to be used with FCA in the prediction of soccer match results. This research is previous to building an expert system for advising sport betting which could detect some kind of regularities in data. Concept lattices, which are computed from attribute values, represent a mathematical structure of relationships among the concepts which are involved in selected sport events to study. Since this method is bet-oriented, its performance is evaluated within a confidence-based reasoning system. This system increases the number of hits in forecasting of soccer matches, discovering

trends by means of data mining, and reasoning using association rules. The analysis of attributes was used in [6] to describe a confidence-based (and contextual) reasoning system for predicting sports betting.

Actually the aim is twofold. On one hand, the complexity of the qualitative representation of soccer competition is investigated. On the other hand, we analyze the attribute selection problem as a feature selection problem, which would shape the behavior of the CS that represents professional soccer leagues. The theoretical framework, on which this model is based, is presented in [13]. Due to the huge amount of information, attribute selection advised by experts is mandatory. In fact, the system can be considered a reasoning model based on bounded rationality and recognition heuristics, and focused on features which were considered important by human experts. Therefore, the system aims to forecast results, but it is designed based on bounded rationality models, instead of pure statistical models. However, in the future hybrid models will be considered.

### 1.1 Bounded Reasoning on Complex Systems

As it has already been mentioned, in order to predict betting trends, attributes have to be selected from publicly available information (Web sites, sport journals, etc.). Sport betters extract the information from these sources that they eventually use to argue a prediction. In this way the aim of the system is to simulate (and to improve if possible) better's behavior.

The system must analyze qualitative features of CS, using FCA. The idea is to isolate qualitative attributes from (past) local interactions among components of complex systems and to apply FCA tools in order to predict properties of system behavior in the near future.

Bounded rationality (BR) is intimately related to the human capacity for making inferences under limited time and knowledge<sup>[14]</sup>. From the viewpoint of the artificial intelligence (AI), BR comprises reasoning techniques that facilitate, for example, context and temporal reasoning. Psychological research on specific heuristics in human inference processing reveals a complex framework where classic (computational) logic is not sound for explaining the success of several of them, for example, recognition heuristic (RH)<sup>[10]</sup>. A number of experiments show that cognitive mechanisms capable of successful performance in the real world do not need to satisfy the classical norms of rational inference (cf. [15]). In fact, an intriguing question from ecological rationality analysis is: How could more knowledge be no better- or worse- than significantly less knowledge<sup>[10]</sup>? Several experiments show that bounded rationality is a successful human inference. One of the key features in BR heuristics is that inference process is concentrated on a limited set of experiences in which objects, properties, and actions are selected. The aim is to model this feature with formal concept analysis (FCA).

### 1.2 Structure of the Paper

Section 2 reviews basic features of FCA. In Section 3, the reasoning model for CS based on FCA is described. Section 4 is devoted to analysing the problem of attribute selection to model bounded reasoning for soccer. In Section 5, the scale-free residue of the concept lattice is presented. Section 6 describes the main issues in attribute selection. Section 7 presents the results obtained by the system and how to estimate the reasoning performance from subnetworks of the concept lattice. In Section 8, some conclusions and insights about future work are made.

## 2 Background: Formal Concept Analysis

According to Wille, FCA<sup>[5]</sup> mathematizes the philosophical understanding of a concept as a unit of thoughts composed of two parts: the extent and the intent. The extent covers all objects

belonging to this concept, while the intent comprises all common attributes valid for all the objects under consideration. It also allows the computation of concept hierarchies from data tables. In this section, we succinctly present basic FCA elements (the main reference is [5]).

A formal context  $M = (O, A, I)$  consists of two sets,  $O$  (objects) and  $A$  (attributes) and a relation  $I \subseteq O \times A$ . Finite contexts can be represented by a 1-0-table (identifying  $I$  with a boolean function on  $O \times A$ ). See Figure 1 for an example of formal context about living beings.

The FCA main goal is the computation of the concept lattice associated with the context. Given  $X \subseteq O$  and  $Y \subseteq A$  it defines

$$X' := \{a \in A \mid oIa \text{ for all } o \in X\} \text{ and } Y' := \{o \in O \mid oIa \text{ for all } a \in Y\}.$$

A (formal) concept is a pair  $(X, Y)$  such that  $X' = Y$  and  $Y' = X$ . For example, concepts from formal context about living beings are depicted in Figure 1 (left), right. Each node is a concept, and its intension (or extension) can be formed by the set of attributes (or objects) included along the path to the top (or bottom), e.g., the node tagged with the attribute Legs represents the concept  $(\{\text{Legs, Mobility, Need Water}\}, \{\text{Cat, Frog}\})$ .

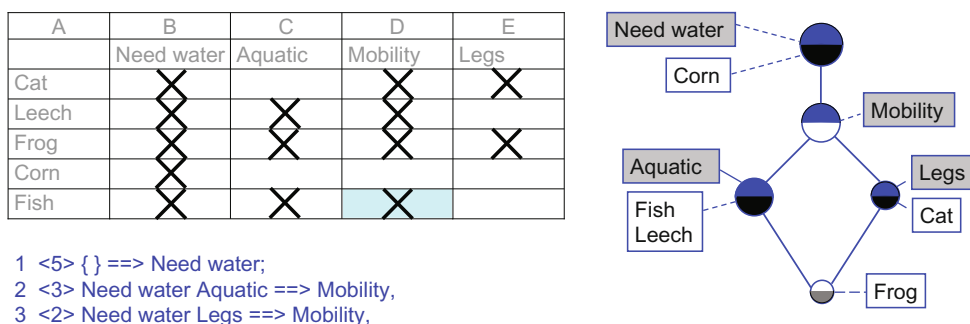


Figure 1 Formal context, associated concept lattice, and stem basis

## 2.2 Implication and Association Rules Between Attributes

Logical expressions in FCA are implications between attributes. An implication is a pair of sets of attributes, written as  $Y_1 \rightarrow Y_2$ , which is true with respect to  $M = (O, A, I)$  according to the following definition. A subset  $T \subseteq A$  respects  $Y_1 \rightarrow Y_2$  if  $Y_1 \not\subseteq T$  or  $Y_2 \subseteq T$ . It says that  $Y_1 \rightarrow Y_2$  holds in  $M$  ( $M \models Y_1 \rightarrow Y_2$ ) if for all  $o \in O$ , the set  $\{o\}'$  respects  $Y_1 \rightarrow Y_2$ . In that case, it is said that  $Y_1 \rightarrow Y_2$  is an implication of  $M$ .

**Definition 1** Let  $\mathcal{L}$  be a set of implications and  $L$  be an implication.

- 1)  $L$  follows from  $\mathcal{L}$  ( $\mathcal{L} \models L$ ) if each subset of  $A$  respecting  $\mathcal{L}$  also respects  $L$ .
- 2)  $\mathcal{L}$  is complete if every implication of the context follows from  $\mathcal{L}$ .
- 3)  $\mathcal{L}$  is non-redundant if for each  $L \in \mathcal{L}$ ,  $\mathcal{L} \setminus \{L\} \not\models L$ .
- 4)  $\mathcal{L}$  is a (implication) basis for  $M$  if  $\mathcal{L}$  is complete and non-redundant.

It can obtain a basis from the pseudo-intents called stem basis (SB)<sup>[16]</sup>. An SB for the formal context on living beings is provided in Figure 1 (down). It is important to remark that SB is only an example of a basis for a formal context. In this paper no specific property of the SB is used, so it can be replaced by any implication basis.

In order to work with formal contexts, stem basis, and association rules, the ConExp software (<http://sourceforge.net/projects/conexp/>) has been selected. It is used as a library to build the module which provides the implications (and association rules) to the reasoning module of our

system. The reasoning module is a production system (designed for [17]). Initially, it works with SB, and entailment is based on the following result.

**Theorem 1** *Let  $\mathcal{S}$  be a basis for  $M$  and  $\{A_1, A_2, \dots, A_n\} \cup Y \subseteq A$ . The following conditions are equivalent:*

- 1)  $\mathcal{S} \cup \{A_1, A_2, \dots, A_n\} \vdash_p Y$  ( $\vdash_p$  is the entailment with the production system).
- 2)  $M \models \{A_1, A_2, \dots, A_n\} \rightarrow Y$ .

Stem basis is an adequate knowledge base (KB) for the production system to reason about attributes and concepts. However, stem basis is designed for entailing true implications only, without any exceptions to the object set nor implications with a low number of counterexamples in the context.

Another more important question arises when working on predictions. In this case the goal is to obtain methods for selecting a result among all entailed conclusions (eventually they are mutually incoherent), and Theorem 1 does not provide such a method. Therefore, it is better to consider association rules (with confidence) instead of true implications and the initial production system must be revised for working with confidence.

Researching on logical reasoning methods for association rules is a relatively recent promising line of research<sup>[18]</sup>. In FCA, association rules are implications between sets of attributes. Confidence and support are defined as usual. Recall that the support of  $X$ ,  $\text{supp}(X)$  of a set of attributes  $X$  is defined as the proportion of objects which satisfy every attribute of  $X$ , and the confidence of an association rule is  $\text{conf}(X \rightarrow Y) = \text{supp}(X \cup Y) / \text{supp}(X)$ . Confidence can be interpreted as an estimate of the probability  $P(Y|X)$ , the probability of an object satisfying every attribute of  $Y$  under the condition that it also satisfies every one of  $X$ . Conexp software provides association rules (and their confidence) for formal contexts.

### 3 Bounded Reasoning on Complex Networks

Our general approach to qualitative reasoning on CS is based on considering local interactions as objects, which have several features (attributes). Therefore, using FCA, a complex network representing some features of the CS is obtained as a concept lattice. Complex networks are a widely used representation of selected features from a CS. The topological structure of the network aids in understanding a considerable number of characteristics of the associated CS. When the goal is to reason with qualitative features, it can be interesting to extract emergent concepts from these interactions. It's here where FCA can play a relevant role.

The selection of FCA as basic mathematic machinery for processing the qualitative information about complex systems lies in the fact that human reasoning (BR) about the dynamics and organization of a complex system has a qualitative nature. Therefore, human reasoning and conjectures about the CS can be expressed in qualitative terms (possibly choosing thresholds and multivalued attributes). Once qualitative hypothesis are presented, non-symbolic mechanisms and reasoning are useful to validate the conjectures. The qualitative reasoning process, in FCA would be depicted as in Figure 2. Once the observer has a collection of relevant attributes to study on the system, he can consider local interactions or nodes as objects of a formal context. This context (often a huge formal context) is built by means of data extraction, database processing, expert observations, data mining, etc. The observer has to select attributes and objects he considers relevant to determine CS dynamics, and the reasoning focuses on the associated subcontext (contextual selection). It is expected that reasoning with the contextual selection gives some information about the CS.

The complex nature of the contexts associated to CS suggests studying its topological structure when there is not a suitable BR method to achieve the aim. When experience and knowl-

edge about the CS suggests a reasoning method, an FCA-based reasoning task is defined and results are obtained. In [6], this general approach was applied using argumentative reasoning on contextual selections. The structure of contextual selection showed some insights on attribute selection<sup>[13]</sup>.

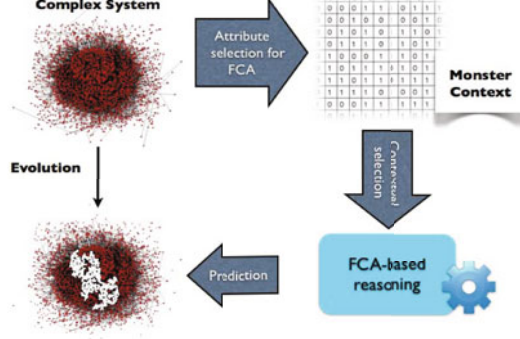


Figure 2 FCA-based model for qualitative reasoning with complex systems

### 3.1 Reasoning Under Contextual Selection: Logical Foundations

The model (described in [13]) is composed of events (objects) which have a number of properties (attributes). They constitute a universal formal context  $\mathbb{M}$  (which we call monster context following the tradition in model theory). Thus,  $\mathbb{M}$  can be considered as the global memory from which subcontexts are extracted. Once the specific context is considered, it is also possible to consider background knowledge  $\Delta$  (in the form of propositional logic formulas) which would be combined with the knowledge extracted from the formal context (stem basis or association rules).

**Definition 2** Let  $\mathbb{M} = (\mathbb{O}, \mathbb{A}, \mathbb{I})$  be the monster context, and let  $O \subseteq \mathbb{O}$ .

- 1) A context on  $O$  is a context  $M = (O_1, A, I)$  where  $O \subseteq O_1 \subseteq \mathbb{O}$ .
- 2) A contextual selection on  $O$  and  $M$  is a map  $s : O \rightarrow \mathcal{P}(O_1) \times \mathcal{P}(A)$ .
- 3) A contextual KB for an object  $o \in O$  w.r.t. a selection  $s$  with confidence  $\gamma$  is a subset of association rules with confidence greater or equal to  $\gamma$  of the formal context associated with  $s(o) = (s_1(o), s_2(o))$ , that is, to the context  $M(s(o)) := (s_1(o), s_2(o), I_{\upharpoonright_{s_1(o) \times s_2(o)}}$ ) (note that when confidence is 1 the contextual KB is a implicational basis).

Contextual KB is useful for entailing attributes on an object. The reasoning model is based on arguments extracted from subcontexts<sup>[13]</sup>.

**Definition 3** Let  $L$  be an implication and  $\Delta$  a background knowledge. It is said that  $L$  is a possible consequence of  $\mathbb{M}$  under  $\Delta$ ,  $\mathbb{M} \models_{\exists}^{\Delta} L$ , if there exists  $M$  a non-empty subcontext of  $\mathbb{M}$  such that  $M \models \Delta \cup \{L\}$ .

Note that by Theorem 1, when  $\Delta$  is a set of implications, it holds that  $\models_{\exists}$  is equivalent to the relation  $\vdash_{\exists}$  defined by:  $\mathbb{M} \vdash_{\exists} L$  if there exists  $M \models \Delta$  a non-empty subcontext of  $\mathbb{M}$  such that  $S \vdash_p L$  (where  $S$  is a stem basis for  $M$ ).

Particularly interesting is the case of inferring properties about future events when the monster model represents attributes of past events. The inference process for  $\vdash_{\exists}^{\Delta}$  consists of the three steps shown below (see Figure 3).

- 1) A question on whether a new event (object) has a property (attribute) is raised. Some properties on the new object are known (attribute values)  $\{A_1, A_2, \dots, A_n\}$ .
- 2) The contextual selection outputs a subcontext of  $\mathbb{M}$ . A contextual KB  $\mathcal{L}$  (for some confidence threshold) is computed for the subcontext. Selection provides a relatively small set

of attributes, selected from own experience and beliefs.

3) The production system is executed on  $\mathcal{L} \cup \{A_1, A_2, \dots, A_n\}$ . The results obtained are the attributes inferred about the new object.

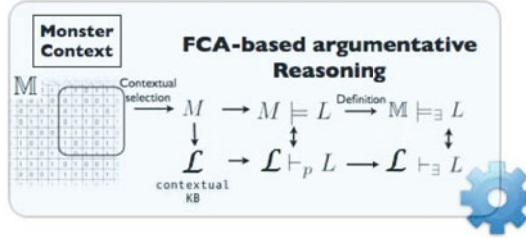


Figure 3 FCA-based reasoning based on  $\vdash_{\exists}$

If  $A'$  is inferred by the production system, then  $\mathbb{M} \models_{\exists}^{\Delta} \{A_1, A_2, \dots, A_n\} \rightarrow \{A'\}$ .

Note that it does not compute all valid implications, only those entailed from the set of attributes selected by the user. Therefore, we need to understand the topology of the lattice to properly choose the attributes to reason.

Attributes are essential components in the contextual selection. Association rules are extracted from the contexts and used by the production system. From these association rules and some initial facts, based on the match we want to forecast, the production system infers the confidence (probability) for each one of the three possible results of the match, home win, draw or away win. Thus attributes constitute one of the most important and sensitive parts of the system. They are sensitive because the accuracy of the inferred results will depend on how they represent the behavior of the teams. Lastly, attribute selection allows isolating a piece of the monster context where argument reasoning is based.

### 3.2 Confidence-Based Reasoning System

The system works on facts  $(a, c)$ , where  $a$  is an attribute and  $c$  is the estimated probability of the trueness of  $a$ , which we also call confidence (by similarity with the same term for association rules). See [6] for a more detailed description of the system. It has a module for a confidence-based reasoning system (Figure 4).

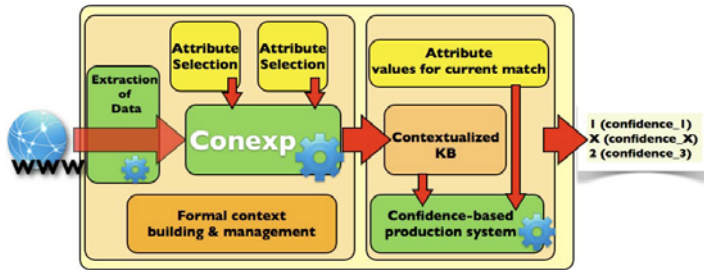


Figure 4 FCA-based reasoning system

Its entries for a match Team<sub>1</sub>-Team<sub>2</sub> are: the contextual KB for a threshold, given as a rule set, and attribute values for the current match (except 1, X, 2) as facts, all of them with a confidence (see below). After the execution of the production system, the output is a triple  $\langle (1, c_1), (X, c_x), (2, c_2) \rangle$  of (attribute, confidence) for the match. The attribute with greater confidence is selected as the prediction. Production system execution is standard, with several

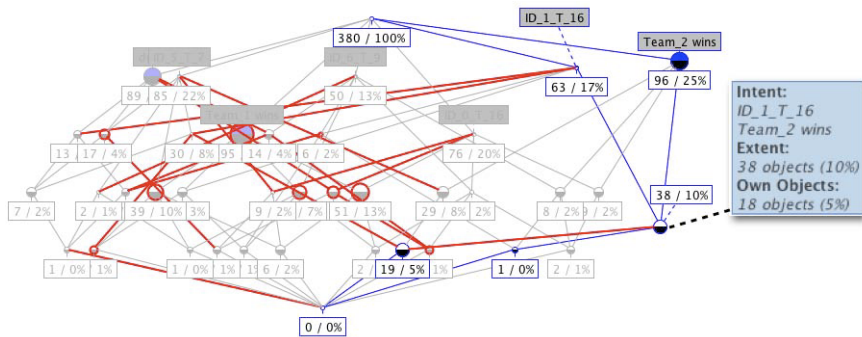
modes for confidence computing based in uncertain reasoning in expert systems<sup>[19]</sup>. Any fact (attribute)  $a$  is initialized with confidence

$$\text{conf}(a) := \frac{|\{o : oIa\}| + 1}{|O| + 1}.$$

## 4 Attributes and Formal Contexts for Soccer Leagues

For both selecting data and building contexts, some assumptions on forecasting in soccer league matches have been considered. Reconsiderations of such decisions can be easily computed in the system. First, we consider that the regularity of a team’s behavior only depends on the contextual selection that has been considered. This contextual selection is obtained by taking matches from the last  $X$  weeks backwards, starting from the week just before the one we want to forecast. Second, FCA methods are used to discover regularity features, thus it does not consider forecasting exceptions (unexpected results). Therefore, the model can be considered a starting point for a betting expert who would adjust attributes, in order to obtain more personalized criteria.

These attributes have to be computed and used to entail a forecasting. In order to select the most interesting attributes for the system, starting from an initial configuration, the user can compute the associated concept lattice and check it. In this way, attributes’ goodness (and thresholds) can be evaluated to reconsider attribute selection. For example, in Figure 5, the concept lattice associated to a contextual selection for Málaga-Sevilla match is shown. This contextual selection is based on a given attribute selection and the matches of the last 38 weeks, just before the queried match date. In this concept lattice, the attribute  $ID\_1\_T\_16$  is defined by: ‘the budget of team<sub>2</sub> is greater than  $\gamma_1$  times the budget of team<sub>1</sub>’, where  $\gamma_1$  is the threshold the expert must estimate. In the concept lattice we can observe that the biggest concept containing the attributes  $team_2\_wins$  and  $ID\_1\_T\_16$  covers about the 10% of the objects owned by the first attribute, therefore it is suggested to use the second attribute for reasoning with association rules to get a prediction.



**Figure 5** Concept lattice for the match Málaga-Sevilla (week 31, season 2009-10)

The system computes the value of an amount of attributes of objects. Experimentally a boolean combination of attributes is possible. Once the temporal context has been computed, the system can build contextual selections by selecting the match and the attribute set. The selection of attributes was made by considering diverse kinds of factors (as for example the difference between team budgets). Seventeen relevant attributes were selected. The attribute set has three special attributes, Team<sub>1</sub> wins (1), Team<sub>2</sub> wins (2), and draws ( $X$ ).



Data are automatically extracted from The RSSSF Archive (<http://www.rsssf.com>). Objects are matches and attribute set is a list of features, including timestamp (week, year). Data was collected for the past twelve years.

#### 4.1 Attribute Selection

As it was remarked before, the attribute selection is based on users's attribute selection, which have to be computed from publically available data (provided by the monster context). Discussion on the soundness of each selection is a human process based on experience. In this section a representative set of attributes is presented. Criteria and expected correlation with match results are also discussed. There are many approaches and studies about the problem of selecting relevant features for soccer matches and competitions (cf. [7, 8]), but in our case we need those associated with bounded rationality. Therefore, attributes have to be computed from publicly available data and well-known relevant features of teams.

**Criteria for selecting attributes.** A set of attributes with many possibilities through customizable parameters was selected. When the parameters are set up with proper values, the attribute set will represent the team's logical behavior.

Recall that formal concept analysis works with qualitative attributes and all teams's information we work with have quantitative nature. Thus it is necessary to convert quantitative attributes into qualitative ones. This task is left to users by choosing a proper threshold for each attribute.

Before choosing the base set of attributes, we have carried out an analysis of information about soccer results. The aim has been to discover which factors are more influential in teams' behavior and which ones are less influential. First of all, we have collected every interesting factor found, and after analyzing each one, individually, we have chosen the most suitable ones. Examined factors can be classified in four different categories (Table 1): those related to league ranking, those related to previous teams's results, those related to historical direct matches, and any other factors. It is worthy to note that to increase possibilities of the attribute set, and considering the boolean nature of formal context attributes, we have added the option to create new ones by means of logical combinations of these attributes.

According to those factors, it considers a base set of 18 attributes, customizable by some parameters. This will let us obtain a diverse set of attributes (Table 2).

1) Threshold: Parameter to be used to translate quantitative attribute values into qualitative ones.

2) Team: Recall that in the formal context considered, objects are matches but attributes refer to team properties. This parameter will set the team from the object (match) by which the attribute will be considered. It has two possible values: {HOME, AWAY}. Thus, usually, we will have each attribute in a context twice, once for the home team and once for the away one.

3) Number of matches: number of past matches to be considered when some attributes are computed, e.g., the ones associated with previous team results.

4) Kind of matches: sets the type of past matches to be taken into account, when computing some attributes, considering home/away team's condition at matches. It has three possible values: {MATCHES AS HOME TEAM, MATCHES AS AWAY TEAM, ALL MATCHES}.

With these parameters, and the possibility of creating compound attributes, it is possible to build a set of detailed attributes. It is worthy to note that experiments showed that the simplest and most logical attributes give a good representation of team behavior. However we consider that a versatile attributes set, as the above described, was necessary because of the huge number of factors that can determine the result of a soccer match. The task of customizing

the attribute set is left to users. It is the most important one in the forecasting process, so a basic soccer knowledge should be required. Its goodness will determine system results.

**Table 1** Factors considered for selecting/building attributes

Factors	Correlation Degree	Used?
Associated to the league ranking		
Team in the first classification level	medium/high	yes
Team in the last classification level	medium/high	yes
Difference between team's classifications	medium/high	yes
Team was in a different league last year	medium	no
Team scored an important number of goals (in the last matches)	medium/low	no
Associated to previous results of the team		
Number of consecutive won matches	high	yes
Number of consecutive lost matches	high	yes
Number of consecutive draws	medium	yes
Number of non-consecutive won matches in previous weeks	high	yes
Number of non-consecutive lost matches in previous weeks	high	yes
Number of non-consecutive draws in previous weeks	medium/high	yes
Points collected in previous weeks	medium/high	yes
Related with directed matches (included previous seasons)		
Number of wins in previous directed matches	medium/high	yes
Number of losts in previous directed matches	medium/high	yes
Number of draws in previous directed matches	medium/high	yes
Other Factors		
Number of red cards collected by the team's players	low	no
Weather the day and the city where the match took place	medium	no (hard to parametrize)
Motivation because of fans support when playing as home team	high	no (hard to parametrize, subjective)
Team hires a new coach	high	no (only useful when new coach hired)
Some players of the team are selected for their National Team	medium/low	no (relevant for some nationalities)
Difference between team's budgets	high	yes
One or more important team's players are injured	medium	no (hard to automatically collect the data)
Cups collected in the lasts years	low	no (only for a few of teams)

## 5 Concept Lattices as Complex Networks: Scale-Free Residue

Given the attribute set of objects of a CS, the concept lattice associated to the CS (CLCS) is the concept lattice built from the monster context associated with the CS. Note that the CLCS can be considered as a directed graph (as Hasse diagram indicates) or as a nondirected graph if necessary. Thus, each node (concept) connectivity can be studied as degree (nondirected), indegree, or outdegree. The analysis of CLCS can show interesting concepts for better

understanding the structure and dynamics of CS. It is also useful to consider the role some attributes play in the qualitative study of the CS. However, a CLCS associated with the monster context is a complex network of semantic relationships which is not bounded by the self language, as in other semantic networks<sup>[20]</sup>, because concepts do not have to be represented by a single language term nor a intelligible definition by the observer. It is expected to face complex networks with extreme structural topology, therefore, prior to applying BR techniques it is advisable to study their topology.

**Table 2** Attributes and parameters

Attribute	Configurable parameters
1) Number of non-consecutive won matches in previous weeks > threshold	<Threshold> <Team> <Number of Matches> <Matches>
2) Number of non-consecutive lost matches in previous weeks > threshold	<Threshold> <Team> <Number of Matches> <Matches>
3) Number of non-consecutive draws in previous weeks > threshold	<Threshold> <Team> <Number of Matches> <Matches>
4) Points collected in previous matches > threshold	<Threshold> <Team> <Number of Matches> <Matches>
5) Position in the classification based on previous matches > threshold	<Threshold> <Team> <Number of Matches> <Matches>
6) Number of positions over the opponent in the classification based on previous matches > threshold	<Threshold> <Team> <Number of Matches> <Matches>
7) Number of positions under the opponent in the classification based on previous matches > threshold	<Threshold> <Team> <Number of Matches> <Matches>
8) Number of wins in previous directed matches (included previous leagues) > threshold	<Threshold> <Team> <Number of Matches> <Matches>
9) Number of losts in previous directed matches (included previous leagues) > threshold	<Threshold> <Team> <Number of Matches> <Matches>
10) Number of draws in previous directed matches (included previous leagues) > threshold	<Threshold> <Number of Matches> <Matches>
11) Position in the classification > threshold	<Threshold><Team><Matches>
12) Number of positions over the opponent in the classification > threshold	<Threshold><Team><Matches>
13) Number of positions under the opponent in the classification > threshold	<Threshold><Team><Matches>
14) Number of consecutive won matches > threshold	<Threshold><Team><Matches>
15) Number of consecutive lost matches > threshold	<Threshold><Team><Matches>
16) Number of consecutive draws > threshold	<Threshold><Team><Matches>
17) Team's budget Y times bigger than opponent's budget (Y > threshold)	<Threshold> <Team>
18) Team's budget Y times smaller than opponent's budget (Y > threshold)	<Threshold> <Team>

A scale-free network is one whose degree distribution follows a power law, at least asymptotically. That is, the fraction  $P(k)$  of nodes in the network having  $k$  connections to other nodes goes for large values of  $k$  as  $P(k) \sim ck^{-\gamma}$ , where  $c$  is a normalization constant and  $\gamma$  is a parameter whose value is typically in the range  $2 < \gamma < 3$ , although occasionally it may lie outside these bounds (as we will see below). The asymptotic behavior means that, in practice, few empirical phenomena obey the power law distribution for all the values<sup>[21]</sup>. It is more common to have this behavior from a certain threshold  $x_{\min}$ . The scale-free residue of a CLCS is the set of its nodes whose degree is greater than  $x_{\min}$ .

CLCS are concept networks based on the attribute selection of observers, who choose what attributes they think are more relevant to reason with the system. The complexity of such

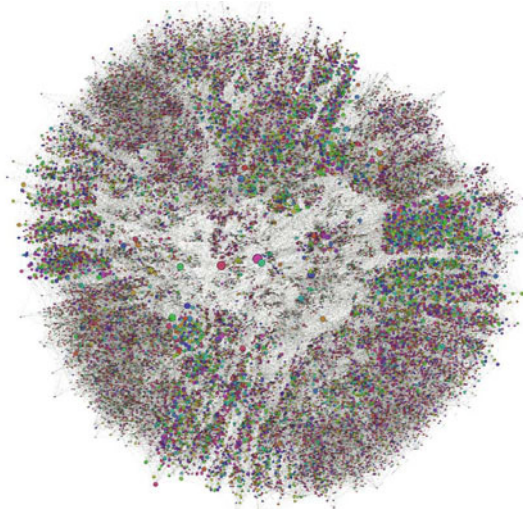
networks lies in the fact that the combinatorial nature of FCA covers every formal concept. We expect that the topological analysis of the dynamics of CLCS shows a big picture of the complex system itself. Normally one will expect that the CLCS has a topology similar to other concept lattices, even similar to concept lattices associated to random formal contexts (with similar density). However, as it is discussed in the following examples, the degree of nodes is not usually very large although there are many attributes in play. As a consequence of this, the architecture of the lattice is very complex and exhibits a very different topological structure than the lattices extracted from random contexts.

However, it is possible to refine the choice of  $x_{\min}^{[21]}$  (See Subsection 5.1 below). In our case we use  $x_{\min}$  as the degree value with the maximum frequency within the CLCS (the maximum of the degree distribution function). Lastly, CLCS are not random networks. Recall that a random network is a network in which each node is randomly connected with other nodes of the network with a certain (fixed) probability. It is known that the degree distributions of these networks follow a Poisson distribution<sup>[22]</sup>.

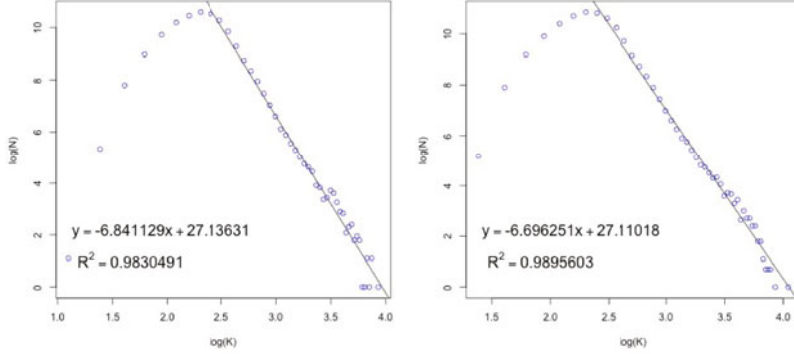
The monster context covers all matches from season 2005/06 to season 2010/11 of Spanish Soccer Premier League. Attributes express properties of the teams involved in the match, and they are booleanized if it were necessary (by taking thresholds, see below for details). Table 3 shows the main parameters of the cumulative CLCS (density is the proportion of  $|\mathbb{I}|$  with respect to  $|\mathbb{O} \times \mathbb{A}|$ ,  $|\text{CL}|$  is the size of the CLCS,  $sc$  is the proportion of concepts in the scale-free residue,  $|\text{SB}|$  is the size of stem basis and  $|\text{CF} > 0.5|$  is the number of association rules with confidence greater than 0.5, and  $\gamma$  is the exponent of the power-law distribution). In Figure 6 the CLCS for the first season only is depicted. Figure 7 shows the log-log chart representing the scale-free residue behavior.

**Table 3** Data on dynamics of both monster contexts and associated concept lattices

Seasons	$ \mathbb{O} $	$ \mathbb{A} $	Density	$ \text{CL} $	avg. degree	$x_{\min}$	$sc$	$ \text{SB} $	$ \text{CF} > 0.5 $	$\gamma$
2005-06	842	94	13.312%	27,434	8.273	7	77.619%	5,772	32,710	5.334
2005-07	1,684	94	13.306%	81,490	9.472	9	60.865%	10,062	90,864	6.084
2005-08	2,526	94	13.267%	140,739	9.976	9	68.199%	14,235	154,011	6.348
2005-09	3,368	94	13.375%	243,959	10.622	10	63.596%	17,840	260,716	6.841
2005-10	4,210	94	13.368%	324,146	10.822	10	66.203%	22,088	344,986	6.696



**Figure 6** Concept lattice associated to the monster context (season 2005-06)



**Figure 7** Degree distribution of the Scale-free residue of CLCS for 4 and 5 seasons, respectively)

It is worthy to note the growth and size of  $\gamma$  in the distributions shown in Table 3. One explanation for this behavior is based on [22], where the asymptotic behavior of scale-free networks is studied. In the work mentioned, two models of limit behavior are shown. In one of them, the number of nodes in the network remains constant, and the distribution tends to a Gaussian one. In our case, if it is supposed that the process starts from a node set with  $2^{94}$  nodes (the power set of the attribute set), and in each step new objects are added, new relations in the lattice emerge. From this point of view, the CLCS show a distribution which is in the middle of the pure scale-free and the Gaussian one<sup>[22]</sup>.

The influence of scale-free distribution on the reasoning system is twofold. Firstly, the target attributes are allocated, in this case, in the tail of the distribution, that is, a great number of attribute sets (concepts intent) influence in such target attributes, hence a direct discard of attributes does not discriminate at first one target attribute from the rest (because the graph remains connected, and the application of association rules by the reasoning system can output any other target). Secondly, since the diameter of the graph is small (for example, the diameter of the graph associated to the monster context for 2005-06 season is 14 and the expected length path is 4.9) the number of rule firings to observe the first phenomenon can be relatively low. Both issues involve a significant number of concepts in the scale-free residue, about the 60% of the concepts.

Therefore, scale-free distributions prevent us to use only graph-based algorithms (for example, deleting concepts) to estimate the result of a match. Furthermore, the evolution of the cumulative context (Table 3) shows that the topological structure is preserved; it is neither specific of a season nor of a set of seasons.

### 5.1 Goodness-of-Fit Test

The methodology presented in [21] states that the usual/common methods for obtaining  $\gamma$  and  $x_{\min}$ , and for testing the power-law distribution in a sample can be faulty in some cases. Therefore, they propose a new methodology, which has been applied to the node's degree distribution of CLCS in order to test whether they really follow a power-law distribution or not.

The process requires a previous step in which optimal values for  $\gamma$  and  $x_{\min}$  are estimated. Then the goodness-of-fit test (see [21] for details) is performed and the  $p$ -value (statistical significance) is obtained. If  $p$ -value  $< 0.1$  then the power-law distribution hypothesis is ruled out, otherwise is accepted. To obtain a precise  $p$ -value, the number of randomly generated samples per test considered is 10,000.

The results of the test showed that the power-law distribution hypothesis could be accepted. (For example, for monster context of seasons 2005-06 with  $x_{\min} = 11$ ,  $\gamma = 5.71$ , the test returned  $p$ -value=0.2855 and for the one of seasons 2005-08 with  $x_{\min} = 15$ ,  $\gamma = 6.53$ , the test returned  $p$ -value=0.2451).

Finally this methodology also presents tests to discard other possible distributions. Those were applied in order to discard them in the same way it is done for many samples in [21].

## 6 Computing Issues Relative to Qualitative Reasoning

The nature of the competition makes it necessary to take into account some special situations when computing attribute values. In this section we describe the main problems that emerged and how they were fixed. Roughly speaking, these main problems concern the initial matches of the season.

**Beginning of a new season: week 0.** This problem is not difficult but, as many others, unavoidable, and a solution becomes essential. It happens when computing an attribute value related to the league standings to forecast the first week of a season. As no match has been played in the previous week, there is no way to build a standing table.

When teams in the current season remain in the same division as in the last one, a trivial solution is to take into account positions and matches in the last weeks of previous season. If the team played in a higher division than in the last season, it will be at the first positions in the current standing. Otherwise, if the team played in a lower division, it will be considered in the last positions.

**Missing matches in attribute computation.** Another problem, closely related to previous one, is when not enough previous matches are available to compute an attribute. The solution consists in taking the last matches of the previous season as if they were in a continuous temporal line. This is not so simple, because some teams were not playing at the same division the previous season. Indeed, when playing in a lower or higher division, the difficulty of the division changes and matches cannot be compared in the same way. Therefore, we need to handle the situation of a team playing in a different division from the one.

Another problematic situation where there are not enough matches for attribute computation is computing results for directed matches between two teams because there are only a few of such matches in the data source.

For these two related situations we offer two solutions. The first one is to compute attributes with a null value, but this way we are giving fake information to the system. We are setting that attribute as not true but, in fact, we don't have enough information to determine it, so a better approach is required. The second option (the chosen solution) is based on adjusting the attribute's threshold. The value of this threshold is decreased proportionally in relation to the number of required matches and the number of available matches. Threshold  $\gamma$  is revised by

$$\gamma_{\text{new}} = \gamma_{\text{old}} \cdot \frac{\text{number of match results available}}{\text{number of match results needed}}.$$

When the number of required matches is too high and the number of available matches is low, it looks like we are giving fake information to the system again, but our experience shows that the collateral effects of this approach are worthless compared to computing attributes with a null value.

## 6.1 Attribute Selection vs Expert System Behavior

In general terms, predictions produced by the current base set of attributes are quite good in regular conditions, that is to say, they detect the logical behavior of teams. Therefore, the attribute set must be specifically customized to detect unexpected results. Even so, some experiments, in order to study attributes' behavior, have been developed.

An attribute is strict when only few objects can satisfy it, because its threshold is too high. By working with sets of strict attributes, we can assure that they estimate the teams' behavior better than other sets. Thus, with strict attributes, we will have very reliable estimates, but only just for very few matches, and not for most others. On the other hand, using less strict attributes, the system will produce less reliable estimations but for a large scope. So, it is essential to find a balance between these two opposite situations: reliability of attribute set against number of matches without information. A good solution could be to build and use different attribute sets, some more strict, and others less so. Thus, less strict attribute sets will be used when strict ones fail at making an estimation

## 6.2 Trends Favoring the Victory of the Home Team

It is a fact that, in soccer, a victory for the home team is more probable than an away team victory. To address this, we offer two different approaches. First, modeling the teams's behavior and second, computing confidence values. For modeling teams behavior (attribute set customization) it is a good practice to use attributes with lower thresholds for the home team and higher thresholds for attributes related to the away team. Therefore, it will be easier for the home team to satisfy an attribute than for the away team.

Around 50% of played matches finish with the victory of home team. This means that the attribute value corresponding to the match result, will be 'home team victory' around 50% of objects from the formal context. As a consequence of this, many rules from the inferred association rules will contain the attribute 'result = home team victory' within their conclusions. Thus when forecasting a match the system will infer, in most cases, 'home team victory' as consequence of overestimation confidence value for this result. It is possible to avoid this effect easily, just applying a decreasing factor over confidence for 'home team victory'. It is estimated by means of experiments.

## 7 Results

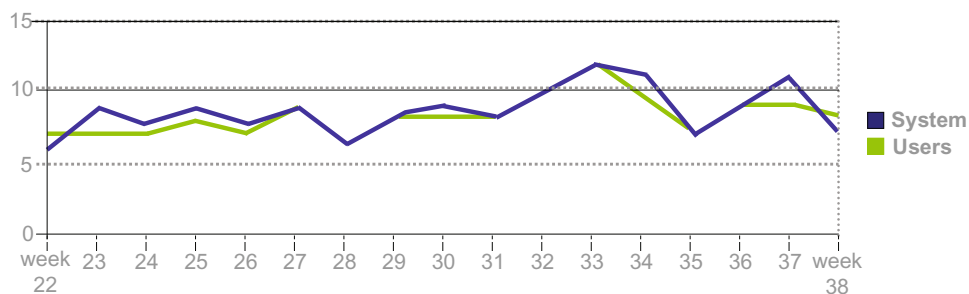
Following the process described above, experiments run for the Spanish Premier Soccer League. Attributes were selected according to an expert, and contextual KB is computed (Figure 8 shows a fragment of the KB for Málaga-Sevilla match). From the selection  $\vdash_{\exists}$  is computed for each match in each week.

```
....
14 < 2 > ID_2_T_0 ID_6_T_9 =[100%]==> < 2 > ID_1_T_16;
15 < 1 > ID_2_T_0 ID_0_T_16 =[100%]==> < 1 > ID_4_T_7 Team1_Wins;
16 < 2 > ID_3_T_2 ID_4_T_7 =[100%]==> < 2 > ID_0_T_16;
17 < 1 > ID_3_T_2 ID_5_T_7 =[100%]==> < 1 > Team1_Wins;
18 < 2 > ID_3_T_2 ID_0_T_16 =[100%]==> < 2 > ID_4_T_7;
19 < 1 > ID_3_T_2 Team2_Wins =[100%]==> < 1 > ID_4_T_7 ID_0_T_16;
20 < 2 > ID_4_T_7 ID_1_T_16 =[100%]==> < 2 > Team2_Wins;
21 < 1 > ID_5_T_7 ID_6_T_9 ID_1_T_16 =[100%]==> < 1 > Drawn;
22 < 1 > ID_5_T_7 ID_6_T_9 Drawn =[100%]==> < 1 > ID_1_T_16;
23 < 10 > ID_2_T_0 Team2_Wins =[90%]==> < 9 > ID_1_T_16;
24 < 8 > ID_2_T_0 ID_5_T_7 =[88%]==> < 7 > ID_1_T_16;
25 < 4 > ID_2_T_0 Drawn =[75%]==> < 3 > ID_1_T_16;
26 < 4 > ID_3_T_2 =[75%]==> < 3 > Team1_Wins;
....
```

Figure 8 KB fragment from Figure 5

**2009-10 season** Experiments with the system show forecasts of about 58.16% by a contextual selection based on the previous 38 matches of each team. Such a percentage of hits for a qualitative reasoning system may be considered as an acceptable result comparable with results conjectured by experts<sup>[9,15]</sup>. Experiments with other contextual selections show an increase in the number of hits by about 7% in the second half of the season. The reason is that data from the first half provides more recent information on teams and past matches.

**2010-11 season** According to the idea mentioned above, we have evaluated the system in the second half of the 2010-11 soccer season. A way to evaluate the effectiveness of this forecasting system is to compare the number of successes in our pool with the most popular betting selections. The popular selections are collected from the most voted results for each match, published on a state agency website that controls soccer pools. In Figure 9, both results are compared. The average number of bets per week from which this popular selection is obtained is 19,064,873. Note that Spanish soccer pools are over 15 matches.



**Figure 9** Correct predictions on the last 17 weeks of the season 2010-11

## 7.1 Analysis of the Reasoning Performance and Comparison with Other Methods

In order to understand the complexity of logical reasoning in a CLCS, it is possible to analyze the topological structure of concept nodes which are involved in frequent deduction tasks. In Figure 10 the deduction graph of the monster model for sport forecasting is depicted. Its nodes are the concepts of CLCS whose intension contains both the left-hand side of at least 10% of SB rules and the right-hand side of at least 10% of the same kind of rules. The graph contains 715 nodes and 1296 arcs. Its diameter is 7 (considered as a directed graph) which suggests a reasonable (upper bound) estimation of rule firings by the production system. The existence of many connected components is useful in deciding which attributes are relevant to achieving a certain result in prediction (by analyzing the attributes which appear in the same connected component as the targeted attributes).

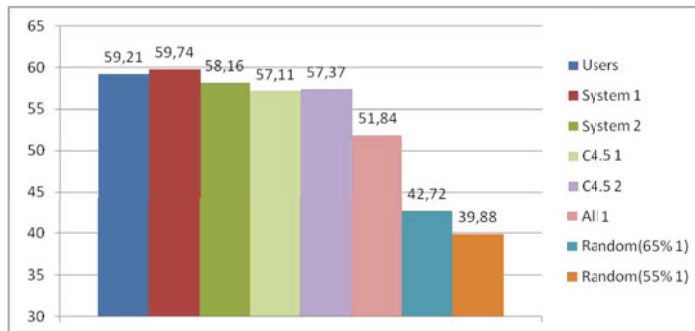
No statistical approaches have been taken into account, because it was not the aim of this paper. However a comparative study of our system against C4.5 classifier has been done. Therefore, two different attribute selections have been considered and used for both C4.5 classifier and our system. The experiment is to forecast all matches (380) in the 2010-11 season. In order to estimate each match result, considering  $N$  (weeks) as a timestamp, previous matches are used to build the contextual selection (or training set in C4.5) from weeks  $N - 1$  to  $N - 19$  (190 objects). Figure 11 shows the percentage of correct predictions for our system and C4.5 classifier, using both attribute selections. Other cols are also shown: ‘user’s most voted results’, ‘the local team always wins’ (All 1) and two that are randomly generated. These ‘randomly generated’ cols were built assuming different weights per result. It means,  $\langle 1 : 55\%, X : 23\%, 2 : 22\% \rangle$  and  $\langle 1 : 65\%, X : 18\%, 2 : 17\% \rangle$  were used, where 1,  $X$ , 2 are the probabilities of forecasting a match with the result: local team wins, draw and away team wins, respectively. The



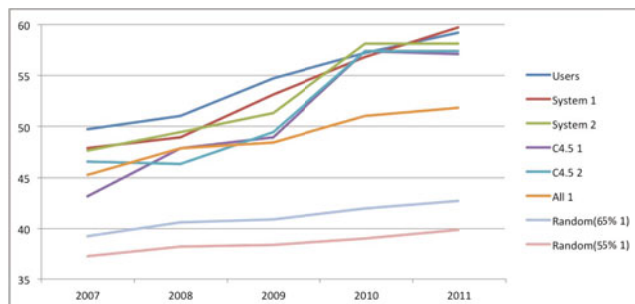
same experiment was repeated for the last five seasons using the same attribute selections and parameters than the former (see Figure 12). It is interesting to remark two features of this experiment: On one hand, the number of correct predictions increases over the years, that is to say, it is more predictable, mostly due to the increase of the team's budget difference in the present years. On the other hand, system predictions fits well with the popular selection in the last year, but worse in former years. System accuracy is worse in previous years due to the fact that the attribute selection was designed regarding current league trends, to be used to forecast results in the near future.



**Figure 10** Deduction graph extracted from CLCS of season 2005-06



**Figure 11** Comparative of correct predictions on the whole season 2010-11



**Figure 12** Comparative of correct predictions on the last five seasons

**Table 4** Statistical significance of non-parametric tests of similarity between correct predictions of the system and popular selection, and hypothesized median for the one sample test

Season	Two paired samples	One sample	Hypothesized median $\theta$ (one sample)
2006-07	0.075	0.545	1.5
2007-08	0.174	0.081	1.5
2008-09	0.344	0.945	1.5
2009-10	0.491	0.107	1
2010-11	0.697	0.943	1

## 7.2 Testing Bounded Rationality Hypothesis

Since the system is based on bounded reasoning, it is expected for both popular selection and system predictions to be similar. The system can be considered as a tool for simulating human behavior while betting in a rational way. In order to justify this assertion, another statistical test was performed to measure the similarity between popular selection and system predictions. Due to the nature of the sample, a non-parametric test should be used, in this way, the Wilcoxon signed rank test has been chosen. The test has been applied in two different modes:

**Wilcoxon signed rank test for two related (paired) samples** Recall this test starts from a null hypothesis stating that the median of the difference between the two paired samples is 0. If the statistical significance of the test is greater than 0.05 (significance level) the null hypothesis can be retained. To apply this test, the paired samples to be considered are the number of correct predictions in each week (from 0 to 10) of both, popular selection and system. It is possible that the number of correct predictions of the system and the popular selection are the same, but having different sets of well predicted matches. Therefore another more exhaustive test should be done.

**Wilcoxon signed rank test for one sample** Recall the test starts from a null hypothesis stating that the median of the sample is equal to a hypothesized value  $\theta$ . If the statistical significance of the test is greater than 0.05 (significance level) the null hypothesis can be retained. To apply this test, the sample to be considered is the symmetric difference between the well predicted matches of the system and the popular selection, for each week. This test is expected to be more exhaustive than the former because it measures the differences individually case by case.

In Table 4, the results of both tests for the last five seasons are shown. It can conclude that the predictions of the popular selection and the system are quite similar. It is worthy to note that the similarity increases from past years to the present, reinforcing the assertions made in Subsection 7.1 (see also Figure 12).

## 8 Conclusions and Future Work

The challenge of detecting emergent concepts to reason on complex systems represents an exciting field of research. Concepts with qualitative nature are extracted from data considering only partial features of complex system dynamics, that is to say, a partial understanding of systems. In this paper, FCA is used with this aim in a specific application. The attribute selection problem using an FCA-based reasoning system for sport forecasting is analyzed. In fact, the reasoning system is a computational logic model for bounded rationality. The model is concerned with association rule reasoning and it does not use -in its current form- more

sophisticated probability tools (as for example, [8, 11, 23]). As it is stated in [24], the theory of probabilistic mental models assumes that inferences about unknown states of the world are based on probability cues<sup>[25]</sup>. It can say that confidence of association rules extracted from subcontexts play the role of probability cues.

It is worth it to note that, while the classifier achieves the highest performance (58, 68%) when the number of matches increases from 190 to 380, our system reaches the highest performance (59, 74%) using only 190 instances. This can be explained by the fact that our system uses a kind of fast and frugal<sup>[24]</sup> method, and these are designed to achieve acceptable results using as few as possible resources. We may assert that our model recognises trends in contexts. Trends (represented as association rules) can be considered a kind of recognizing method. Although the system is based on bounded rationality models instead of statistic models, hybrid models will be considered in future work. It could be interesting to use the system to detect rigged results ('likelihood evidence' of illegal behavior) in soccer. Recently some work about a similar question has been published (see, e.g., [26]).

In the short term, we carry on extending the system to combine the results of two or more different attribute sets. Therefore, the system will return a single, more reliable result. In the long term, we aim to extend the model in order to obtain a general system to detect emergent concepts in complex systems.

Finally, from the point of view of CS research, current and future authors' work on CLCS show that those associated to CS where rational activities are relevant (for example, reasoning and classifying) present similar regularities which do not appear in those without human intervention nor in random contexts. In some way, this phenomenon would be useful to estimate rationality influence in a CS.

## 9 Acknowledgements

The authors would like to thank the anonymous reviewers and Prof. Rafael Blanquero-Bravo from University of Seville for their valuable comments and suggestions to improve the quality of the paper.

## References

- [1] Why Spain will win . . . , *Engineering & Technology*, 5 June–18 June 2010.
- [2] Andersson P, Memmert D, and Popowicz E, Forecasting outcomes of the World Cup 2006 in football: Performance and confidence of bettors and laypeople, *Psychology of Sport & Exercise*, 2009, **10**(1): 116–123.
- [3] Schumaker R P, Solieman O K, and Chen H, Sports data mining, *Integrated Series in Information Systems*, 2010, **26**: 55–63.
- [4] Vlastakis N, Dotsis G, and Markellos R, How efficient is the European football betting market? Evidence from arbitrage and trading strategies, *Journal of Forecasting*, 2009, **28**(5): 426–444.
- [5] Ganter B and Wille R, *Formal Concept Analysis — Mathematical Foundations*, Springer, Berlin, Heidelberg, 1999.
- [6] Aranda-Corral G A, Borrego-Díaz J, and Galán J, Confidence-based reasoning with local temporal formal contexts, *Lecture Notes in Computer Science*, 2011, **6692**: 461–468.
- [7] Carling C, Williams A M, and Reilly T, *Handbook of Soccer Match Analysis*, Routledge Press, New York, 2005.
- [8] Oberstone J, Differentiating the top English premier league football clubs from the rest of the pack: Identifying the keys to success, *Journal of Quantitative Analysis in Sports*, 2009, **5**(3): Article 10.
- [9] Goldstein D G and Gigerenzer G, Fast and frugal forecasting, *Int. J. of Forecasting*, 2009, **25**(4):

- [10] Goldstein D G and Gigerenzer G, Models of ecological rationality: The recognition heuristic, *Psychological Review*, 2002, **109**(1): 75–90.
- [11] Min B, Kim J, Choe C, Eom H, and McKay R I, A compound framework for sports results prediction: A football case study, *Know. Based Syst.*, 2008, **21**(7): 551–562.
- [12] Imberman S P, Domanskiand B, and Orchard R A, Using booleanized data to discover better relationships between metrics, *Proc. Int. CMG Conference*, 1999, 530–539.
- [13] Aranda-Corral G A, Borrego-Díaz J, and Galán-Páez J, Bounded rationality for data reasoning based on formal concept analysis, *Proc. Int. Workshop on Database and Expert Systems Applications*, 2011, 350–354.
- [14] Simon H A, *Models of Bounded Rationality*, MA: MIT Press, Cambridge, 1982.
- [15] Andersson P, Ekman M, and Edman J, Forecasting the fast and frugal way: A study of performance and information-processing strategies of experts and non-experts when predicting the World Cup 2002 in soccer, *Working Paper Series in Business Administration*, 2003, **9** Stockholm School of Economics.
- [16] Guigues V and Duquenne J L, Familles minimales d' implications informatives resultant d'un tableau de donnees binaires, *Math. Sci. Humaines*, 1986, **95**: 5–18.
- [17] Aranda-Corral G A and Borrego-Díaz J, Reconciling knowledge in social tagging web services, *Lecture Notes in Artificial Intelligence*, 2010, **6077**: 383–390.
- [18] Balcázar J L, Redundancy, deduction schemes, and minimum-size bases for association rules, *Logical Methods in Computer Science*, 2010, **6**(2): 1–23.
- [19] Giarratano J C and Riley G D, *Expert Systems: Principles and Programming*, Brooks/Cole Publishing Co, Pacific Grove, CA, 2005.
- [20] Motter A E, de Moura A P S, Lai Y, and Dasgupta P, Topology of the conceptual network of language, *Physical Review E*, 2002, **65**(6): 065102.
- [21] Clauset A, Shalizi C R, and Newman M E J, Power-law distributions in empirical data, *SIAM Review*, 2009, **51**(4): 661–703.
- [22] Albert R and Barabási A L, Statistical mechanics of complex networks, *Reviews of Modern Physics*, 2002, **74**(1): 47–97.
- [23] Carmichael F, Thomas D, and Ward R, Team performance: The case of English premiership football, *Managerial and Decision Economics*, 2000, **21**(1): 31–45.
- [24] Goldstein D G and Gigerenzer G, Reasoning the fast and frugal way: Models of bounded rationality, *Psychological Review*, 1996, **103**(4): 650–669.
- [25] Brunswik E, Representative design and probabilistic theory in a functional psychology, *Psychological Review*, 1955, **62**(3): 193–217.
- [26] DiFatta G, Haworth G, and Regan K, Skill rating by Bayesian inference, *Proc. 2009 IEEE Symposium on Computational Intelligence and Data Mining*, 2009, 89–94.