Bounded Rationality for Data Reasoning based on Formal Concept Analysis

Gonzalo A. Aranda-Corral
Department of Information Technology
Universidad de Huelva
Palos de La Frontera, Spain
Email:gonzalo.aranda@dti.uhu.es

Joaquín Borrego-Díaz and Juan Galán-Páez

Department of Computer Science and Artificial Intelligence

Universidad de Sevilla

Sevilla, Spain

Email:jborrego@us.es, juangalan@us.es

Abstract—Formal Concept Analysis (FCA) is a theory whose goal is to discover and extract Knowledge from qualitative data. It also provides tools for sound reasoning (implication basis and association rules). The aim of this paper is to apply FCA to a new model for bounded rationality based on the implicational reasoning over contextual knowledge bases which are obtained from contextual selections. A contextual selection is a selection of events and attributes about them which induces partial contexts from a global formal context. In order to avoid inconsistencies, association rules are selected as reasoning engine. The model is applied to forecast sport results.

Keywords-Formal Concept Analysis, Bounded Rationality, Confidence Reasoning

I. Introduction

Bounded Rationality (BR) is intimately related with the human capacity for making inferences under limited time and Knowledge [1]. From the viewpoint of 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 traditional approaches to classical logic is not sound for explaining the success of several of them, as for example Recognition Heuristic (RH) [2]. 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. [3]; see also [4]). In fact, an intriguing question from ecological rationality analysis is: How could more knowledge be no better—or worse—than significantly less knowledge? [2]. One of the key features in BR is that inference process is concentrated on a limited set of experiences in which objects, properties and actions are selected. In this paper we aim to model this feature with Formal Concept Analysis.

Formal Concept Analysis (FCA) [5] is a mathematical theory for data analysis, using formal contexts and concept lattices as key tools. Domains can be formally modelled according to the extent and the intent of each formal concept. In FCA, the basic data structure is a formal context (with a qualitative nature) which represents a set of objects and their properties. It is useful both to detect and to describe regularities and the relationship structures among concepts.

It also provides a sound formalism for reasoning with such structures, mainly implication basis and association rules.

Roughly speaking, formal contexts represent weak structures easily built from experience that allow the extraction of knowledge from them. Despite its simple data structure, formal contexts are useful structures for knowledge extraction (cf. [5]) and reasoning. Moreover, in BR it is well known that in several cases simple statistical forecasting rules, which are usually simplifications of models, have been shown to make better predictions than more complex rules, especially when the future values of a criterion are highly uncertain [6]. The thesis of the paper is that association rules associated to formal contexts can be an interesting source for BR.

The aim is to present a logical model of BR based on reasoning on subcontexts of a predetermined (global) context which plays the role of global memory/qualitative dataset. The model is based on the existence of some selective processes (named contextual selection here) which induce specific contexts, and implicational basis are extracted from them (namely Stem Basis [7] and association rules). The reasoning with these Knowledge Bases (KB) (called contextual KB) is the model reasoning proposed in the paper. Logical combination of contextual KBs in order to avoid inconsistencies with background Knowledge can be made. The model has been used in [8] to describe a confidencebased (and contextual) reasoning system for forecasting sports betting. In this paper we analyse the soundness of the formal model as one of bounded rationality, presenting the theoretical framework.

The structure of the paper is as follows. The next section reviews the main elements of FCA and its logical features. In section 3 the role of formal contexts as basic bricks for a model of bounded rationality is presented. Section 4 reviews an experiment by using the model for forecasting in sports betting. Section 5 is devoted to describe future work.

II. BACKGROUND: FORMAL CONCEPT ANALYSIS

According to R. Wille, FCA [5] 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 of all common attributes valid for all the

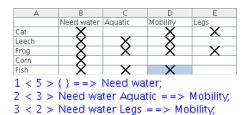


Figure 1. partial context from observation and Stem Basis

objects under consideration. It also allows the computation of concept hierarchies from data tables. In this section, we succinctly present basic FCA elements (see [5] for details).

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 Fig. 1 for an example of formal context about live beings.

The FCA main goal is the computation of the concept lattice from the context. For $X \subseteq O$ and $Y \subseteq A$ we define

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

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

A (formal) concept is a pair (X,Y) such that X'=Y and Y' = X.

In this paper it works with logical relations on attributes which are valid in the context and the standard implicational logic in FCA (see, e.g., [5]), called implications between attributes.

Definition. 1 An implication between attributes is a pair of sets of attributes, written as $Y_1 \rightarrow Y_2$,

An implication is true with respect to a formal context M = (O, A, I) according to the following definition. A subset $T \subseteq A$ respects $Y_1 \to Y_2$ if $Y_1 \not\subseteq T$ or $Y_2 \subseteq T$. It says that $Y_1 \to Y_2$ holds in M $(M \models Y_1 \to Y_2)$ if for all $o \in O$, the set $\{o\}'$ respects $Y_1 \to Y_2$. In that case, it is said that $Y_1 \to Y_2$ is an implication of M.

Definition. 2 Let \mathcal{L} be a set of implications and L be an implication of M.

- 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) If \mathcal{L} is a (implication) basis for M is complete and non-redundant.

A well-known method for computing specific implicational basis, called Stem Basis (SB), exists [7]. It is implemented into Conexp (http://sourceforge.net/projects/conexp/) software. A SB for live beings' formal context is provided in Fig. 1. It is important to remark that SB is only an example of a basis for a formal context. In this paper any specific property of the SB can be used, so it can be replaced by any implication basis.

It is possible to extend \models in relation to any propositional formula with propositional variables in A, by considering each object $o \in \mathbb{M}$ as a valuation v_o on \mathbb{A} defining

$$v_o(A) = 1 \iff (o, A) \in \mathbb{I}$$

So $M \models F$ if and only if $v_o \models F$ for any $o \in O$.

By defining \vdash_A as the proof relation induced by Armstrong rules [9],

$$R1: \frac{X \to X}{X \cup X} \ R2: \frac{X \to Y}{X \cup Z \to Y} \ R3: \frac{X \to Y, \ Y \cup Z \to W}{X \cup Z \to W}$$

it holds that the implicational bases are \vdash_A -complete (a straightforward consequence of Armstrong's result [9]):

Theorem 3 Let \mathcal{L} be a basis for a formal context M, and L an implication. Then $M \models L$ if and only if $\mathcal{L} \vdash_A L$.

In order to work with formal contexts, stem basis and association rules, the 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 based on what was designed for [10]. Initially it works with Stem Basis and entailment is based on the following result:

Theorem 4 Let \mathcal{L} be a basis for the context M and $\{A_1,\ldots,A_n\}\cup Y\subseteq A$. The following conditions are equivalent:

- 1) $S \cup \{A_1, \dots A_n\} \vdash_p Y (\vdash_p \text{ is the entailment with the }$ production system).
- 2) $S \vdash_A A_1, \dots A_n \to Y$ 3) $M \models \{A_1, \dots A_n\} \to Y$.

A. Association rules for a a formal context

We can consider a Stem Basis as an adequate knowledge base for a production system in order to reason. However, Stem Basis is designed for entailing true implications only, without any exceptions in the object set nor implications with a low number of counterexamples in the context.

Another more important question arises when it works on predictions. In this case we are interested in obtaining methods for selecting a result among all obtained results (even if they are mutually incoherent). Theorem 4 does not provide such a method. Therefore, it is better to consider association rules (with confidence) instead of true implications. Moreover, the initial production system must be revised for working with confidence.

Investigations on sound logical reasoning methods with association rules is a relatively recent research line with promising applications [11]. In FCA, association rules are implications among sets of attributes. Confidence and support are defined as usual. Recall that the *support* of X,

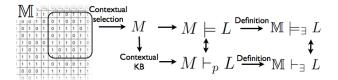


Figure 2. Model for reasoning based on ⊢¬

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 a association rule is $conf(X \to Y) = supp(X \cup Y)/supp(X)$. 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.

III. FORMAL CONTEXTS AS KNOWLEDGE STRUCTURES

Global memory is composed of events (objects) which have a number of properties (attributes). They constitute a global formal context $\mathbb{M} = (\mathbb{O}, \mathbb{A}, \mathbb{I})$ (which we call monster context following the tradition in Model Theory) from which subcontexts are extracted. Once the specific subcontext is considered, it is also possible to consider background knowledge Δ which would be combined with the KB extracted from formal context (Stem basis or association rules).

Definition. 5 *Let* \mathbb{M} *be a monster context, and let* $O \subseteq \mathbb{O}$.

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

The reasoning model on \mathbb{M} is argumentative, where the argument is based on KBs extracted from subcontexts. Anagously to [12], the existential arguments are considered, but replacing the consistent set by subcontext:

Definition. 6 Let L be an implication and Δ a background knowledge. It is said that L is a possible consequence of \mathbb{M} under the background knowledge Δ , $\mathbb{M} \models^{\Delta}_{\exists} L$, if there exists M a nonempty subcontext of \mathbb{M} such that $M \models \Delta \cup \{L\}$.

Note that by theorem 4, when Δ is a set of implications, it holds that \models_\exists is equivalent to \vdash_\exists which is defined by: $\mathbb{M} \vdash_\exists L$ if there exists $M \models \Delta$ a subcontext of \mathbb{M} such that $S \vdash_p L$ (where S is a SB for M). In the example described

in Sect. 4, the reasoning model is based on \vdash_p on contextual KBs for an object $o \in O$ w.r.t. a selection given by an expert. To compute all consequences by $\vdash_{\exists}^{\Delta}$ implies to consider the entire model. However we only need consequences entailed by a submodel. See section IV bellow.

Given $M_i = (O_i, A_i, I_i)$, i = 1, 2 two subcontexts of \mathbb{M} the intersection of M_1 and M_2 , $M_1 \cap M_2$ is

$$(O_1 \cap O_2, A_1 \cup A_2, I_1 \cap ((O_1 \cap O_2) \times A_1) \cup I_2 \cap ((O_1 \cap O_2) \times A_2))$$

In order to study \vdash_\exists under background knowledge, it is necessary to study the relationship among arguments based on distinct contexts. Two compatibility notions can be used.

Definition. 7 Let $M_i = (O_i, A_i, I_i)$, i = 1, 2 be two subcontexts of \mathbb{M} , and let Δ be a background propositional knowledge on the language of $A_1 \cap A_2$.

- It is said that M₁ and M₂ are compatible w.r.t Δ if there exists a supercontext M of M₁ and M₂ such that M ⊨ Δ.
- It is said that M₁ and M₂ are downward compatible w.r.t Δ if M₁ ∩ M₂ |= Δ.

Compatible contexts are also downward compatible. Therefore, it can jointly extend downward compatible contexts but they can not be restricted with logical reliability. Thus, formal contexts can be extended in order to refine results. It is also possible to work with any context whose objects satisfy background knowledge Δ to obtain \vdash_\exists consequences.

Proposition. 8 If two contexts are compatible then they are downward compatible

Proof: Suppose that M_1 and M_2 are compatible. Let M be the supercontext for M_1 and M_2 . By considering each object $o \in \mathbb{M}$ as a valuation v_o on \mathbb{A} defined by

$$v_o(A) = 1 \iff (o, A) \in \mathbb{I}$$

the objects in $M_1 \cap M_2$ are models of Δ . Thus $M_1 \cap M_2 \models \Delta$

The reciprocal is not true: Consider the context M=(O,A,I) with $O=\{o_1,o_2,o_3\}$ and $A=a_1,a_2,a_3$ and let $I=\{(o_1,a_1),(o_1,a_3),(o_2,a_2),(o_3,a_1),(o_3,a_3)\}$. Let M_1 be the subcontext with $O_1=\{o_1,o_2\}$ and $A_1=\{a_1,a_2\}$ and let M_2 be the subcontext with $O_2=\{o_2,o_3\}$ and $A_2=a_2,a_3$. The intersection $M_3=M_1\cap M_2$ has $O_3=\{o_2\}$ and $A_3=a_1,a_2,a_3$ and $I_3=\{(o_2,a_2)\}$. Since $M_3\models a_2$, we have that M_1 and M_2 are downward compatible w.r.t. $\{a_2\}$ (seen as a propositional formula), yet there is no supercontext of M_1 and M_2 satisfying a_2 .

The inference process for $\vdash_{\exists}^{\Delta}$ has tree steps (Fig. 2):

1) A question on whether a new event (object) has a property (attribute) is raised. On the new object some properties are known (attribute values) $\{A_1, \ldots A_n\}$.

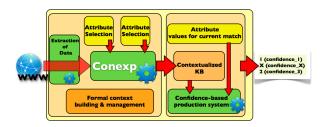


Figure 3. Context based reasoning system

- 2) A contextual selection outputs a subcontext of \mathbb{M} . A contextual KB \mathcal{L} (for some confidence threshold) is computed for the subcontext. Selection made by a user is composed by a small set of attributes.
- 3) The production system is executed on $\mathcal{L} \cup \{A_1, \dots A_n\}$. The results obtained are the attributes inferred about the event.

It has that, if A' is inferred by the production system, then

$$\mathbb{M} \vdash^{\Delta}_{\exists} \{A_1, \dots A_n\} \to \{A'\}$$

Note that it does not compute all implications, only those which are entailed from the attributes selected by the user.

A. Incompatible attributes

Implication logics do not suffer from inconsistency issues. However, in FCA it can be usual to consider incompatible attributes. A pair A_1, A_2 of incompatible attributes verifies that $\mathbb{M} \models \neg (A_1 \land A_2)$. If such a formula is included in background knowledge Δ it is possible to deal with inconsistency issues, because \vdash_\exists is an argumentative entailment which works on subcontexts (see def. of \vdash_\exists in [12]).

Two options have been considered for solving this problem by FCA. The first one is based on avoiding inconsistencies using *conservative retraction* [13]. The aim is to remove one attribute from incompatible pairs in contextual KB, next to present related attribute proof. We have selected a second option: to use association rules and do not consider any background knowledge. In this way, inferred attribute with maximum confidence is selected [8].

IV. EXAMPLE: DATA ON SOCCER LEAGUE MATCHES

We have applied the model for soccer betting. The project starts with the hypothesis that past data hides trends of soccer teams that experts use for forecasting. The monster model is composed of data (past and current) on matches (the objects, with temporal stamp). The experiment carries out seven stages (see Fig. 3):

- 1) **Selection of the set of relevant attributes** to consider. It is made by the user.
- 2) Data extraction. With this data the system is capable of building any subcontext needed. In the data time stamps are important, because a number of attributes deal with past matches. Data has been extracted from

the RSSSF Archive (http://www.rsssf.com) from the past ten years. Objects are matches (with temporal stamp), and attributes are computed for each object. The relevant properties (attributes) that experts selected was 17, several of them are parametrized (for example, ranking difference above a threshold). An attribute with threshold can produce a large number of binary attributes by changing the threshold. Thus the explicit computation of \mathbb{M} is not feasible. It has three distinguished attributes, corresponding of $Team_1$ wins (1), $Team_2$ wins (2) and draws (X).

- 3) Selection of a (future) match.
- Contextual selection, based on the selection of thresholds for attributes.
- Computing of attribute values for the match from data (to build the subcontext), except obviously the value of distinguished attributes.
- 6) Execution of the System (association rules as KB and attributes for the object as facts). Several modes for confidence computing, based on uncertain reasoning techniques by Expert Systems are considered [8].
- 7) **Result:** a triple $<(1, c_1), (X, c_x), (2, c_2)>$ of pairs (attribute, confidence), for the selected match.

V. EXPERIMENTS

Two experiments were launched for Spanish soccer league, on 2009-10 and 2010-11 seasons. Attributes were selected according to authors' knowledge about Spanish soccer league (which are not experts). From this contextual selection, \vdash_\exists was computed for all matches and weeks.

2009-10 season: Experiments with the system show forecasts of about 58.16% by a contextual selection based on the previous 38 matches. Such a percentage of hits for a qualitative reasoning system may be considered as an acceptable result comparable with expectable results of experts [14]. Experiments also shows 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: A way to evaluate how good is this forecasting sistem is comparing number of successes in our pool with the most popular betting selections. This popular selections are collected from the most voted results for each match, published at state agency web that controls soccer pools. In Fig. 4 both results are compared. Our hits are in blue and popular ones in green and last seventeen weeks from 2010-11 season are represented. Note that Spanish soccer pools are over 15 matches.

VI. CONCLUDING REMARKS AND FUTURE WORK

The model presented is concerned with association rule reasoning and it does not use -in its current form- more sophisticated probability tools (se e.g. [15]). As is stated in [16], the theory of probabilistic mental models assumes

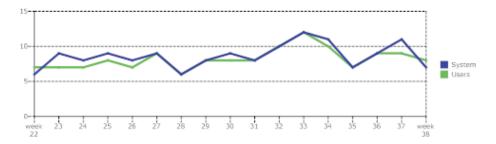


Figure 4. Correct predictions on the last 17 weeks of the season 2010-11 compared with popular the most popular bets

that inferences about unknown states of the world are based on probability cues [17]. In some sense, association rules's confidence plays the role of probability cues in the model.

The relationship of our proposal with RH [2] (roughly speaking, if one of the possibilities is recognized and the other is not, then infer that the recognized object has the higher value with respect to the criterion) is not clear. We may assert that our model recognises trends in contexts. Trends (represented as association rules or implication basis) can be considered as a kind of recognizing method, though.

It is worth noting that it only uses \vdash_\exists because the aim is to simulate bounded reasoning. Other entailment relationships from argumentative framework, as for example \vdash_\forall , have not been considered in this paper, because it requires of an exhaustive exploration of \mathbb{M} .

Part of our ongoing work includes two research lines. The first one is the analysis of conservative retraction method for working with incompatible attributes. The second one is to simulate the attribute learning process in our model, applying non monotone reasoning techniques.

ACKNOWLEDGMENTS

Supported by TIN2009-09492 project of Spanish Ministry of Science and Innovation, and *Excellence project* TIC-6064 of *Junta de Andalucía* cofinanced with FEDER founds.

REFERENCES

- [1] H. A. Simon, *Models of bounded rationality*. MIT Press, 1982.
- [2] D. G. Goldstein and G. Gigerenzer, "Models of ecological rationality: the recognition heuristic." *Psychological Review*, vol. 109, no. 1, pp. 75–90, 2002.
- [3] D. Kahneman, P. Slovic, and A. Tversky, (eds.) Judgment under uncertainty: Heuristics and biases. Cambridge University Press, 1982.
- [4] T. Pachur and G. Biele, "Forecasting from ignorance: the use and usefulness of recognition in lay predictions of sports events." *Acta Psychologica*, vol. 125, no. 1, pp. 99–116, 2007.
- [5] B. Ganter and R. Wille, Formal Concept Analysis: Mathematical Foundations. Springer, 1999.

- [6] D. G. Goldstein and G. Gigerenzer, "Fast and frugal forecasting," *International Journal of Forecasting*, vol. 25, no. 4, pp. 760–772, 2009.
- [7] J.-L. Guigues and V. Duquenne, "Familles minimales d'implications informatives resultant d'un tableau de donnees binaires," *Mathematiques et Sciences Humaines*, vol. 95, pp. 5–18, 1986.
- [8] G. A. Aranda-Corral, J. Borrego-Díaz, and J. Galán-Páez, "Confidence-based reasoning with local temporal formal contexts," in *IWANN* (2), 2011, pp. 461–468.
- [9] W. W. Armstrong, "Dependency structures of data base relationships." in *IFIP Congress*, 1974, pp. 580–583.
- [10] G. A. Aranda-Corral and J. Borrego-Díaz, "Reconciling knowledge in social tagging web services," in *HAIS* (2), 2010, pp. 383–390.
- [11] J. L. Balcázar, "Redundancy, deduction schemes, and minimum-size bases for association rules," *Logical Methods* in Computer Science, vol. 6, no. 2, 2010.
- [12] A. Hunter, "Reasoning with inconsistency in structured text," Knowledge Engineering Review, vol. 15, pp. 200–0, 1999.
- [13] G. A. Aranda-Corral, J. Borrego-Díaz, and M. M. Fernández-Lebrón, "Conservative retractions of propositional logic theories by means of boolean derivatives: Theoretical foundations," in *Calculemus/MKM*, 2009, pp. 45–58.
- [14] P. Andersson, M. Ekman, and J. Edman, "Forecasting the fast and frugal way: A study of performance and informationprocessing strategies of experts and non-experts when predicting the world cup 2002 in soccer," Stockholm School of Economics, Tech. Rep. 2003:9, May 2003.
- [15] B. Min, J. Kim, C. Choe, H. Eom, and R. Bobmckay, "A compound framework for sports results prediction: A football case study," *Knowledge-Based Systems*, vol. 21, no. 7, pp. 551–562, 2008.
- [16] G. Gigerenzer and D. G. Goldstein, "Reasoning the fast and frugal way: models of bounded rationality." *Psychological Review*, vol. 103, no. 4, pp. 650–669, 1996.
- [17] E. Brunswik, "Representative design and probabilistic theory in a functional psychology." *Psychological Review*, vol. 62, no. 3, pp. 193–217, 1955.