Interactive visual exploration of association rules with rule-focusing methodology
— **Source link** ⧉

Julien Blanchard, Fabrice Guillet, Henri Briand

Institutions: École polytechnique de l'université de Nantes

Topics: Association rule learning, Knowledge extraction, User interface, Expert system and Interactive visualization

Related papers:

- FIsViz: a frequent itemset visualizer

- WiFIsViz: Effective Visualization of Frequent Itemsets

- Pruning and visualizing generalized association rules in parallel coordinates

- Mining association rules between sets of items in large databases

- Query, analysis, and visualization of hierarchically structured data using Polaris

Interactive visual exploration of association rules with rule-focusing methodology

Julien Blanchard, Fabrice Guillet, Henri Briand

Interactive Visual Exploration of Association Rules with Rule Focusing Methodology

Julien Blanchard, Fabrice Guillet and Henri Briand

KnOwledge & Decision Team (KOD)
LINA – FRE CNRS 2729, Polytechnic School of Nantes University, France

Abstract. On account of the enormous amounts of rules that can be produced by data mining algorithms, knowledge post-processing is a difficult stage in an association rule discovery process. In order to find relevant knowledge for decision-making, the user (a decision-maker specialized in the data studied) needs to rummage through the rules. To assist him/her in this task, we here propose the *rule focusing* methodology, an interactive methodology for the visual post-processing of association rules. It allows the user to explore large sets of rules freely by focusing his/her attention on limited subsets. This new approach relies on rule interestingness measures, on a visual representation, and on interactive navigation among the rules. We have implemented the rule focusing methodology in a prototype system called *ARVis*. It exploits the user's focus to guide the generation of the rules by means of a specific constraint-based rule-mining algorithm.

Keywords: knowledge discovery in databases, association rules, post-processing, interactive visualization, rule focusing, constraint-based mining, interestingness measures, neighborhood of rules

1. Introduction

Among the knowledge models used in Knowledge Discovery in Databases (KDD), association rules (Agrawal et al, 1993) have become a major concept and have received significant research attention. Association rules are implicative tendencies $X \rightarrow Y$ where X and Y are conjunctions of items (boolean variables of the form *databaseAttribute=value*). The left-hand side X is the antecedent of the rule and

the right-hand side Y the consequent. Such a rule means that most of the records which verify the antecedent in the database verify the consequent too. For instance, in market basket analysis where the data studied are the customers' transactions in a supermarket, an association rule $\{pizza, crisps\} \rightarrow \{beer\}$ means that if a customer buys a pizza and crisps then (s)he most probably buys beer too. Since the pioneering algorithm of Agrawal, called Apriori (Agrawal and Srikant, 1994), many algorithms have been proposed for association rule mining (cf. Hipp et al (2000) for a survey). They generally produce very large amounts of rules. This is due to the unsupervised nature of association rule discovery. Indeed, because the user does not know precisely enough what (s)he is looking for to express it with the data terminology, (s)he does not make his/her goals explicit and does not specify any endogenous variable. Thus, the algorithms search all the valid associations existing in the database and generate an amount of rules exponentially growing with the number of items.

A crucial step in association rule discovery is post-processing, i.e., the interpretation, evaluation and validation of the rules in order to find interesting knowledge for decision-making. Because of the oversized amounts of rules, the post-processing stage often turns out to be a second mining challenge called "knowledge mining". While data mining is automatically computed by combinatorial algorithms, the knowledge mining stage is manually done by the user (a decision-maker specialized in the data studied). In practice, it is very difficult for users to rummage through the rules and find interesting ones in a corpus that can hold hundreds of thousands of rules, or even millions of rules with large business databases.

Many authors have stressed that the KDD process is by nature highly iterative and interactive and requires user involvement (Silberschatz and Tuzhilin, 1996) (Fayyad et al, 1996). In particular, Brachman and Anand (1996) have pointed out that in order to efficiently assist the users in their search for interesting knowledge, the KDD process should be considered not from the point of view of the discovery algorithms but from that of the users', as a human-centered decision support system. The human-centered approaches aim at creating a retroaction loop between the user and the system which constantly takes into account the information processing capacities of the user (cf. Bisdorff (2003) for examples of applications). Adopting Brachman & Anand's point of view, in this article we propose the *rule focusing* methodology, a human-centered methodology for the post-processing of association rules. The rule focusing methodology allows the user to explore large sets of rules by focusing his/her attention on successive limited subsets. The methodology relies on several *neighborhood relations* that connect the rules among them according to the user's semantics. With these relations, the user can navigate freely among the subsets of rules and thus drive the post-processing. In this way, a voluminous set of rules is explored subset by subset so that the user does not need to appropriate it entirely. Our approach combines:

– rule interestingness measures to filter and sort the rules,

– a visual representation to make comprehension easier,

– interactivity based on the neighborhood relations to guide the post-processing.

The rule focusing methodology can be used in two ways. First, it can be applied after association rule mining, as a pure post-processing technique. This is also called post-analysis or a posteriori filtering of rules (Hipp and Gntzer, 2002).

Secondly, it can be applied during association rule mining, as an interactive mining technique conducted by the user. Effectively, the rule focusing methodology induces a constraint-based rule-mining algorithm. A constraint-based rule-mining algorithm exploits constraints that the user gives to specify which kind of rules (s)he wants to find (cf. for example Srikant et al (1997), Ng et al (1998), Goethals and Van den Bussche (2000), Jeudy and Boulicaut (2002), Ordonez et al (2006)). Syntactic constraints (constraints specifying the items that must occur or not in the rule) and interestingness measure threshold constraints are the most commonly used constraints, but more general studies concern the so-called anti-monotone and succinct constraints (Ng et al, 1998), and the monotone constraints (Grahne et al, 2000) (Bonchi et al, 2005). Constraints allow to significantly reduce the exponentially growing search space of association rules[1]. Thus, the constraint-based algorithms can mine dense data more efficiently than the classical Apriori-like algorithms (the FP-growth-based algorithms of (Han et al, 2000) can mine dense data too, but they use a condensed representation of the data and require that it holds in memory). Besides, with appropriate constraints, the constraint-based algorithms can discover very specific rules which cannot be mined by the Apriori-like algorithms (the constraint-based algorithms can use low support thresholds for which Apriori-like algorithms are intractable). These rules are often very valuable for the users because they were not even thought of beforehand (Freitas, 1998). For these reasons, we use a constraint-based rule-mining algorithm to implement the rule focusing methodology in the prototype system described in this article. This specific algorithm extracts the rules interactively according to the user's focus. Note that using the rule focusing methodology as a pure post-processing technique or as an interactive mining technique is only a choice of implementation. The methodology does not depend on it.

The remainder of this article is organized as follows. In the next section we present a survey on association rule evaluation, exploration, and visualization. Then we describe the Information Visualization field of research, and in particular we compare 2D and 3D visualizations. Section 4 is dedicated to the study of cognitive constraints of the user during rule post-processing. From these constraints, in section 5, we define the rule focusing methodology. Section 6 describes the prototype system implementing our methodology: *ARVis*, a visual tool for association rule mining and post-processing . In section 7, we give an example of rule post-processing with *ARVis*. It comes from a study made with the firm PerformanSe SA on human resource management data. Finally we give our conclusion in section 8.

2. Survey on association rule evaluation, exploration, and visualization

At the output of the data mining algorithms, the sets of association rules are simple text lists. Each rule consists of a set of items for the antecedent, a set of items for the consequent (sets of items are called itemsets), and the numerical values of two interestingness measures, support and confidence (Agrawal et al, 1993). Support is the proportion of records which verify a rule in the database;

---

[1] Choosing the best way of harnessing multiple constraints whatever the data is still an open problem.

it evaluates the generality of the rule. Confidence (or conditional probability) is the proportion of records which verify the consequent among those which verify the antecedent; it evaluates the validity of the rule (success rate).

Three kinds of approaches aim at helping the user appropriate large sets of association rules:

- the user can filter and order the rules with other interestingness measures;
- the user can browse the large sets of rules with interactive tools or query languages;
- the user can visualize the rules.

2.1. Rule interestingness measures

It is now well-known that the support-confidence framework is rather poor to evaluate the rule quality (Silverstein et al, 1998) (Bayardo and Agrawal, 1999) (Tan et al, 2004). Numerous rule interestingness measures have been proposed to complement this framework. They are often classified into two categories: the subjective (user-oriented) ones and the objective (data-oriented) ones. Subjective measures take into account the user's a priori knowledge of the data domain (Liu et al, 2000) (Silberschatz and Tuzhilin, 1996) (Padmanabhan and Tuzhilin, 1999). On the other hand, the objective measures do not depend on the user but only on objective criteria such as data cardinalities or rule complexity. Depending on whether they are symmetric (invariable by permutation of antecedent and consequent) or not, they evaluate correlations or rules.

There exist two significant configurations in which the rules appear non-directed relations and therefore can be considered as neutral or non-existing (Blanchard, 2005):

- the *independence*, i.e., when the antecedent and consequent are independent;
- what we call the *equilibrium*, i.e., when examples and counter-examples are equal in numbers (maximum uncertainty of the consequent given that the antecedent is true).

Thus we distinguish two different but complementary aspects of the rule interestingness: the deviation from independence and the deviation from equilibrium. The objective measures of interestingness can be classified into two classes (Blanchard et al, 2005) (Blanchard et al, 2005):

- the measures of deviation from independence, which have a fixed value at independence, such as rule-interest (Piatetsky-Shapiro, 1991), lift (Silverstein et al, 1998), conviction (Brin et al, 1997), Loevinger index (Loevinger, 1947), implication intensity (Gras, 1996) (Blanchard et al, 2003);
- the measures of deviation from equilibrium, which have a fixed value at equilibrium, such as confidence (Agrawal et al, 1993), Sebag and Schoenauer index (Sebag and Schoenauer, 1998), IPEE (Blanchard et al, 2005).

These two kinds of measures are complementary (in particular, they do not create the same preorder on rules) (Blanchard, 2005) (Blanchard et al, 2005). A rule can have a good deviation from independence with a bad deviation from equilibrium, and conversely. Regarding the deviation from independence, a rule $A \rightarrow B$ with a good deviation means "When A is true, then B is **more** often true" (more than usual, i.e. more than without any information about A). On

the other hand, regarding the deviation from equilibrium, a rule $A \rightarrow B$ with a good deviation means "When A is true, then B is **very** often true". Deviation from independence is a comparison relatively to an expected situation, whereas deviation from equilibrium is an absolute statement. The measures of deviation from independence are useful to discover associations between antecedent and consequent (do the truth of A influence the truth of B?), while the measures of deviation from equilibrium are useful to take decisions or make predictions about B (knowing or supposing that A is true, is B true or false?) (Blanchard, 2005).

2.2. Interactive browsing of rules

Interactive tools of the type "rule browser" have been developed to assist the user in the post-processing of association rules. First, Klemettinen et al (2004) present a browser with which the user reaches interesting rules by adjusting thresholds on interestingness measures and applying syntactic constraints (templates). Secondly, Liu et al (1999) propose a rule browser which is based on subjective interestingness measures and exploits the user's a priori knowledge of the data domain to present the rules. The user expresses his/her knowledge under the form of relations and then the tool classifies the rules in different categories according to whether they confirm or not the user's beliefs. Finally, in (Ma et al, 2000), the user explores a summary of the rules. (S)he can access the rules by selecting elements in the summary. The main limit of all these tools lies in the textual representation of the rules which does not suit the study of large amounts of rules described by numerous interestingness measures.

More recently, a rule browser equipped with numerous functionalities has been presented in (Fule and Roddick, 2004). It allows to filter the rules with syntactic constraints that are more or less general since they can take into account an item taxonomy. The tool also enables the user to program any interestingness measure to order and filter the rules. Besides, the user can save the rules that (s)he judges interesting during the exploration. Another rule browser is presented in (Tuzhilin and Adomavicius, 2002), but it is not a generic tool. It is dedicated to the analysis of gene expression data coming from DNA microarrays, and relies on a very complete system of syntactic constraints which can take into account a gene taxonomy.

Within the framework of inductive databases (Imielinski and Mannila, 1996), several rule query languages have been proposed, such as DMQL (Han et al, 1996), MINE RULE (Meo et al, 1998), MSQL (Imielinski and Virmani, 1999), or XMINE (Braga et al, 2002). They allow to mine (by means of constraint-based algorithms) and post-process rules interactively under the user's guidance. However, as regards rule post-processing, the query languages are not very user-friendly (cf. (Botta et al, 2002) for an experimental study).

2.3. Visualizing the rules

Visualization can be very beneficial to KDD (Fayyad et al, 2001). Visualization techniques are indeed an effective means of introducing human subjectivity into each stage of the KDD process while taking advantage of the human perceptual capabilities. The information visualization techniques can either be used as knowledge discovery methods on their own, which is sometimes called "visual
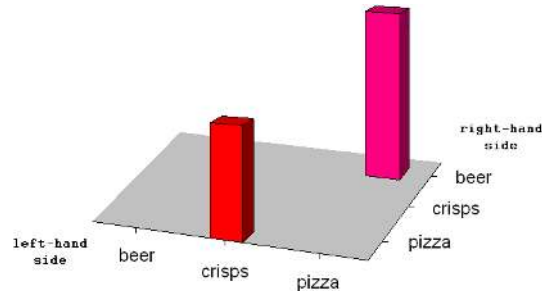
Fig. 1. An item-to-item matrix showing the rules $\{crisps\} \rightarrow \{pizza\}$ and $\{pizza\} \rightarrow \{beer\}$
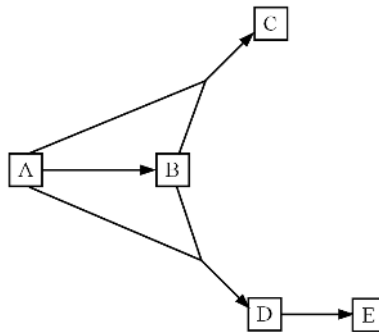

Fig. 2. A rule graph with items as nodes, showing the rules $A \rightarrow B$, $AB \rightarrow C$, $AB \rightarrow D$, and $D \rightarrow E$

data mining" (Keim, 2002), or they can collaborate with data mining algorithms to facilitate and speed up the analysis of data, intermediate results, or discovered knowledge (Aggarwal, 2002) (Schneiderman, 2002) (Han et al, 2003). Association rule visualization comes within this latter case. It must be noticed that the methods and tools presented below are generally supplied with basic functionalities for ordering and filtering the rules on items and on a few interestingness measures.

A first rule visualization method consists in using a matrix representation. Hofmann and Wilhelm (2001) and the Quest research group (Agrawal et al, 1996), as well as the software programs DBMiner (Han et al, 1997), SGI MineSet (Brunk et al, 1997), DB2 Intelligent Miner Visualization (**?**), and Enterprise Miner (**?**), give different implementations of it. In an item-to-item matrix (figure 1), each line corresponds to an antecedent item and each column to a consequent item. A rule between two items is symbolized in the matching cell by a 2D or 3D object whose graphical characteristics (generally size and color) represent the interestingness measures. This visualization technique has been improved into rule-to-item matrices (Wong et al, 1999) whose cluttering is lower and which allow a more efficient representation of rules with more than two items. The main limit of these approaches is that the matrices reach considerable sizes in case of large sets of rules over numerous items.

Sets of association rules can be also visualized by using a directed graph (Klemettinen et al, 2004) (Han et al, 1997) (Rainsford and Roddick, 2000) (**?**), the nodes and edges respectively representing the items and the rules (cf. figure
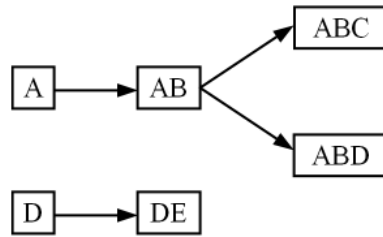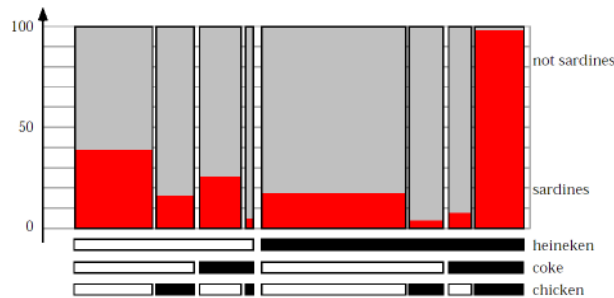
Fig. 3. A rule graph with itemsets as nodes, showing the rules $A \rightarrow B$, $AB \rightarrow C$, $AB \rightarrow D$, and $D \rightarrow E$

The represented rules involve the items *heineken*, *coke*, and *chicken* in antecedent, and *sardines* in consequent. Antecedent and consequent are respectively read on the horizontal and vertical axes. The first rectangle on the left shows that the rule $\{heineken = 0, coke = 0, chicken = 0\} \rightarrow \{sardines = 1\}$ has a confidence of about 40% (red part of the rectangle), while the opposite rule $\{heineken = 0, coke = 0, chicken = 0\} \rightarrow \{sardines = 0\}$ has a confidence of 60% (grey part of the rectangle). The support of the rules is proportional to the area (red or grey) of the rectangles. The rectangle on the right indicates that the confidence of the rule $\{heineken = 1, coke = 1, chicken = 1\} \rightarrow \{sardines = 1\}$ is about 100%.

Fig. 4. Mosaic display for association rules (from Hofmann et al (2000))

2 where letters denote items). The interestingness measures are symbolized on the edges, for instance with color or thickness. The graph representation is very intuitive but it has two main drawbacks. First, it makes transitivity appear among the rules whereas, in the general case, the rules are not transitive (with most measures, rule interestingness does not spread transitively). Secondly, it does not suit the visualization of large sets of rules over numerous items either. Indeed, the graph is then overloaded with nodes and crossing edges, all the more when rules with more than two items are considered. To improve the rule visualization, the same representation method has been used in 3D with a self-organization algorithm to guarantee a more efficient graph layout (Hao et al, 2001). Also we have proposed in (Kuntz et al, 2000) a dynamic rule graph which is a subgraph of the itemset lattice. In this graph, the nodes do not represent the items but the itemsets so that a rule $AB \rightarrow C$ is symbolized by an edge between the nodes AB and ABC (figure 3). The resulting graph is acyclic with more nodes but fewer edge crossings. The user can dynamically develop the graph as (s)he wishes by clicking on the nodes.

All the visual representations described so far are based on rule syntax (i.e. the items). A different approach is proposed in (Unwin et al, 2001), where the

representation is based on interestingness measures. This representation is a scatterplot between support and confidence where each point is colored according to density estimation. The user can query any point to display the names of the rules represented by the point (rules with close supports and confidence). The main advantage of such a representation is that it can contain a great number of rules. However, several rules can be represented by one and only one point, which does not facilitate the task of the users when they search for rules using items as criteria. This approach is the closest to the one we propose in this paper, which also uses a spatial mapping to highlight the interestingness measures.

Other methods have been proposed to represent association rules. Nevertheless, they do not deal with the visualization of the whole rule set but with the visualization of a pattern of rules (a group of rules with given items in antecedent and consequent). These methods allow a thorough study of a restricted number of rules, making their interpretation easier and helping to understand their occurrence context. We can quote for example Hofmann et al.'s mosaic plots (2000) for rules with categorical attributes (figure 4), or Fukuda et al (2001) and Han et al (2003) for numerical rules. Also some techniques inspired from parallel coordinates have been considered to visualize patterns of classification rules (Han et al, 2000) or association rules (Kopanakis and Theodoulidis, 2001).

3. Information visualization

3.1. Context

Information visualization (Card et al., 1999) (Spence, 2000) consists in representing abstract data under a visual form in order to improve cognition for a given task, that is to say the acquisition and use of new knowledge. The core of information visualization is visual encoding, i.e., the mapping of data tables to visual structures in a 2D or 3D space (Card et al., 1999). The visual structures have several graphical properties such as position, length, area, hue, brightness, saturation, shape, texture, angle, curvature... They can be zero-dimensional (points), one-dimensional (lines), two-dimensional (surfaces), or three-dimensional (volumes).

Several authors proposed classifications of visual encodings in order to show which ones are appropriate according to the data variables to be represented. Among these works, those of Cleveland and McGill (1984), Tufte (1983), and then Wilkinson (1999) are references for statistical graphs (charts). A second trend stems from cartography, with the works of McEachren (1995) and Bertin whose *Semiology of Graphics* (Bertin, 1967-1983) is considered as the first and most influential structural theory of graphics (Wilkinson, 1999). As regards the visual representation of quantitative variables, the two trends agree that the best encodings are done with position (Bertin, 1967-1983) (Cleveland and McGill, 1984) (McEachren, 1995) (Card et al., 1999) (Wilkinson, 1999). However, the two trends mainly differ about the use of surfaces to represent quantitative variables: this use is not advisable with statistical graphs whereas it is standard practice in cartography. In particular, Cleveland and McGill (1984) propose a hierarchy of visual encodings saying that surface is little appropriate (less than length) to represent quantities. This point of view is based on Stevens's law in psychophysics according to which the perceived quantities are not linearly related to the actual quantities with surface (Baird, 1970). On the other hand, Bertin points out that

before the variation of length, the variation of surface is the sensitive stimulus of the variation of size (Bertin, 1967-1983).

The visualization we propose in this article is not a map, and even less a statistical graph: this is a 3D virtual world. With the increase in the capacities of personal computers, the 3D virtual worlds have become common in information visualization (Chen, 2004). Associated with navigation operators (viewpoint controls), they have shown to be efficient for browsing wide information corpuses such as large file system hierarchies with Silicon Graphics' FSN (re-used in Mine-Set for the visualization of decision trees), hypertext document graphs with Harmony (Andrews, 1995), or OLAP cubes with DIVE-ON (Ammoura et al, 2001)] (cf. Chen (2004) for other examples of applications). While a 2D representation is restricted to the two dimensions of the screen, the additional dimension in a 3D virtual world offers a viewpoint towards infinity, creating a wide workspace capable of containing a large amount of information (Card et al., 1999). In this workspace, the most important information can be placed in the foreground (most visible objects) and thus be highlighted compared to the less important information placed behind it (less visible objects). This is the reason why the 3D representations are sometimes considered as focus+context approaches. Moreover, 3D enables to exploit volumes as objects in the visualization space. It allows to benefit from more graphical properties for the objects and thus to represent even more information.

3.2. 2D or 3D?

The choice between 2D and 3D representations for information visualization is still an open problem (Card et al., 1999) (Chen, 2004). This is especially due to the fact that the efficiency of a visualization is highly task-dependent (Carswell et al, 1991). Besides, while 3D representations are often more attractive, 2D has the advantage of a long and fruitful experience in information visualization. In fact, few research works are dedicated to the comparison between 2D and 3D. As regards the static (non interactive) visualization of statistical graphs, the 3D representations have generally not been advisable since the influential publications of Tufte (1983) and Cleveland and McGill (1984). Nevertheless, the psychophysics experiments of Spence (1990) and Carswell et al (1991) show that there is no significant difference of accuracy between 2D and 3D for the comparison of numerical values. In particular, Spence points out that this is not the apparent dimensionality of visual structures which counts (2 for a surface, 3 for a volume) but the actual number of parameters that show variability (Spence, 1990). In his experiments, whatever the apparent dimensionality of visual structures, Stevens's law is almost always the same when only one parameter actually varies (Stevens's law exponents are very close to 1). Under some circumstances, information may even be processed faster when represented in 3D rather than in 2D. As regards the perception of global trends in data (increase or decrease), the experimental results of Carswell et al (1991) also show an improvement in the answer times with 3D but to the detriment of accuracy.

Other works compare 2D and 3D within the framework of interactive visualization. Cockburn and McKenzie (2001) study the storage and retrieval of bookmarked web-pages in a 2D or 3D visualization space. With the 2D interface, the processing times of the users are shorter but not significantly. On the other hand, the subjective assessment of the interfaces shows a significant preference

for 3D (which Spence (1990) and Carswell et al (1991) also sense but without assessing it). Finally, Ware and Franck (1996) compare the visualization of 2D graphs and 3D graphs. Their works show a significant improvement in intelligibility with 3D. More precisely, their experiment consists in asking users whether there is a path of length two between two nodes randomly chosen in a graph. With the 3D graphs, the error rate is reduced by 2.2 for comparable answer times. With stereoscopic display, the error rate is even reduced by 3. One generally considers that only stereoscopy allows fully exploiting the characteristics of the 3D representations.

4. Cognitive constraints of the user during rule post-processing

4.1. User's task

During the post-processing of the rules, the user is faced with long lists of rules described by interestingness measures. The user's task is then to rummage through the rules in order to find interesting ones for decision-making. To do so, (s)he needs to interpret the rules in the business semantics and to evaluate their quality. The two decision indicators are therefore the rule syntax and the interestingness measures. The user's task is difficult for two reasons. First, the profusion of rules at the output of the data mining algorithms prevents any exhaustive exploration. Secondly, on account of the unsupervised nature of association rule discovery, it is generally not feasible for the user to obviously formulate constraints which would isolate relevant rules directly.

4.2. Cognitive hypotheses of information processing

On account of the human "bounded rationality" hypothesis (Simon, 1979), a decision process can be seen as a search for a dominance structure. More precisely, the decision-maker faced with a set of multiattribute alternatives tries to find an alternative (s)he considers dominant over the others, i.e., an alternative (s)he thinks better than the others according to his/her current representation of the decision situation (Montgomery, 1983). This type of models of decision process can be transferred to the post-processing of association rules by considering the rules as a particular kind of alternatives with items and interestingness measures as attributes. According to Montgomery, the decision-maker isolates a limited subset of potentially useful alternatives and makes comparisons among them. This can be done iteratively during the decision process. More precisely, he has pointed out that: "The decision process acquires a certain directionality in the sense that certain alternatives and attributes will receive more attention than others [...] The directionality of the process may be determined more or less consciously. Shifts in the directionality may occur several times in the process, particularly when the decision-maker fails to find a dominance structure".

Furthermore, a KDD methodology called "attribute focusing" has been proposed in (Bhandari, 1994). It results from experimental data concerning the user's behavior in the discovery process. This methodology is based on a filter which automatically detects a small number of potentially interesting attributes.
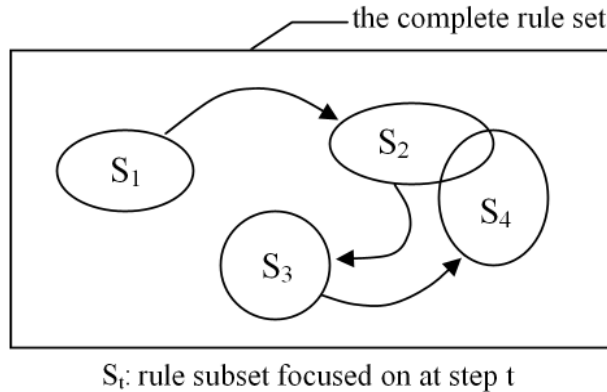
Fig. 5. Navigation among successive subsets of rules with the rule focusing methodology

The filter guides the user's attention on a small, and therefore more intelligible, subset of the database. The importance of focusing on a small number of attributes in human information processing has also been widely confirmed with works on decision strategies (cf. for example the moving basis heuristics in (Barthelemy and Mullet, 1992)). Indeed, on account of his/her limited cognitive abilities, the decision-maker examines only a small amount of information at each moment.

From these different works on human information processing, we establish three principles on which our rule focusing methodology relies:

P1. enabling the user to focus his/her attention on a limited subset of rules with a small number of attributes (items and interestingness measures),

P2. enabling the user to make comparisons among the rules in the subset,

P3. enabling the user to shift the subset of rules (s)he is focusing on at any time during the post-processing, until (s)he is able to validate some rules and reach a decision.

5. Rule focusing methodology

The idea of developing the *rule focusing* methodology has arisen from our earlier works on the visualization of rule sets by graphs (Kuntz et al, 2000). The methodology consists in letting the user navigate freely inside the large set of rules by focusing on successive limited subsets via a visual representation of the rules and their measures. In other words, the user gradually drives a series of visual local explorations according to his/her interest for the rules (figure 5). Thus, the rule set is explored subset by subset so that the user does not need to appropriate it entirely. At each navigation[2] step, the user must make a decision to choose which subset to visit next. This is the way subjectivity is introduced into the post-processing of the rules. The user acts here as an exploration heuristics.

---

[2] We call "navigation" the fact of going from one subset to another, while "exploration" refers to the whole process supervised by our methodology, i.e., the navigation among the subsets and the visits (local explorations) of the subsets.
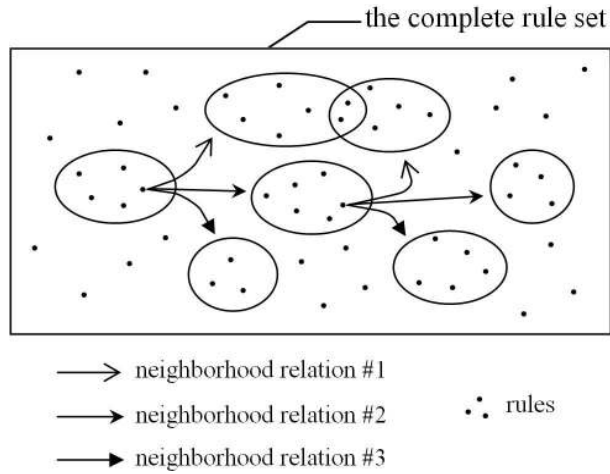
Fig. 6. A neighborhood relation associates each rule to a subset of rules

Exploiting a human heuristics is coherent since the function to be optimized, i.e. the user's interest, is subjective.

The rule focusing methodology integrates the cognitive principles of section 4.2 in the following way:

– Relations allow to focus on the subsets and to navigate among them (principles P1 and P3). We call them *neighborhood relations.*
– The user visualizes the subsets to visit them, and in particular to compare the rules (principle P2).

Both the neighborhood relations and the visualization technique must take into account the two decision indicators involved in the user's task: the rule syntax and the interestingness measures (cf. section 4.1).

5.1. Neighborhood relations among rules

The neighborhood relations determine the way the subsets of rules are focused on (cognitive principle P1) and the way the user can go from one subset to another (cognitive principle P3). They are a fundamental element of the rule focusing methodology since they are the vectors of the navigation for the user. These relations are defined in the following way: a neighborhood relation associates each rule from the complete set of rules to a limited subset of rules called neighbors (figure 6). So with x relations, the user can reach x neighboring subsets of rules from one rule, and from a subset containing y rules (s)he can reach $x.y$ possible neighboring subsets. To navigate from one subset to another, the user must make two choices: which neighborhood relation to apply, and on which rule.

In mathematical terms, the neighborhood relations are binary relations over the complete set R of the rules extracted by the data mining algorithms. Still with the aim of being appropriate to the user's task, we choose neighborhood relations which have a pertinent meaning for the user:

$\forall (r_1, r_2) \in R^2,$
$neighborOf(r_1, r_2) \Leftrightarrow$ (the user judges that r_1 is close to r_2 from a point of view)

This introduces user semantics into the navigation among the rules. Any relation could be considered provided it makes sense for the user. Consequently, the relations have to be defined with his/her help before starting the rule post-processing.

Here are for example four possible neighborhood relations $neighborOf(r_1, r_2)$:

1. r_1 is neighbor of r_2 if and only if r_1 and r_2 have the same conclusion;

2. r_1 is neighbor of r_2 if and only if r_1 is an exception of r_2;

3. r_1 is neighbor of r_2 if and only if the antecedent of r_1 is more general than that of r_2;

4. r_1 is neighbor of r_2 if and only if r_1 has the same support and confidence as r_2 to within about 0.05.

The neighborhood relations 1, 2, and 3 are based on the rule syntax, while relation 4 is based on two interestingness measures. Furthermore, relation 1 is an equivalence relation, whereas relation 2 is neither reflexive, nor symmetric, nor transitive. Relation 3 is only transitive, and relation 4 is reflexive and symmetric but not transitive.

Let us assume that the user applies a neighborhood relation Π. From a rule r, (s)he can reach the subset S of all the rules that are neighbors of r according to Π. We call r the "transitional rule" because it allows to navigate from one subset to another. Depending on the reflexivity of the relation Π chosen, S can or cannot contain the transitional rule r.

The originality of our methodology in comparison with the existing rule exploration techniques (described section 2.2) mainly lies in the concept of neighborhood relation. With a query language or an interactive interface like a rule browser, the user can reach any subset of rules but (s)he must explicitly specify the constraints which delimit it. With the rule focusing methodology, the constraint specification is implicit since it is hidden in the neighborhood relations. Actually, the neighborhood relations can be seen as generalizations of constraints (classes of constraints). We think that the user's task is made easier with neighborhood relations than with explicit constraints. For example, let us imagine the following exploration scenario:

The user finds an interesting rule $ABC \rightarrow D$ (where letters denote items). (S)he thinks that the combination of these four items is very pertinent but (s)he would like to change the order of the items and verify whether the rules $ABD \rightarrow C$, $ACD \rightarrow B$, and $BCD \rightarrow A$ (and why not the rules $AB \rightarrow CD$, $AC \rightarrow BD$ and so on as well) are better evaluated by the interestingness measures.

With the rule focusing methodology, the user can carry out this scenario in just one interaction. (S)he only needs the neighborhood relation "r_1 is neighbor of r_2 if and only if r_1 and r_2 have the same items". On the other hand, with a query language or a rule browser, the user has to write a series of appropriate queries or to specify a series of constraints manually with the graphical interface. This can be a tedious and time-consuming task.

5.2. Rule visualization

In order to help the user to visit the subsets of rules, we provide him/her with a visual representation instead of poorly intelligible textual lists. The visual representation facilitates and speeds up comprehension, and in particular it makes the comparisons among the rules easier (cognitive principle P2). Most of the techniques proposed for rule visualization have been developed to represent the whole set of rules produced by the data mining algorithms. Nevertheless, in the rule focusing methodology, we can take advantage of the user's focus strategy by representing only the current subset of rules at each navigation step. This reduces the number of rules to draw and above all largely improves the representation intelligibility. Visually, the user's point of view on the complete rule set is thus always local.

The rule visual representations are generally based on the rule syntax and handle interestingness measures as additional information (except for the approach of Unwin et al (2001), cf. section 2.3). However, the interestingness measures are also decision indicators fundamental to the user's task. So that the user can quickly assess and compare the rules, the representation must highlight the interestingness measures and make the best rules clearly recognizable. Also the visualization must be able to integrate numerous measures (not only support and confidence as it often happens), to dynamically filter the rules according to thresholds set by the user, and to support large amounts of rules having any number of items inside. Finally, the visualization must integrate interactive operators allowing the user to trigger the neighborhood relations.

Research works in visual perception show that a human being has first a global perception of a scene, before taking an interest in details. This is what motivated the development of the approaches named overview+details and focus+context (Card et al., 1999). Thus, in the rule focusing methodology, the user has to be able to easily change between global and detailed views of the rules by interacting with the visualization.

6. *ARVis*, a visual tool for association rule mining and post-processing

In this section, we present *ARVis (Association Rule Visualization)*, an experimental prototype implementing the rule focusing methodology. It was originally developed for the firm PerformanSe SA in order to find knowledge for decision support in human resource management. *ARVis* considers rules with single consequents (one item only in the consequent). This choice is usual in association rule discovery. Indeed, in association rule discovery in general and in our applications with PerformanSe SA in particular, the users are often interested spontaneously in this kind of rules because they are more intelligible than rules with multi-item consequents. However, considering only rules with single consequents is not a limitation to our approach. This choice could be easily changed.

At least three interestingness measures are calculated in *ARVis*: support, confidence (Agrawal et al, 1993), and entropic implication intensity (respectively noted sp, cf and eii). We choose support and confidence because they are the basic indexes to assess association rules. As for implication intensity, it is an asymmetric probabilistic index which evaluates the statistical significance

of the rules by quantifying the unlikelihood of the number of counter-examples (Guillaume et al, 1998) (Gras, 1996). The entropic version of this index also takes into account the imbalances between examples and counter-examples for both the rule and its contrapositive (Blanchard et al, 2003). The entropic implication intensity is a powerful measure since it takes into account both the deviation from independence and the deviation from equilibrium. This is the reason why we have chosen to integrate it into *ARVis*. But here again, this choice of measures is not a limitation to our approach, and others can be added. Each measure is associated to minimum and maximum thresholds set by the user: min_{sp}, min_{cf}, min_{eii}, max_{sp}, max_{cf}, max_{eii}. Although most of the tools for association rule mining do not provide them, the maximum thresholds improve the user's focus. For example, rules with high support and high confidence are often already known by the users; removing them allows highlighting more interesting rules.

In *ARVis*, we have opted for neighborhood relations mainly based on items and for a visualization technique mainly based on interestingness measures. We think this is the way to the most user-friendly solutions for rule exploration.

6.1. Neighborhood relations

Eight neighborhood relations are implemented in *ARVis*, most of them being generalization-type relations or specialization-type relations. Two of the most fundamental human cognitive mechanisms for generating new rules are indeed generalization and specialization (cf. the study of the reasoning processes in (Holland et al, 1986)).

Given the set I of items relative to the data studied, the rules are of the form $X \rightarrow y$ where X is an itemset $X \subset I$ and y is an item $y \in I - X$. The complete set of rules with single consequents that can be built with the items of I is noted R. In order to simplify the notations, we note $X \cup y$ instead of $X \cup \{y\}$ and $X - y$ instead of $X - \{y\}$. For the same simplicity reason, we define the neighborhood relations not as binary relations over R but as functions Π from R to 2^R which associate each rule with the subset composed of its neighbors:

$$\forall r_1 \in R, \Pi(r_1) = \{r_2 \in R \mid neighborOf(r_1, r_2)\}$$

Each of the eight neighborhood relations below induces two kinds of constraints:

- syntactic constraints, which specify the items that must occur or not in the antecedent and in the consequent;
- interestingness measure constraints, which specify minimum and maximum thresholds for the measures.

The syntactic constraints are peculiar to each neighborhood relation. On the other hand, the interestingness measure constraints are shared by all the relations. We group them together into the boolean function $interesting()$:

$$\forall r \in R, interesting(r) \Leftrightarrow \left| \begin{array}{l} min_{sp} \leq sp(r) \leq max_{sp} \\ min_{cf} \leq cf(r) \leq max_{cf} \\ min_{eii} \leq eii(r) \leq max_{eii} \end{array} \right.$$

A rule is said interesting if the three measures respect the minimum and maximum thresholds.

Specialization-type relations

$$agreement\_specialization(X \rightarrow y) = \left\{ X \cup z \rightarrow y \;\middle|\; \begin{array}{l} z \in I - (X \cup y) \\ interesting(X \cup z \rightarrow y) = true \end{array} \right\}$$

$$exception\_specialization(X \rightarrow y) = \left\{ X \cup z \rightarrow not(y) \;\middle|\; \begin{array}{l} z \in I - (X \cup \ not(y)) \\ interesting(X \cup z \rightarrow not(y)) \end{array} \right\}$$

$$forward\_chaining(X \rightarrow y) = \left\{ X \cup y \rightarrow z \;\middle|\; \begin{array}{l} z \in I - (X \cup \ y) \\ interesting(X \cup y \rightarrow z) \end{array} \right\}$$

Holland et al (1986) point out that a too general rule can be specialized into two kinds of complementary rules: exception rules and agreement rules. Exception rules aim at explaining the counter-examples of the general rule, while agreement rules aim at better explaining the examples. For instance, a rule "If α is a dog then α is friendly" can be specialized into the rules "If α is a dog and α is muzzled then α is mean" and "If α is a dog and α is not muzzled then α is friendly". The interest of exception rules in KDD has been widely confirmed (cf. for example (Hussain et al, 2000) (Suzuki, 2002)). On the basis of these two kinds of specialization, we propose the neighborhood relations *agreement_specialization* and *exception_specialization*[3] in *ARVis*. The third specialization-type relation is inspired by forward chaining in inference engines for expert systems: when a rule $X \rightarrow y$ is fired, the concept y becomes active and can be used with X to fire new rules and deduce new concepts z. Backward chaining cannot be considered with rules with single consequent.

Generalization-type relations

$$generalization(X \rightarrow y) = \left\{ X - z \rightarrow y \;\middle|\; \begin{array}{l} z \in X \\ interesting(X - z \rightarrow y) = true \end{array} \right\}$$

$$antecedent\_generalization(X \rightarrow y) = \left\{ X - z \rightarrow z \;\middle|\; \begin{array}{l} z \in X \\ interesting(X - z \rightarrow z) = true \end{array} \right\}$$

generalization relies on the condition-simplifying generalization mechanism described in (Holland et al, 1986). This relation is complementary to *agreement_specialization* and *exception_specialization*. It consists in deleting an item in the antecedent. The relation *antecedent_generalization* is complementary to *forward_chaining*. After applying *forward_chaining* on a rule r, one can effectively come back to r by applying *antecedent_generalization*.

Other relations

$$same\_antecedent(X \rightarrow y) = \left\{ X \rightarrow z \;\middle|\; \begin{array}{l} z \in I - X \\ interesting(X \rightarrow z) = true \end{array} \right\}$$

---

[3] To extend the notations to non-boolean attributes, $not(y)$ refers to any item coming from the same attribute as y but involving a different attribute value. For example, if y is the item *eye_color=blue* then $not(y)$ can be *eye_color=brown* or *eye_color=green*.

$$same\_consequent(X \to y) = \left\{ z \to y \,\middle|\, \begin{array}{l} z \in I - y \\ interesting(z \to y) = true \end{array} \right\}$$

$$same_i tems(X \to y) = \left\{ (X \cup y) - z \to z \,\middle|\, \begin{array}{l} z \in X \cup y \\ interesting((X \cup y) - z \to z) = true \end{array} \right\}$$

The relations *same_antecedent* and *same_consequent* preserve the antecedent and change the consequent, or vice versa. The relation *same_items* allows to reorder the items in a rule. All the rules produced by this relation concern the same population of records in the database.

6.2. Quality-oriented visualization

Each subset of rules is visualized in a 3D space, which we call a *world*. The representation is built in the following way: each rule is symbolized by an object composed of a sphere perched on top of a cone. Three straightforward graphical characteristics are thus obtained to represent the interestingness measures: the sphere diameter, the cone height, and the color. The representation size depends only on the number of rules in the subset and not on the amount of items. In order to facilitate the navigation (viewpoint control) inside the world, a ground and a sky are represented. As pointed out by Chen (2004), such visual landmarks make the navigation task easier by facilitating the acquisition of spatial knowledge, and more generally by facilitating the building of the cognitive map by the user (mental model of the world).

In a visual representation, the perceptually dominant information is the spatial position (Card et al., 1999). Therefore, in order to be emphasized, the interestingness measures which are fundamental for decision-making are represented by the object position in the world. Since several rules can present the same interestingness, only two measures can be symbolized by spatial position, so that the third dimension is free for scattering the objects. All things considered, we have chosen to use only one axis to place the objects in space and so to spatially represent only one interestingness measure. Indeed, the objects are laid out in the 3D world on an arena (a transparent half-bowl), which means that the further an object is, the higher it is placed (figure 7). This arena allows a better perception of the depth dimension and reduces occultation of objects by other objects. It can hold at most around 250 objects. A similar choice is made in the document manager Data Mountain of Microsoft Research, where web pages are laid out on an inclined plane (Robertson et al, 1998).

Weighing it all up, we have opted for the following visual metaphor to represent each subset of rules by highlighting the interestingness measures (figure 8):

− the object position represents the entropic implication intensity,
− the sphere visible area represents the support,
− the cone height represents the confidence,
− the object color is used redundantly to represent a weighted average of confidence and entropic implication intensity, which gives a synthetic idea of the rule interestingness.

This visual metaphor facilitates comparisons among the rules. It stresses the good rules whose visualization and access are made easier compared to the less good
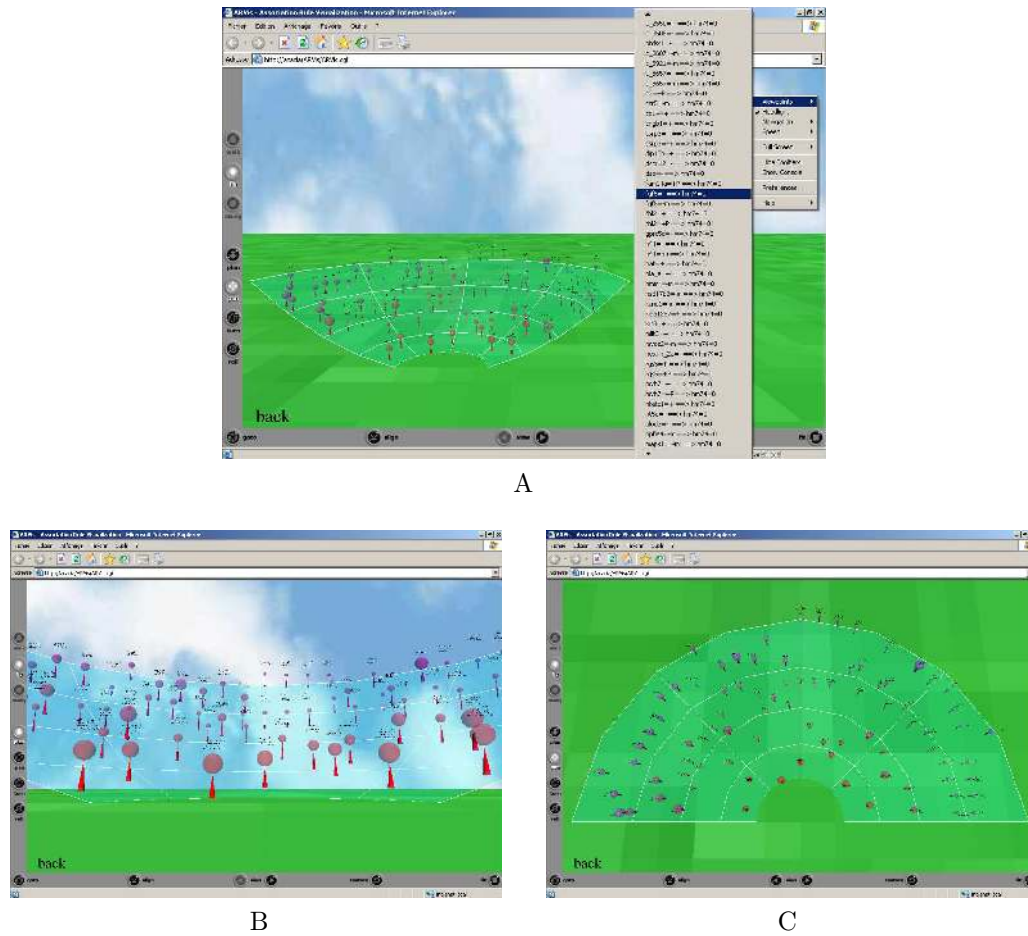
A

B   C

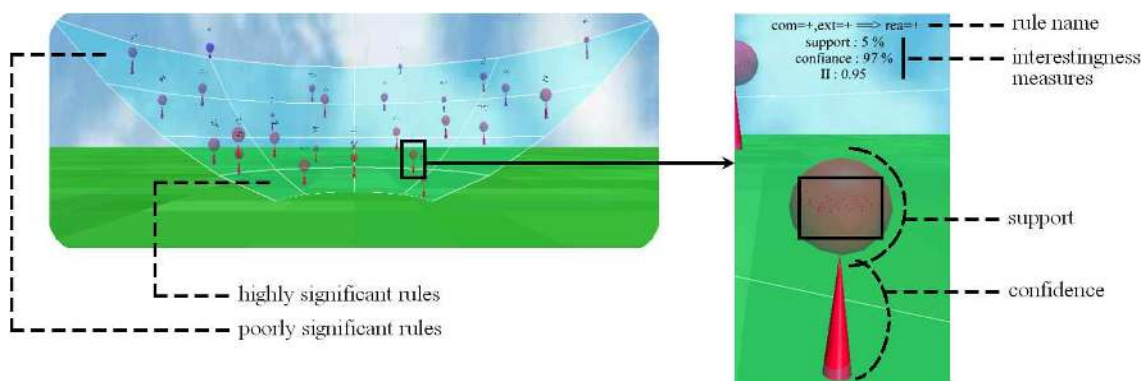**Fig. 7.** Each subset of rules is represented by a 3D world
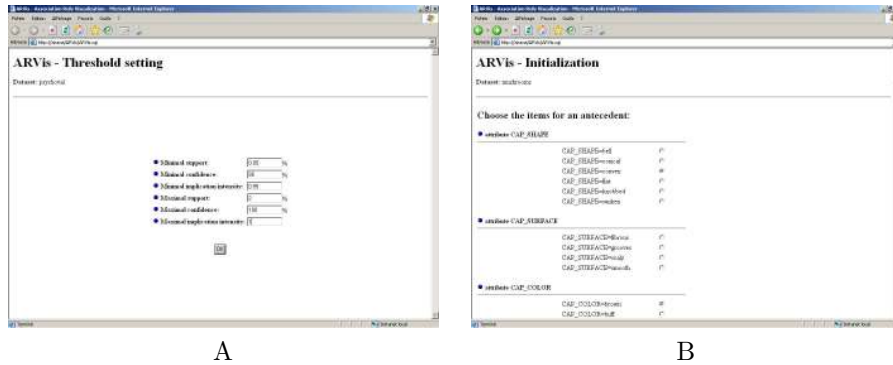

Fig. 8. Visual metaphor

Fig. 9. Exploration initialization interface

rules. More precisely, a large red sphere perched on top of a tall cone placed at the front of an arena, on the lower steps, represents a rule whose support, confidence and entropic implication intensity are high. On the other hand, a little blue sphere perched on top of a small cone placed at the back, on the upper steps, represents a rule whose three measures are weak. This metaphor is a choice among the many possible combinations. It can be adapted by the user. One can choose for instance to change the mappings, or to represent more interestingness measures with color or by using two axes for the spatial mapping.

Furthermore, some complementary text labels appear above each object to give the name of the corresponding rule. They provide the numerical values for support, confidence and entropic implication intensity (noted "EII") too and thus complete the qualitative information given by the representation.

6.3. Interactions with the user

The user can interact in three different ways with the visual representation: by visiting a subset of rules, by filtering the rules on the interestingness measures in a subset, and by navigating among the subsets to discover new rules.

For each subset of rules, at the beginning of a visit, the user is placed in the 3D world in front of the arena so that (s)he benefits from an overall and synthetic view of the rules. With this comprehensive vision, it is easy to locate the best rules. Then the user can wander freely over the world to browse the rules, and zoom in on them to examine them more closely. (S)he just has to click on an object to move in front of it. In each 3D there also exist predefined viewpoints providing overall visions of the arena (cf. for instance, the viewpoint from the top in figure 7.C). If the user looks for a rule with particular items in it, (s)he can search it in a menu (figure 7.A) which lists all the rule names of the subset and allows to move directly in front of the object wanted. In the final analysis, *ARVis* enables the user to find the rules that interest him/her in a subset whether his/her search criteria are based on interestingness measures or on items.

At any time during the visit of a subset, the user can filter the rules on the interestingness measures by adjusting the thresholds min_{sp}, min_{cf}, min_{eii}, max_{sp}, max_{cf}, and max_{eii} (figure 9.A). Only the rules whose measures respect
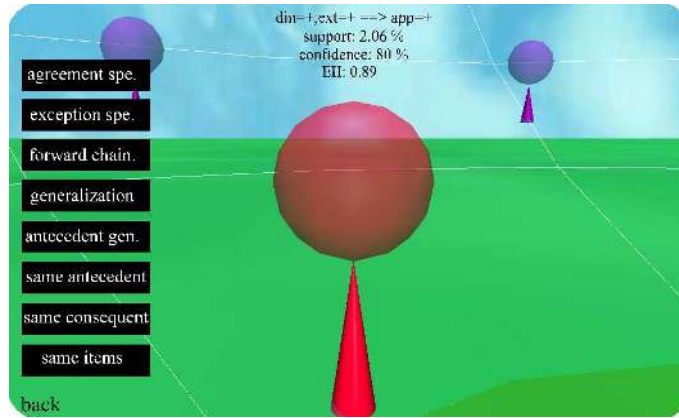
Fig. 10. For each rule, a menu can be displayed to choose the neighborhood relation to be triggered

the thresholds are represented. This makes objects appear or disappear in the world.

Finally, the user can navigate among the subsets of rules via a menu providing the eight neighborhood relations (figure 10). By applying a neighborhood relation on a rule, the current subset is replaced by a new subset. Visually, the current world is replaced by a new world, which gives the impression of virtually moving inside the whole set of rules. At any time during the navigation, the user can go back to the previous subsets (and worlds) with the "back" operator.

Let us assume that the user applies a neighborhood relation Π on a rule r. This generates a new subset $S = \Pi(r)$ containing all the rules neighbors of r according to Π. In *ARVis*, we systematically add the transitional rule r into the new subset S. Visually, the transitional rule r can be easily located in the world since its object flashes. This enables the user to make comparisons between the transitional rule r and its neighbor rules. For instance, with the neighborhood relation *agreement_specialization*, it is interesting to compare a rule r to its neighbors in order to see whether or not the addition of a new item in r tends to improve the rule interestingness. Reciprocally, with the relation *generalization*, comparing a rule r to its neighbors allows to detect the superfluous items in r (those whose removal does not reduce the quality of the rule).

To start or restart the navigation among the subsets, the user can choose the first subset to focus on with an exploration initialization interface (figure 9.B). This interface is an "itemset browser" working with inclusive templates: it enables to build the itemset of one's choosing and then to display the world of the rules that include this itemset in the antecedent, or in the consequent, or in both. Furthermore, the exploration initialization interface allows to choose the database and the table to be studied, and to choose the set I of the items to be used during the exploration.

6.4. *ARVis* implementation

6.4.1. Constraint-based rule-mining algorithm

When the user triggers a neighborhood relation, *ARVis* runs a constraint-based algorithm which dynamically computes the appropriate subset of rules with the interestingness measures. As seen in section 6.1, each of the eight neighborhood relations induces two kinds of constraints: syntactic constraints and interestingness measure constraints. These constraints are "pushed" into association rule mining to reduce the exponentially growing search space. The general structure of the constraint-based algorithm is given below.

```
(1)  Procedure LocalMining
(2)  Input:   rule                   //the transitional rule
(3)           Π                      //the neighborhood relation
(4)           thresholds             //the 6 thresholds on interestingness measures
(5)           database               //connection to the database
(6)  Output: subset_{rules,measures}  //subset of rules with interestingness measures

(7)       subset_{rules} = ∅          //subset of rules (without measures)
(8)       cardinalities = ∅           //cardinalities of the itemsets

(9)       //STEP 1: Generate the candidate rules with the syntactic constraints
(10)      subset_{rules} = SyntacticGeneration(rule,Π)

(11)      //STEP 2: Count the itemsets of the candidate rules (database scan)
(12)      cardinalities = RetrieveCardinalities(subset_{rules},database)

(13)      //STEP 3: Calculate the interestingness measures
(14)      subset_{rules,measures} = CalculateMeasures(subset_{rules},cardinalities)

(15)      //STEP 4: Eliminate the candidate rules which do not respect
(16)      //the interestingness measure constraints
(17)      subset_{rules,measures} = Filter(subset_{rules,measures},thresholds)
(18)      return(subset_{rules,measures})
```

Only step 1 depends on the neighborhood relation Π chosen. This step needs no database scan since it deals only with the syntax of the rules. The syntactic constraints induced by the neighborhood relations of *ARVis* are powerful constraints which drastically reduce the number of rules to be produced. Effectively, the syntactic constraints are verified by at most $|I|$ rules, whatever the relation chosen. It is easy to enumerate these candidate rules and therefore to enumerate all the itemsets that must be counted in the database during step 2 (Ng et al (1998) pointed out the interest of such itemset enumeration procedures in constraint-based itemset-mining). Thus, whatever the neighborhood relation chosen, the whole constraint-based algorithm has a polynomial complexity[4] in $O(|I|)$.

In this polynomial algorithm, the most time-consuming step is step 2. It consists in counting the cardinalities of the itemsets by scanning the database. To improve the response times of the algorithm, a *progressive save system* is

---

[4] Except for the neighborhood relation *exception_specialization* for which the number of rules in a subset is bounded by $m.|I|$ and the complexity is polynomial in $O(m.|I|)$ where $m < |I|$ is the maximum of values for the attributes.
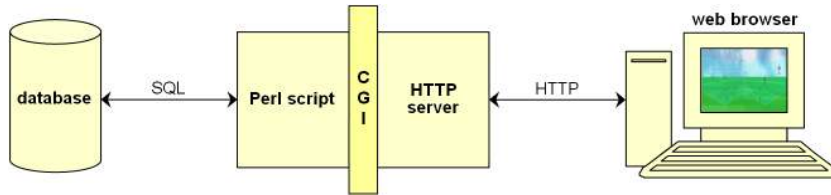
Fig. 11. *ARVis* architecture

implemented in the procedure RetrieveCardinalities: each time an itemset is counted, its cardinality is saved to avoid counting it another time during the remainder of the exploration. In this way, the greater number of times the algorithm is run over the same database, the more the itemset cardinalities are saved, and the more probable it is that the algorithm runs faster. The cardinalities of the itemsets are saved in the database in specific tables. There is one table for each itemset length (number of items in the itemset), so that all the itemsets of the same length are saved in the same table. For each table, the retrieval of the cardinalities uses a B-tree index.

Furthermore, for most of the neighborhood relations of *ARVis*, the constraint-based algorithm can be optimized by "pushing" the interestingness measure constraints into step 1. With the progressive save system, one can indeed anticipate that some candidate rules do not respect the thresholds. Eliminating these candidate rules allows to reduce the number of itemsets that must be counted in the database during step 2. For example, with the neighborhood relation *same_consequent*, it is useless to generate a candidate rule *antecedent* $\rightarrow$ *consequent* if the cardinality of *antecedent* has already been counted and does not respect $cardinality \geq n*min_{sp}/max_{cf}$ and $cardinality \leq n*max_{sp}/min_{cf}$ where n is the number of transactions in the database.

6.4.2. Architecture

ARVis is built on a client/server architecture with a thin client (figure 11). The main block is a CGI program in Perl divided into two parts:

– the constraint-based algorithm which dynamically extracts the subset of rules with their interestingness measures from the database,
– a procedure which dynamically generates the corresponding 3D world in VRML (this procedure is not time-consuming since no database access is needed).

The user visits the worlds with a web browser equipped with a VRML plug-in. The exploration initialization interface is a series of web pages generated by the CGI program.

6.4.3. Response time

The figure 12 shows the response times obtained on three datasets (presented table 1) by executing an exploration scenario with *ARVis*, i.e., a series of neighborhood relations. For each relation triggered by the user, the response time is the time required by *ARVis* to compute the subset of rules with the constraint-based algorithm and to display the corresponding world on the screen. The minimum and maximum thresholds chosen in the scenarios are given in table 2. For the

	# of items	# of transactions	average transaction length
MUSHROOMS	119	8416	23
T10.I4.D100k	100	100000	10
T20.I6.D100k	40	100000	20

Table 1. Dataset characteristics

	MUSHROOMS	**T10.I4.D100k**	**T20.I6.D100k**
min_{sp}	1%	0.05%	0.05%
max_{sp}	100%	100%	100%
min_{cf}	70%	70%	70%
max_{cf}	100%	100%	100%
min_{eii}	0.5	0	0
max_{eii}	1	1	1

Table 2. Minimum and maximum thresholds used in the exploration scenarios

experiments, the server of *ARVis* was running on an SGI Origin 2000 server equipped with four 250 MHz RISC R10000 processors and with 512 MB of memory. The DBMS was PostgreSQL. The tables storing the itemset cardinalities were empty at the beginning of the scenarios.

The first dataset is the MUSHROOMS dataset from the UCI Repository (Blake and Merz, 1998). It is small but it is known to be highly correlated. The exploration scenario that was used with this dataset is given in table 3. The two other datasets are two large synthetic ones: T10.I4.D100k and T20.I6.D100k. They were generated by the procedure proposed by Agrawal and Srikant (1994) (the number of patterns was set to 1000). The dataset T20.I6.D100k is deliberately very dense (on average, each transaction contains 43 % of the items). The exploration scenarios for these two datasets are similar to the one given in the table 3 but we do not present them since the data have no real meaning.

As can be seen on figure 12, the response times tend to decrease as the scenario unfolds. This is due to the progressive save system of the constraint-based algorithm of *ARVis*. The peaks in the response time curve (for example for t=6 and t=11 in the MUSHROOMS scenario) appear when the algorithm needs lots of itemsets that have not been counted yet. In this case, like in any classical procedure for frequent itemset mining, the algorithm has to scan the database, which is time-consuming.

The experiment on the dataset T20.I6.D100k shows that *ARVis* can efficiently mine dense data. In particular, during this experiment, very specific rules containing up to 15 items and presenting a support of 0.07 % have been computed. With a levelwise exhaustive algorithm, such specific rules could never be extracted from a dense database.

Time	Neighborhood relation	Transitional rule (on which the neighborhood relation is applied)	Number of rules generated
t=1	$forward\_chaining$	$CLASS = edible \rightarrow GILL\_SIZE = broad$	3
t=2	$forward\_chaining$	$CLASS = edible, GILL\_SIZE = broad \rightarrow ODOR = none$	4
t=3	$same\_items$	$CLASS = edible, GILL\_SIZE = broad, ODOR = none \rightarrow STALK\_SHAPE = tapering$	3
t=4	$antecedent\_generalization$	$GILL\_SIZE = broad, ODOR = none, STALK\_SHAPE = tapering \rightarrow CLASS = edible$	3
t=5	$same\_antecedent$	$GILL\_SIZE = broad, STALK\_SHAPE = tapering \rightarrow ODOR = none$	7
t=6	$same\_consequent$	$GILL\_SIZE = broad, STALK\_SHAPE = tapering \rightarrow RING\_NUMBER = one$	54
t=7	$forward\_chaining$	$CLASS = edible, GILL\_SIZE = broad, STALK\_SHAPE = tapering \rightarrow ODOR = none$	10
t=8	$back + antecedent\_generalization$	$CLASS = edible, GILL\_SIZE = broad, STALK\_SHAPE = tapering \rightarrow ODOR = none$	3
t=9	$min_{cf} = 60\% + antecedent\_generalization$	$GILL\_SIZE = broad, STALK\_SHAPE = tapering \rightarrow CLASS = edible$	2
t=10	$forward\_chaining$	$STALK\_SHAPE = tapering \rightarrow GILL\_SIZE = broad$	8
t=11	$exception\_specialization$	$GILL\_SIZE = broad, STALK\_SHAPE = tapering \rightarrow CLASS = edible$	4
t=12	$forward\_chaining$	$GILL\_SIZE = broad \rightarrow CLASS = edible$	3
t=13	$forward\_chaining$	$CLASS = edible, GILL\_SIZE = broad \rightarrow ODOR = none$	4
t=14	$forward\_chaining$	$CLASS = edible, GILL\_SIZE = broad, ODOR = none \rightarrow STALK\_SHAPE = tapering$	10

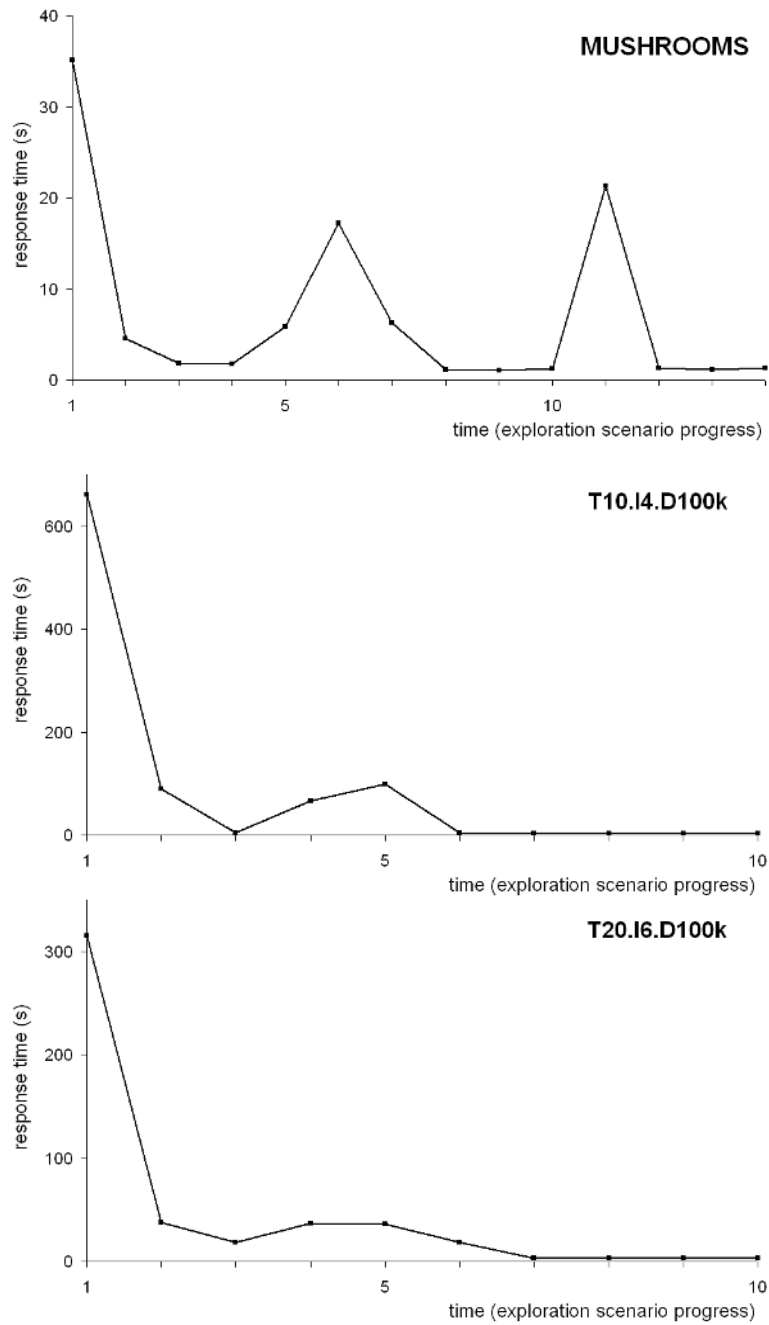**Table 3.** Exploration scenario for the MUSHROOMS dataset

Fig. 12. Response times obtained on three exploration scenarios with $ARVis$

sta	stability	eas	motivation for easiness
fis	fighting spirit	pro	motivation for protection
ext	extroversion	rea	motivation for realization
que	questioning	mem	motivation for membership
wil	willpower	pow	motivation for power
rec	receptiveness	imp	improvisation
rig	rigor	dyn	dynamism
inc	intellectual conformism	com	communication
anx	anxiety	aff	affirmation
soc	spirit of conciliation	ind	independence

Table 4. Meaning of the attributes

7. Example of rule exploration with *ARVis*

The example presented in this section comes from a case study made with the firm PerformanSe SA on human resource management data. The data are a set of workers' psychological profiles used to calibrate decision support systems. It contains around 4000 individuals described by 20 categorical attributes (table 4). Each attribute has three possible values: "+", "0", and "-". In the example, since flashing objects cannot be seen on the figures, a transitional rule is represented in the worlds by an object with a white sphere.

The user begins by studying people that are extrovert and motivated by power. By means of the exploration initialization interface, he displays the world that contains the rules with the itemset $\{ext=+, pow=+\}$ in the antecedent (figure 13.A). The user explores the world. There are three rules with high confidence and high entropic implication intensity at the bottom of the arena, and one of them especially interests the user: $\{ext=+, pow=+\} \rightarrow \{rec=-\}$. To know more characteristics of this not very receptive population, he applies the neighborhood relation *forward_chaining* on this rule. The newly displayed world contains the rules with $\{ext=+, pow=+, rec=-\}$ in the antecedent (figure 13.B). The user finds a rule which he thinks very pertinent: $\{ext=+, pow=+, rec=-\} \rightarrow \{mem=+\}$. To know the other rules verified by these extrovert, not very receptive, and motivated by power and membership people, he applies the neighborhood relation *same_items* on the rule. In the new world, the user sees the four rules that can be built with the four items (figure 13.C). One is the transitional rule, two others are bad rules, and the fourth is a little better than the transitional rule: this is $\{ext=+, mem=+, rec=-\} \rightarrow \{pow=+\}$. To know whether all the items in the antecedent are useful in this rule, he applies the relation *generalization* on it. In the new world (figure 13.D), there is the rule $\{ext=+, mem=+\} \rightarrow \{pow=+\}$ that is as good as the transitional rule, which means that the item $\{rec=-\}$ was superfluous. Next, the user continues his exploration by examining the exceptions of the rule (figure 13.E).

For another exploration, the user is interested in rigorous people. He starts with the world containing the rules with $\{rig=+\}$ in the antecedent (figure 14.A). His attention is drawn by the rule: $\{rig=+\} \rightarrow \{anx=+\}$. This is quite a good rule, but he wants to verify whether other characteristics could better predict strong anxiety. To do so, he applies the neighborhood relation *same_consequent*. The new world contains the rules that conclude on $\{anx=+\}$ and shows that there is no better rule than $\{rig=+\} \rightarrow \{anx=+\}$ (figure 14.B). So the user comes back to the previous world and applies the relation *agreement_specialization* on $\{rig=+\} \rightarrow \{anx=+\}$ to know whether an additional item could improve strong
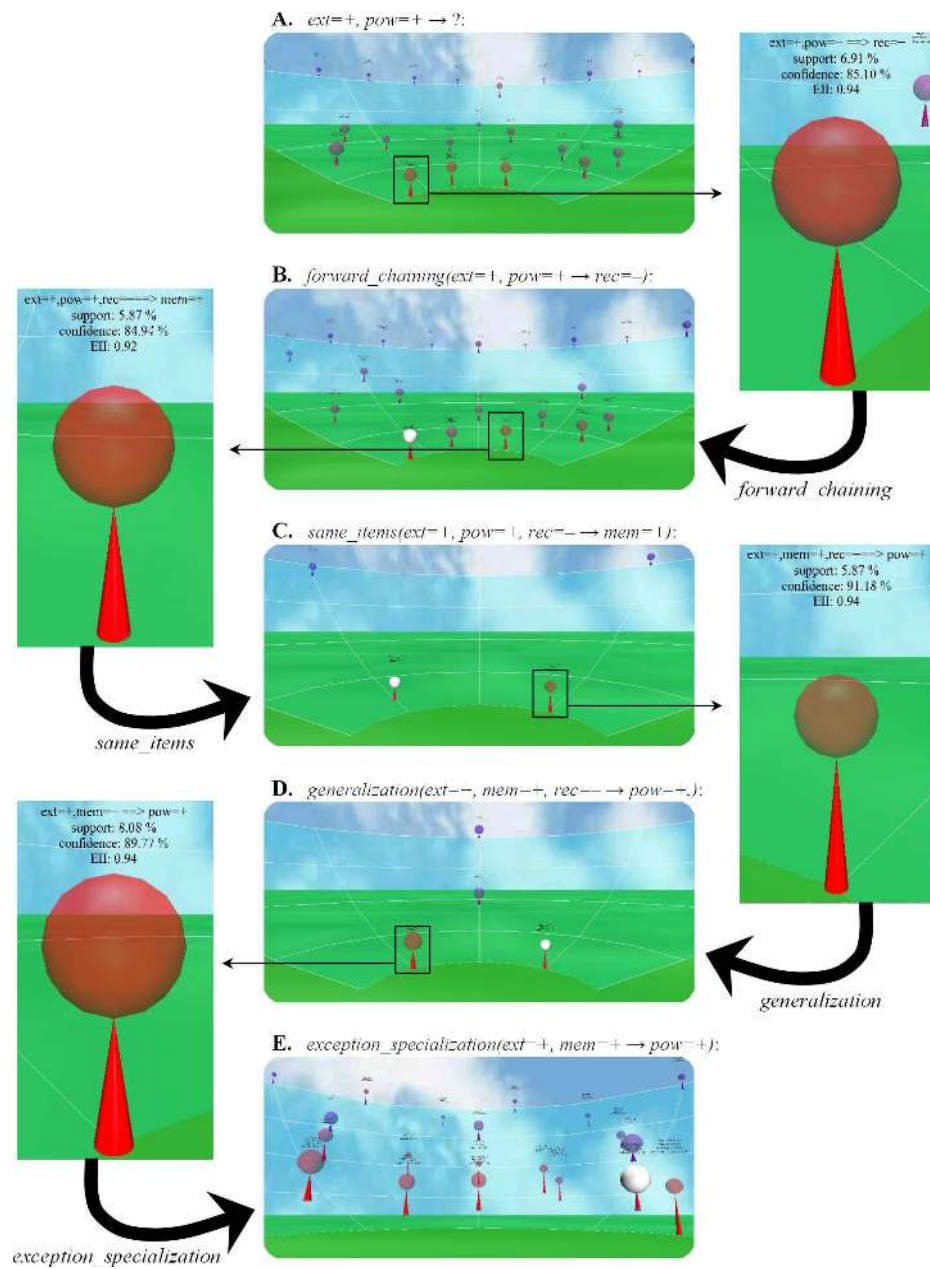
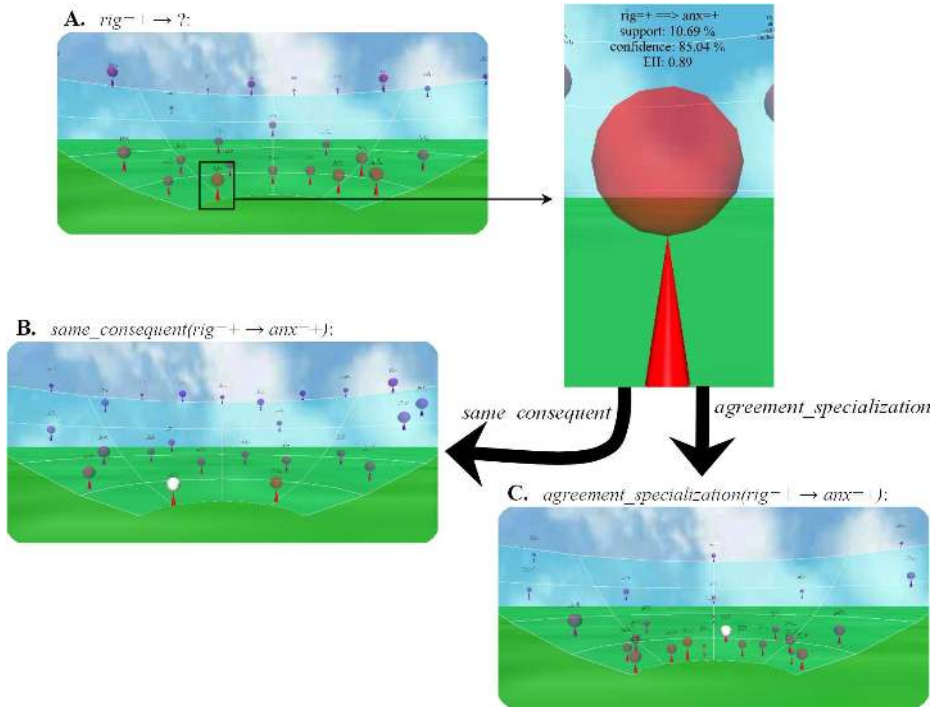**Fig. 13.** An exploration with *ARVis*

Fig. 14. An exploration with *ARVis*

anxiety prediction. The new world presents some better rules effectively (figure 14.C).

8. Conclusion

In this article, we have presented the *rule focusing* methodology for the post-processing of association rules. It enables the user to explore large sets of rules freely by focusing his/her attention on interesting subsets. The methodology relies on:

– rule interestingness measures which filter and sort the rules,
– a visual representation which speeds up comprehension and makes the comparisons among the rules easier,
– several *neighborhood relations* which connect the rules among them and underlie the interactive navigation among the rules.

We have also presented the prototype system *ARVis* which implements the rule focusing methodology by means of a 3D representation, of neighborhood relations meaningful for the user, and of a specific constraint-based rule-mining algorithm. *ARVis* takes advantage not only of the rule syntax, used in the neighborhood relations, but also of the interestingness measures, highlighted in the

representation. This dual approach is original compared to the other rule visualization methods. Moreover, *ARVis* generates the rules dynamically along the exploration by the user. Thus, the user's guidance during association rule post-processing is also exploited during association rule mining to reduce the search space and avoid generating huge amounts of rules.

The experiments we made with *ARVis* have pointed out that the tool can, on the one hand, strengthen the user in certain hypotheses and, on the other hand, provide the user with new ideas. In particular, lots of unknown rules that the user meets along the exploration arouse his/her curiosity and influence the rest of the navigation. Our future works will mainly consist in developing additional neighborhood relations among the rules. For example, a project we have with the French medical research center INSERM on cardiac pathology data requires neighborhood relations which rely on item hierarchies and generate rules with multi-item consequents. Moreover, we think that the analysis of the exploration logs of *ARVis* should reveal some "patterns of navigation" useful to create new neighborhood relations.

References

Aggarwal C.C. (2002) Towards effective and interpretable data mining by visual interaction. SIGKDD Explorations, ACM Press, vol. 3, num. 2, pp 11–22

Agrawal R., Imielinski T., and Swami A. (1993) Mining association rules between sets of items in large databases. In Proc. of the 1993 ACM SIGMOD international conference on management of data, ACM Press, pp 207–216

Agrawal R. and Srikant R. (1994) Fast algorithms for mining association rules. In Proc. of the 20th international conference on very large data bases (VLDB), Morgan Kaufmann, pp 487–499

Agrawal R., Arning A., Bollinger T., Mehta M., Shafer J., and Srikant R. (1996) The Quest data mining system. In Proc. of the 2nd ACM SIGKDD international conference on knowledge discovery and data mining, AAAI Press, pp 244-249, www.almaden.ibm.com/software/quest/

Ammoura A., Zaiane O.R., and Ji Y. (2001) Immersed Visual Data Mining: walking the walk. In BNCOD 18: Proc. of the 18th British National Conference on Databases, Springer-Verlag, pp 202–218

Andrews K. (1995) Visualising cyberspace: information visualisation in the Harmony internet browser. In Proc. of the 1995 IEEE symposium on Information Visualization, IEEE Computer Society, pp 97–104

Baird J. C. (1970) Psychophysical Analysis of Visual Space. Pergamon Press

Barthelemy J.-P. and Mullet E. (1992) A model of selection by aspects. Acta Psychologica, Elsevier Science Publishers, vol. 79, num. 1, pp 1–19

Bayardo R.J. Jr. and Agrawal R. (1999) Mining the most interesting rules. In Proc. of the 5th ACM SIGKDD international conference on knowledge discovery and data mining, ACM Press, pp 145–154

Bertin J. (1967) Smiologie Graphique. Gauthier-Villars, English translation by Berg W. J. as Semiology of Graphics (1983), University of Wisconsin Press

Bhandari I. (1994) Attribute focusing: machine-assisted knowledge discovery applied to software production process control. Knowledge Acquisition, Academic Press Ltd., vol. 6, num. 3, pp 271–294

Bisdorff R. (editor) (2003) Proc. of the mini-EURO conference on human centered processes HCP'2003. University of Luxembourg

Blake C.L. and Merz C.J. (1998) UCI Repository of machine learning databases. Irvine, CA: University of California, Department of Information and Computer Science. www.ics.uci.edu/~mlearn/MLRepository.html

Blanchard J., Kuntz P., Guillet F., and Gras R. (2003) Implication intensity: from the basic statistical definition to the entropic version. In Bozdogan H. (editor.), Statistical data mining and knowledge discovery, Chapman & Hall/CRC Press, pp 473–485

Blanchard J. (2005) A visualization system for interactive mining, assessment, and exploration of association rules. PhD thesis, University of Nantes (in French)

Blanchard J., Guillet F., Briand H., and Gras R. (2005) Assessing rule interestingness with a probabilistic measure of deviation from equilibrium. In Proc. of the 11th international symposium on applied stochastic models and data analysis ASMDA-2005, ENST, pp 191–200

Blanchard J., Guillet F., Briand H., and Gras R. (2005) Using information-theoretic measures to assess association rule interestingness. In Proc. of the 5th IEEE international conference on data mining ICDM'05, IEEE Computer Society, pp 66–73

Bonchi F., Giannotti F., Mazzanti A., Pedreschi D. (2005) Efficient breadth-first mining of frequent pattern with monotone constraints. Knowledge and Information Systems, Springer-Verlag, vol. 8, num. 2, pp 131–153

Botta M., Boulicaut J.F., Masson C., and Meo R. (2002) A comparison between query languages for the extraction of association rules. In Proc. of the 4th international conference on data warehousing and knowlege discovery (DaWaK 2002), Springer-Verlag, Lecture Notes in Computer Science 2454, pp 1–10

Brachman, J.R., and Anand, T. (1996) The process of knowledge discovery in databases: a human-centered approach. In Fayyad U.M., Piatetsky-Shapiro G., Smyth P., Uthurusamy R. (editors), Advances in knowledge discovery and data mining, AAAI/MIT Press, pp 37–58

Braga D., Campi A., Klemettinen M., and Lanzi P.L. (2002) Mining association rules from XML Data. In Proc. of the 4th international conference on data warehousing and knowlege discovery (DaWaK 2002), Springer-Verlag, Lecture Notes in Computer Science 2454, pp 21–30

Brin S., Motwani R., Ullman J.D., and Tsur S. (1997) Dynamic itemset counting and implication rules for market basket data. SIGMOD Record, ACM Press, vol. 26, num. 2, pp 255–264

Brunk C., Kelly J., and Kohavi R. (1997) MineSet: An integrated system for data mining. In Proc. of the 3rd ACM SIGKDD international conference on knowledge discovery and data mining, AAAI Press, pp 135-138

Card S.K., Mackinlay J.D., and Schneiderman B. (editors) (1999) Readings in Information Visualization: Using vision to think. Morgan Kaufmann

Carswell C.M., Frankenberger S., and Bernhard D. (1991) Graphing in depth: perspectives on the use of three-dimensional graphs to represent lower-dimensional data. Behaviour and Information Technology, vol. 10, num. 6, pp 459–474

Chen C. (2004) Information Visualization: beyond the horizon. Springer-Verlag

Cleveland W.S. and McGill R. (1984) Graphical perception: theory, experimentation, and application to the development of graphical methods. Journal of the American Statistical Association, vol. 79, num. 387, pp 531–554

Cockburn A. and McKenzie B. (2001) 3D or not 3D? Evaluating the effect of the third dimension in a document management system. In CHI'01: Proc. of the SIGCHI conference on Human factors in computing systems, ACM Press, pp 434–441

Fayyad U.M., Piatetsky-Shapiro G., and Smyth P. (1996) From data mining to knowledge discovery: an overview. In Fayyad U.M., Piatetsky-Shapiro G., Smyth P., Uthurusamy R. (editors.), Advances in knowledge discovery and data mining, AAAI/MIT Press, pp 1–34

Fayyad U.M., Grinstein G.G., and Wierse A. (2001) Information visualization in data mining and knowledge discovery. Morgan Kaufmann

Freitas A.A. (1998) On objective measures of rule surprisingness. In Proc. of the 2nd European conference on principles of data mining and knowledge discovery (PKDD'98), Springer-Verlag, L.N.A.I. 1510, pp 1–9

Fukuda T., Morimoto Y., Morishita S., and Tokuyama T. (2001) Data mining with optimized two-dimensional association rules. ACM Transactions on Database Systems, ACM Press, vol. 26, num. 2, pp 179–213

Fule P. and Roddick J. F. (2004) Experiences in building a tool for navigating association rule result sets. In Hogan J., Montague P., Purvis M., Steketee C. (editors), CRPIT'04: Proc. of the second Australasian workshop on data mining and web intelligence, Australian Computer Society Inc., pp 103–108

Goethals B. and Van den Bussche J. (2000) On Supporting interactive association rule mining. In Proc. of the 2nd international conference on data warehousing and knowledge discovery (DaWaK), Springer-Verlag, L.N.C.S. 1874, pp 307–316

Grahne G., Lakshmanan L.V.S., and Wang X. (2000) Efficient mining of constrained correlated

sets. In Proc. of the sixteenth international conference on data engineering (ICDE), IEEE Computer Society, pp 512–521

Gras R. (1996) L'implication statistique : nouvelle mthode exploratoire de donnes. La Pense Sauvage Editions (in French)

Guillaume S., Guillet F., and Philippe J. (1998) Improving the discovery of association rules with intensity of implication. In Proc. of the 2nd European conference on principles of data mining and knowledge discovery (PKDD'98), Springer-Verlag, L.N.A.I. 1510, pp 318–327

Han J., Fu Y., Wang W., Koperski K., and Zaiane O. (1996) DMQL: a data mining query language for relational databases. In Proc. of the 1996 SIGMOD workshop on research issues on data mining and knowledge discovery (DMKD)

Han J., Chiang J.Y., Chee S., Chen J., Chen Q., Cheng S., Gong W., Kamber M., Koperski K., Liu G., Lu Y., Stefanovic N., Winstone L., Xia B., Zaiane O.R., Zhang S., and Zhu H. (1997) DBMiner: A system for data mining in relational databases and data warehouses In Proc. of CASCON'97: Meeting of Minds, pp 249–260

Han J., Pei J., and Yin Y. (2000) Mining frequent patterns without candidate generation. In Proc. of the ACM SIGMOD international conference on management of data, ACM Press, pp 1–12

Han J., An A., and Cercone N. (2000) CViz: an interactive visualization system for rule induction. In AI'00: Proc. of the 13th Biennial Conference of the Canadian Society on Computational Studies of Intelligence, Springer-Verlag, pp 214–226

Han J., Hu X., and Cercone N. (2003) A visualization model of interactive knowledge discovery systems and its implementations. Information Visualization, vol. 2, num. 2, Palgrave Macmillan, pp 105–125

Hao M.C., Dayal U., Hsu M., Sprenger T., and Gross M.H. (2001) Visualization of directed associations in e-commerce transaction data. In Proc. of VisSym 2001, pp 185–192

Hipp J., Gntzer U., and Nakhaeizadeh G. (2000) Algorithms for association rule mining - A general survey and comparison. SIGKDD Explorations, ACM Press, vol. 2, num. 1, pp 58–64

Hipp J. and Gntzer U. (2002) Is pushing constraints deeply into the mining algorithms really what we want? An alternative approach for association rule mining. SIGKDD Explorations, ACM Press, vol. 4, num. 1, pp 50–55

Hofmann H., Siebes A.P., and Wilhelm A.F. (2000) Visualizing association rules with interactive mosaic plots. In Proc. of the 6th ACM SIGKDD international conference on knowledge discovery and data mining, ACM Press, pp 227–235

Hofmann H. and Wilhelm A. (2001) Visual comparison of association rules. Computational Statistics, Physica-Verlag, vol. 16, num. 3, pp 399–415

Holland J.H. , Holyoak K.J., Nisbett R.E., and Thagard P.R. (1986) Induction : Processes of inference, learning and discovery. MIT Press

Hussain F., Liu H., Suzuki E., and Lu H. (2000) Exception rule mining with a relative interestingness measure. In Proc. of the 4th Pacific-Asia conference on knowledge discovery and data mining (PAKDD2000), Springer-Verlag, Lecture Notes in Computer Science 1805, pp 86–97

IBM (2006) DB2 Intelligent Miner Visualization. www.ibm.com/software/data/iminer/visualization/index.html

Imielinski T. and Mannila H. (1996) A database perspective on knowledge discovery. Communications of the ACM, ACM Press, vol. 39, num. 11, pp 58–64

Imielinski T. and Virmani A. (1999) MSQL: A query language for database mining. Data Mining and Knowledge Discovery, Kluwer Academic Publishers, vol. 3, num. 4, pp 373–408

Jeudy B. and Boulicaut J.-F. (2002) Optimization of association rule mining queries. Intelligent Data Analysis, IOS Press, vol. 6, num 4, pp 341–357

Keim D.A. (2002) Information visualization and visual data mining. IEEE Transactions on Visualization and Computer Graphics, IEEE Educational Activities Department, vol. 8, num. 1, pp 1–8

Klemettinen M., Mannila H., Ronkainen P., Toivonen H., and Verkamo A.I. (1994) Finding interesting rules from large sets of discovered association rules. In Proc. of the 3rd international conference on information and knowledge management (CIKM), ACM Press, pp 401–407

Kopanakis I. and Theodoulidis B. (2001) Visual data mining and modeling techniques. In Proc. of the KDD-2001 workshop on visual data mining

Kuntz P., Guillet F., Lehn R., and Briand H. (2000) A user-driven process for mining association rules. In Proc. of the 4th European conference on principles of data mining and knowledge discovery (PKDD-2000), Springer-Verlag, pp 483–489

Liu B., Hsu W., Wang K., and Chen S. (1999) Visually aided exploration of interesting association rules. In Proc. of the 3rd Pacific-Asia conference on knowledge discovery and data mining (PAKDD1999), Springer-Verlag, Lectures Notes in Artificial Intelligence 1574, pp 380–389

Liu B., Hsu W., Chen S., and Ma Y. (2000) Analyzing the subjective interestingness of association rules. IEEE Intelligent Systems, IEEE Educational Activities Department, vol. 15, num. 5, pp 47–55

Loevinger J. (1947) A systematic approach to the construction and evaluation of tests of ability. Psychological Monographs, vol. 61, num. 4

Ma Y., Liu B., and Wong C.K. (2000) Web for data mining: organizing and interpreting the discovered rules using the Web. SIGKDD Explorations, ACM Press, vol. 2, num. 1, pp 16–23

McEachren A. M. (1995) How Maps Work: representation, visualization, and design. The Guilford Press

Meo R., Psaila G., and Ceri S. (1998) An extension to SQL for mining association rules. Data Mining and Knowledge Discovery, Kluwer Academic Publishers, vol. 2, num. 2, pp 195–224

Montgomery H. (1983) Decision rules and the search for a dominance structure: towards a process model of decision making. In Humphreys P.C., Svenson O., Vari A. (editors.), Analysing and aiding decision processes, Amsterdam:North Holland, pp 343–369

Ng R.T., Lakshmanan L.V.S., Han J., and Pang A. (1998) Exploratory mining and pruning optimizations of constrained associations rules. In Proc. of the 1998 ACM SIGMOD international conference on management of data, ACM Press, pp 13–24

Ordonez C., Ezquerra N., Santana C.A. (2006) Constraining and summarizing association rules in medical data. Knowledge and Information Systems, Springer-Verlag, vol. 9, num. 3, pp 1–2

Padmanabhan B. and Tuzhilin A. (1999) Unexpectedness as a measure of interestingness in knowledge discovery. Decision Support Systems, Elsevier Science Publishers, vol. 27, num. 3, pp 303–318

Piatetsky-Shapiro G. (1991) Discovery, analysis, and presentation of strong rules. In Piatetsky-Shapiro G., Frawley W.J. (editors.), Knowledge discovery in databases, AAAI/MIT Press, pp 229–248

Rainsford C.P. and Roddick J.F. (2000) Visualisation of temporal interval association rules. In Proc. of the 2nd international conference on intelligent data engineering and automated learning (IDEAL 2000), Springer-Verlag, pp 91–96

Robertson G., Czerwinski M., Larson K., Robbins D.C., Thiel D., and van Dantzich M. (1998) Data mountain: using spatial memory for document management. In UIST'98: Proc. of the 11th annual ACM symposium on user interface software and technology, ACM Press, pp 153–162

SAS (2006) Enterprise Miner. www.sas.com/technologies/analytics/datamining/miner/

Sebag M. and Schoenauer M. (1988) Generation of rules with certainty and confidence factors from incomplete and incoherent learning bases. In Proc. of the European knowledge acquisition workshop EKAW'88, Gesellschaft fr Mathematik und Datenverarbeitung mbH, pp 28.1–28.20

Schneiderman B. (2002) Inventing discovery tools: combining information visualization with data mining. Information Visualization, Palgrave Macmillan, vol. 1, num. 1, pp 5–12

Silberschatz A. and Tuzhilin A. (1996) User-assisted knowledge discovery: how much should the user be involved. In Proc. of the 1996 SIGMOD workshop on research issues on data mining and knowledge discovery (DMKD)

Silberschatz A. and Tuzhilin A. (1996) What makes patterns interesting in knowledge discovery systems. IEEE Transactions on Knowledge and Data Engineering, IEEE Educational Activities Department, vol. 8, num. 6, pp 970–974

Silverstein C., Brin S., and Motwani R. (1998) Beyond market baskets: Generalizing association rules to dependence rules. Data mining and knowledge discovery, Kluwer Academic Publishers, vol. 2, num. 1, pp 39–68

Simon H.A. (1979) Models of Thought. Yale University Press

Spence I. (1990) Visual psychophysics of simple graphical elements. Journal of Experimental Psychology: Human Perception and Performance, vol. 16, num. 4, pp 683–692

Spence R. (2000) Information Visualization. Addison Wesley

Srikant R., Vu Q., and Agrawal R. (1997) Mining association rules with item constraints. In Proc. of the 3rd ACM SIGKDD international conference on knowledge discovery and data mining, AAAI Press, pp 67–73

Suzuki E. (2002) Undirected discovery of interesting exception rules. International Journal of Pattern Recognition and Artificial Intelligence, World Scientific Publishing Company, vol. 16, num. 8, pp 1065–1086

Tan P.-N., Kumar V., and Srivastava J. (2004) Selecting the right objective measure for association analysis. Information Systems, Elsevier Science Publishers, vol. 29, num. 4, pp 293–313

Tufte E. (1983) The Visual Display of Quantitative Information. Graphics Press

Tuzhilin A. and Adomavicius G. (2002) Handling very large numbers of association rules in the analysis of microarray data. In KDD'02: Proc. of the eighth ACM SIGKDD international conference on knowledge discovery and data mining, ACM Press, pp 396–404

Unwin A. R., Hofmann H., and Bernt K. (2001) The TwoKey plot for multiple association rules control. In Proc. of 5th European conference on principle and practice of knowledge discovery in databases (PKDD'01), Springer-Verlag, pp 472–483

Ware C. and Franck G. (1996) Evaluating stereo and motion cues for visualizing information nets in three dimensions. ACM Transactions on Graphics, vol. 15, num. 2, pp 121-140

Wilkinson L. (1999) The Grammar of Graphics. Springer-Verlag

Wong P.C., Whitney P., and Thomas J. (1999) Visualizing association rules for text mining. In Proc. of the 1999 IEEE symposium on information visualization, IEEE Computer Society, pp 120–123

Author Biographies

Dr. **Julien Blanchard** earned his Ph.D. in 2005 from Nantes University (France) and is currently an assistant professor at Polytechnic School of Nantes University. He is the author of a book chapter and 7 journal and international conference papers in the areas of visualization and interestingness measures for data mining.

Assistant Prof. **Fabrice Guillet** is currently a member of the LINA laboratory (CNRS 2729) at Polytechnic graduate School of Nantes University (France). He hold a Ph.D. in computer science in 1995 from the Ecole Nationale Supérieure des Télécommunications de Bretagne. He is author of 35 international publications in data mining and knowledge management. He is a founder and a permanent member of the Steering Committee of the annual EGC French-speaking conference.

Prof. **Henri Briand** earned his Ph.D. in 1983 from Paul Sabatier University located in Toulouse (France) and has over 100 publications in database systems and database mining. He was the head of the Computer Engineering Department at Polytechnic School of Nantes University. He was in charge of a research team in the data mining domain. He is responsible for the organization of the Data Mining Master in Nantes University.

Correspondence and offprint requests to: Julien Blanchard, École Polytechnique de l'Université de Nantes, Département Informatique, La Chantrerie, BP 50609, Nantes cedex 3, 44306, France. Email: julien.blanchard@polytech.univ-nantes.fr