Modeling default induction with conceptual structures

VELCIN J. and GANASCIA J.-G.

LIP6, Université Paris VI
8 rue du Capitaine Scott
75015 PARIS
FRANCE
{julien.velcin,jean-gabriel.ganascia}@lip6.fr

Abstract. Our goal is to model the way people induce knowledge from rare and sparse data. This paper describes a theoretical framework for inducing knowledge from these incomplete data described with conceptual graphs. The induction engine is based on a non-supervised algorithm named default clustering which uses the concept of stereotype and the new notion of default subsumption, the latter being inspired by the default logic theory. A validation using artificial data sets and an application concerning an historic case are given at the end of the paper.

1 Introduction

We aim to model the way it is possible to induce knowledge from rare and sparse data using a default reasoning. In this way, we propose a new induction engine using non-supervised learning techniques and the conceptual graph formalism as described by J. Sowa [1]. The induction mechanism is based on the notion of default subsumption, the latter having been inspired from the default logic theory of R. Reiter [2]. This new model has been designed both to deal with heterogeneous and incomplete databases and to understand the way people build stereotypes from incomplete information, as it can be found for example in newspaper articles.

On the one hand, such databases have to be automatically completed with a default reasoning to become comparable. On the other hand, our hypothesis is that popular inductions are not only due to the lack of facts, but also to the poor description of the existing facts. This sparseness is particularly favorable to the use of background knowledge –like theories– and to the elaboration of caricatural representations we called *stereotypes*.

In such cases, we simulate the erection of categories from incomplete information by using machine learning and data mining techniques. These categories are formed with a new relation we introduce, the default subsumption, and named thanks to the concept of stereotype, which is defined below. In the past, some meaningful results have been obtained by using supervised learning techniques and applying them to model pre-scientific reasoning both in the field of medicine and in some cases of dissemination of social misrepresentations [3, 4].

Our paper is divided in two parts. The first section presents the logical framework modeling inductive reasoning from sparse descriptions. This framework makes use both of the notion of default subsumption, which is analogous to default logic, and of the concept of stereotype, which models the way sparse descriptions may be categorized. Details tools and strategies are then detailed in order to build such sets of stereotypes. The second section is dedicated to the validation of the model from artificial data sets and to a real application dealing with social misrepresentations.

2 Logical framework

2.1 Default logic

During the eighties, there were many attempts to model deductive reasoning in presence of implicit informations. A lot of formalisms [5, 6, 2] have been developed to encompass the inherent difficulties of such models, especially their non-monotony: close-world assumption, circumscription, default logic, etc. Since our goal here is to model the way people induce empirical knowledge from partially and non homogeneously described facts, we face a very similar problem: in both cases, it is to reason in presence of implicit information. Therefore, it is natural to make use of similar formalisms.

In this case, we choose the default logic formalism, which were developed in the eighties by R. Reiter [2]. This logic for default reasoning is based on the notion of default rules, which permits to infer new formulas when the hypotheses are not inconsistent.More generally, a default rule has always the following form: $A : B_1, B_2, ...B_n/C$ where A is called the prerequisite, B_i the justifications and C the conclusion. This default rule can be interpreted as follows: if A is known to be true and if it is consistent to assume $B_1, B2, ...B_n$ then conclude C.

For instance, let us consider the next default rule:

`politician(X) ∧ introducedAbroad(X) : ¬ diplomat(X) / traitor(X)`

This rule translates a usual way of reasoning for people living in France during the end of the 19th century; it means that one can suspect all politicians who are introduced abroad to be traitors towards their own countries, except for diplomats. In other words, it expresses that the conclusion traitor(X) can be derived if X is a politician who is known to be introduced abroad while we cannot prove that he is a diplomat.

Let us note that information conveyed by default rules refers to implicit connotations. As example, the antinomy among patriots and internationalists or the rule that assimilates almost all the politicians involved with foreigners to traitors correspond to connotations and may facilitate the completion of partial descriptions. The key idea is that people have in mind stereotypes that correspond to strong images stored in memory and that partial descriptions evoke such stereotypes. The following sections are dedicated to this concept of stereotype; before, it is necessary to introduce the notion of default subsumption.

In the rest of this section, we use the framework designed by Sowa in [1]. A short introduction can be found at [7].

2.2 Default subsumption

First of all, let us assume that a stereotype is a specific description, which will be in this paper a conceptual graph, and consider the description function $\delta : F \rightarrow D$ which associates a conceptual graph $\delta(f) \in D$ to each fact f from the set of initial facts F. Let us next consider that a fact subsumes another fact if it is the result of the generalization operators.

A stereotype stored in the memory is said to subsume a fact by default if it has no contradictory features with the description of the fact considered, i.e. $\delta(f)$. So it can be used to complete this description without adding any incoherence. In other words, the fact can be completed in such a way that its description can now be subsumed by the stereotype.

Let us consider now the graph g associated with a fact in which there is a very large number of missing data. The missing data can be guessed and completed to obtain a more specific graph g_S. We follow here the notations given by Sowa in [1] (definition 3.5.1): $g_S \leq g$ means that g_S is a specialization of g and g is a generalization of g_S, i.e. g subsumes g_S. Now, let s be one stereotype belonging to the structured memory. If this stereotype is more general than g_S, ie $g_S \leq s$, then it subsumes g *by default*. More formally:

Definition 1. *Let f be a fact represented by the conceptual graph $g = \delta(f)$ and s a stereotype. s* subsumes g by default *if and only if there exists a graph g_S with $g_S \leq g$ and $g_S \leq s$. g_S is therefore a graph formed by the join operator performed on the graphs g and s.*

Fig. 1 presents the fact *The meal of Jules is composed of steak, red wine, and ends with a cup of very hot coffee* which can be subsumed by default by the stereotype *The meal is composed of steak with potatoes and French bread, and ends with a cup of coffee* because the fact can be completed to *The meal of Jules is composed of steak with potatoes and French bread, red wine, and ends with a cup of very hot coffee.* If the stereotype had presented a meal ending with a liqueur, it would not match the fact and so could not subsume it by default.

Property 1. The notion of default subsumption is more general than that of classic subsumption. Let g and g' be two conceptual graphs. If g subsumes g' then g subsumes g' by default.

Property 2. The default subsumption is a symmetrical relation. Let u_1 and u_2 be two conceptual graphs. If u_1 subsumes u_2 by default, then u_2 subsumes u_1 by default too.

Let us note that the notion of default subsumption may appear strange for people accustomed to classical subsumption since it is symmetrical. As a consequence, it does not define an ordering relationship on the space of description.

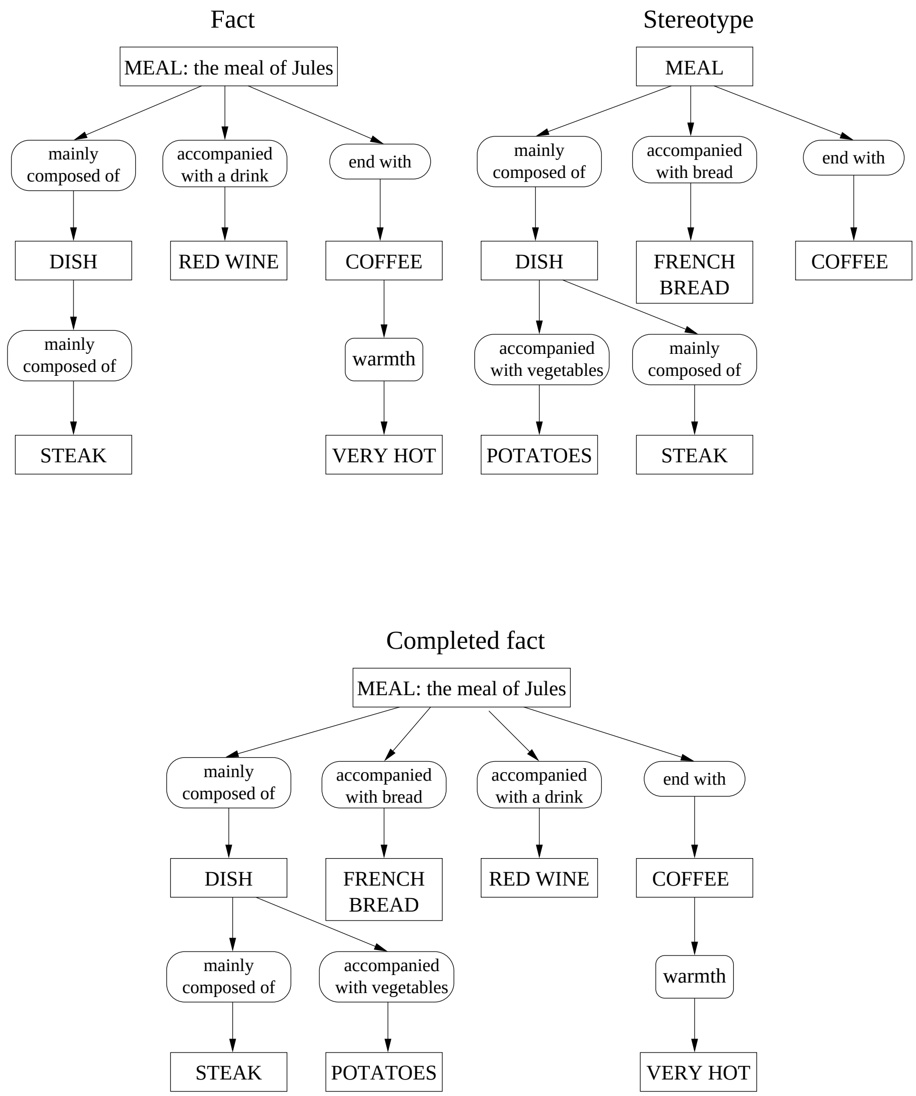

Fig. 1. The stereotype *subsumes by default* the fact description. The description below is the result of the join operator, i.e. the completed fact.

2.3 Concept of stereotype

Eleanor Rosch saw the categorization itself as one of the most important issues in cognitive science [8]. She observed that children learn how to classify first in terms of concrete cases rather than through defining features. She therefore introduced the concept of prototype as the ideal member of a category. Ownership to a class is then defined by the proximity to the prototype and not by the number of shared features.

For example, a robin is closer to the bird prototype than an ostrich, but they are both closer to it than they are to the prototype of a fish, so we call them both birds. However, it takes longer to say an ostrich is a bird than it takes to say a robin is a bird because the ostrich is further from the prototype. Sowa defines a prototype as a typical instance formed by joining one or more schemata. Instead of describing a specific individual, it describes a typical or "average" individual.

From a computational point of view the concept of prototype is difficult to manage since many complete observations have to be considered in order to construct such an ideal fact. Furthermore, it is not really appropriate in order to classify new observations and predict missing information. We therefore propose to adopt the concept of stereotype, which is quite close to that of prototype but more adapted to missing values.

The concept of stereotype was introduced by W. Lippman in a book about public opinion [9]. We define it here as a specific and imaginary fact that combines features found in the facts it subsumes by default. Since there is no contradiction between a fact and its related stereotype, it may be used to complete its description. In other words, a stereotype s is said to cover a fact described by g if and only if s subsumes by default the graph g. As a consequence, g may be completed by the conceptual graph s. This point will be detailed later.

2.4 Set of stereotypes

Groups of people share implicit knowledge, which makes them able to understand each other without having to express everything explicitly. This sort of knowledge can be expressed in terms of erudite theories (e.g. the "blocking perspiration" theory in [3]) or use a more "naive" formulation. Our second hypothesis is that this implicit knowledge can be stored in terms of sets of stereotypes. This means that many people have in mind the sets of stereotypes and that they use them to reason in a stereotyped way by associating new facts to stereotypes they have in mind.

To formalize this idea, let first suppose that a description space D and a set of facts F are given. Then, a measure of dissimilarity M_D is defined on D. Previous work deals with graph matching and an interesting method to calculate the similarity between two conceptual graphs is proposed in [10]. However, in the present context we consider sufficient a simplier measure that adds the differences between graphs.

First let us recall the definition of *compatibility* given in [1]:

Definition 2. *Let conceptual graphs* u_1 *and* u_2 *have a common generalization* v *with projections* $\pi_1 : v \rightarrow u_1$ *and* $\pi_2 : v \rightarrow u_2$*. The two projections are said to be* compatible *if for each concept* c *in* v*, the following conditions are true:*

1. $type(\pi_1 c) \cap type(\pi_2 c) > \perp$.
2. *The referents of* $\pi_1 c$ *and* $\pi_2 c$ *conform to* $type(\pi_1 c) \cap type(\pi_2 c)$.
3. *If* $referent(\pi_1 c)$ *is the individual marker* i*, then* $referent(\pi_2 c)$ *is either* i *or* $*$.

We now consider that there is always only one least common generalization, i.e. only two projections that are compatible and maximally extended. It is easy to generalize our model with graphs having several least common generalizations.

The following theorem is stated in order to link the notions of compatibility and default subsumption:

Theorem 1. *Let conceptual graphs* u_1 *and* u_2 *have the least common generalization* v *with projections* $\pi_1 : v \rightarrow u_1$ *and* $\pi_2 : v \rightarrow u_2$*.* π_1 *and* π_2 *are compatible if and only if* u_1 *subsumes* u_2 *by default.*

Proof. If π_1 and π_2 are compatible then there exists a common specialization w of u_1 and u_2 (cf. theorem 3.5.7). According to definition 1, u_1 subsumes u_2 by default. Reciprocally, if u_1 subsumes u_2 by default then there exists a common specialization w. Suppose that π_1 and π_2 are not compatible. There therefore exists at least one concept in v with $type(\pi_1 c) \cap type(\pi_2 c) = \perp$, or with the referent of $\pi_1 c$ or $\pi_2 c$ not conform to $type(\pi_1 c) \cap type(\pi_2 c)$, or with $referent(\pi_1 c) = i$ and $referent(\pi_2 c) = j$, $i \neq j$. These three cases are absurd because they contradict the construction of w. Therefore, π_1 and π_2 are compatible.

Consider now the measure M_D counting the dissimilarities between two graphs u_1 and u_2. Let v be the least common generalization graph with projections $\pi_1 : v \rightarrow u_1$ and $\pi_2 : v \rightarrow u_2$. If π_1 and π_2 are not compatible then the measure $M_D(u_1, u_2)$ is fixed by convention with an infinite value noted M_∞ because one graph can't be subsumed by default by the second one (cf. theorem 1). Otherwise $M_D(u_1, u_2)$ counts all the differences between the concepts and relations of u_1 and those of u_2. The measure is thus defined:

Definition 3. *Let conceptual graphs* u_1 *and* u_2 *have the least common generalization* v *with projections* $\pi_1 : v \rightarrow u_1$ *and* $\pi_2 : v \rightarrow u_2$*. The* measure of dissimilarities $M_D(u_1, u_2)$ *is equal to:*

1. M_∞ *if* π_1 *and* π_2 *are not compatible.*
2. $C + T(u_1) + T(u_2)$ *otherwise, where:*
   - $C = |\{concept\ c \in v / type(\pi_1 c) \neq type(\pi_2 c)\ or\ referent(\pi_1 c) \neq referent(\pi_2 c)\}|$.
   - $T(u) = card(u) - card(v)$; $card(g)$ *corresponds to the number of nodes (concepts and relations) of graph* g.

This measure presents the following properties:

Property 3. For any conceptual graph u, $M_D(u, u) = 0$.

Property 4. For any conceptuals graphs u and v, $M_D(u, v) = M_D(v, u)$.

Let us now define what is a set of stereotypes :

Definition 4. *In the framework of the conceptual graphs, a set of stereotypes is a tuple of n graphs $(s_1, s_2, ...s_n)$.*

2.5 Completion of facts

Being given a set of facts F and a set of stereotypes $(s_1, s_2, ...s_n)$, it is possible to complete the descriptions of almost any fact f. More precisely, the completion is possible when there exists at least one stereotype s_i belonging to the set of stereotype $(s_1, s_2, ...s_n)$ such that s_i subsumes by default the description $\delta(f)$ of the fact f. In other words, thinking by stereotypes is possible when new descriptions are so sparse that they seem consistent with existing stereotypes. This capacity to classify and to complete the descriptions is characteristic of the concept of stereotype as introduced by Lippman.

When one and only one stereotype s_i covers by default the fact f, the description of f, $\delta(f)$, may be completed by the stereotype s_i. However, it happens that facts may be covered by two or more stereotypes s_i. Then, the stereotype associated with a fact f is the one that minimizes the measure of disimilarity M_D, i.e. it is the stereotype s_i which both covers $\delta(f)$ by default and minimizes $M_D(\delta(f), s_i)$. It is called the relative cover of f, thanks to the measure of dissimilarity M_D and to a set of stereotypes $S = (s_1, s_2, ...s_n)$.

Definition 5. *In a more formal way, the relative cover of a fact f, with respect to a set of stereotype $S = (s_1, s_2, ...s_n)$, noted $C_S(e)$, is the stereotype s_i if and only if:*

1. $s_i \in (s_1, s_2, ...s_n)$,
2. $M_D(\delta(f), s_i) \neq M_\infty$,
3. $\forall k \in [1, n], k \neq i, M_D(\delta(f), s_i) < M_D(\delta(f), s_k)$.

It may also happen that no stereotype covers the new fact f, which means that $\delta(f)$, the description of f, is inconsistent with all s_i. In this case, there may not be any completion and thinking by stereotype is impossible.

2.6 Extraction of stereotypes

Implicit reasoning is formalized here with both the default subsumption and sets of stereotypes which structure our memory. Up to now, these sets of stereotypes were supposed to be given. This section shows how our memories can be organized into sets of stereotypes. In other words, it is to model the way facts aggregate into structures which render implicit reasoning possible. From a technical point of view, this memory organization process can be seen as a non-supervised learning task we call *default clustering*, which can be summarized as follows.

Being given a set of facts F described with conceptual graphs, a non-supervised learning algorithm is supposed to organize the initial set of facts F into a structure, for instance a hierarchy, a lattice or a pyramid. In the present case, we restrain to partitions of the training set, which correspond to sets of stereotypes. Let us recall that $(F_1, F_2, ...Fn)$ constitutes a partition of the set F if and only if:

1. $\forall i \in [1, n], F_i \subset F$
2. $\cup_{i\in[1,n]}, F_i = F$
3. $\forall (i, j) \in [1, n]^2, F_i \cap F_j = \emptyset$

This partition may be generated by n conceptual graphs $\{g_1, g_2, ...g_n\}$: it is sufficient to associate to each g_i the set F_i of facts belonging to F and covered by g_i relative to $(g_1, g_2, ...g_n)$.

To choose among the numerous possible structures, even with simple structures like partitions, a non-supervised algorithm requires a distance. The usual way is to minimize the so-called intra-class distance – i.e. the average distance between examples belonging to the same class – and/or to maximize the inter-class distance – i.e. the sum of distances between pairs of examples belonging to different classes. – The key point is to have a distance among examples of the learning set and to extend it to intra and inter-class distances.

The first distance considered used probabilities. It is very similar to the Category Utility measure [12] which is used in the COBWEB system [13] to evaluate good partitions. But in practice this measure is not really appropriate for sparse descriptions. Moreover, runtime cost was rather high, which made the learning algorithm very inefficient.

Referring to the definition of both "sets of stereotypes" and the relative cover, it appears natural to make use of a distance close to the measure of dissimilarity M_D. This is exactly what we do by introducing a cost function h based on M_D:

Definition 6. *F being a set of facts, the so-called training set, $S = (s_1, s_2, ...s_n)$ a set of stereotypes and C_S the function that associates to each fact f its relative cover, i.e. its closest stereotype with respect to M_D and S, the cost function h is defined as follows:*

$$h(S) = \sum_{f \in F} M_D(\delta(f), C_S(f))$$

Once the cost function h has been defined, the non-supervised learning algorithm has to build the set of stereotypes $(s_1, s_2, ...s_n)$ that minimizes h. In other words, the non-supervised learning problem is reduced to an optimization problem.

There are several methods for exploring such search space. One is incremental and very similar to the one used by Fisher in COBWEB. It starts from an empty set with no stereotypes, considers at each step a new individual to be covered and updates the set of stereotypes with some specific operators. For instance, one of them creates a new stereotype equal to the considered individual; another modifies one existing stereotype to cover it; the “merge” operator merges two stereotypes belonging to the set of stereotypes and the “split” operator splits one active stereotype. The search in COBWEB is a “hill-climbing” strategy; its robustness is largely due to these last two specific operators, the “merge” and the “split”. However, in case of sparse descriptions and especially with graphs, the “merge” and “split” operators cannot be easily implemented. Therefore, it is difficult to apply this algorithm here.

The second option is to search for the best set of stereotypes using general optimization techniques. We chose a “tabu” strategy, which is a classical meta-heuristic technique used in operational research. It seems quite well adapted to solve our problem, as we shall see in the next section. From a technical point of view, a neighborhood is calculated from the current solution with the assistance of permitted movements. These movements can be of low influence (enrich one stereotype with a descriptor, remove a descriptor from another) or of high influence (add or retract one stereotype from the current set of stereotypes). As in almost all local search techniques, there is a trade-off between exploitation, i.e. choosing the best movement, and exploration, i.e. choosing a non optimal state. The search uses short and long-term memory to avoid loops and to intelligently explore the search space. We shall not detail here the “tabu” search algorithm since it is a classical one (see [11]); we shall just evaluate its robustness on artificial data in the next section.

3 Experiments

This section validates our approach in the Attributes/Values formalism, before proposing a real application dealing with a famous French affair translated into conceptual graphs.

3.1 Validation on artificial data sets

Evaluation validates on artificial data sets the robustness of the non-supervised learning algorithm, which builds sets of stereotypes from a learning set F and a description language D. Let us recall that stereotypes are supposed to be more or less shared by many people living in the same society. Since use of stereotypes is the way to model implicit reasoning, it could explain why prejudges and presupposes are almost identical in a group. Our second hypothesis is that people reason from sparse descriptions that they are always able to complete and to organize into a set of stereotypes in their memory. These two hypotheses entail that people, who shared different experiences, and who read different news, are able to build very similar sets of stereotypes from very different learning sets. Therefore, our attempt to model construction of sets of stereotypes with a non-supervised learning algorithm ought to have this stability property. We evaluate it here on artificial data.

Let us now consider the Attributes/Values formalism. Being given this description language, we introduce some full consistent descriptions, e.g. (d_1, d_2, d_3), which stand for the description of a set of stereotypes. Let us note as n_s the number of such descriptions. These n_s descriptions may be randomly generated; the only points are that they need to be full and consistent.

The second step of the artificial set generation is to duplicate these description n_d times, for instance, 50 times, making $n_s \times n_d$ artificial examples. Then, these $n_s \times n_d$ descriptions are arbitrarily degraded: descriptors belonging to these duplications are chosen randomly to be destroyed. The only parameter is the percentage of degradation, i.e. the ratio of the number of destroyed descriptors on the total number of descriptors. The generated learning set contains $n_s \times n_d$ example descriptions, which all correspond to degradations of the n_s initial descriptions.

The default clustering algorithm is tested on these artificially generated and degraded learning sets. Then, the stability property is evaluated by weighing the set of stereotypes built by the non-supervised algorithm against the n_s descriptions initially given when generating the artificial learning set.

Our first evaluation consists in comparing the quality –i.e. the percentage of descriptors– and the number of generated stereotypes to the initial descriptions, n_s, while the percentage of degradation increases from 0% up to 100%. It appears that up to more than 85% of degradation, the sets of stereotypes corresponds most of the time to the initial ones (see figure 2).

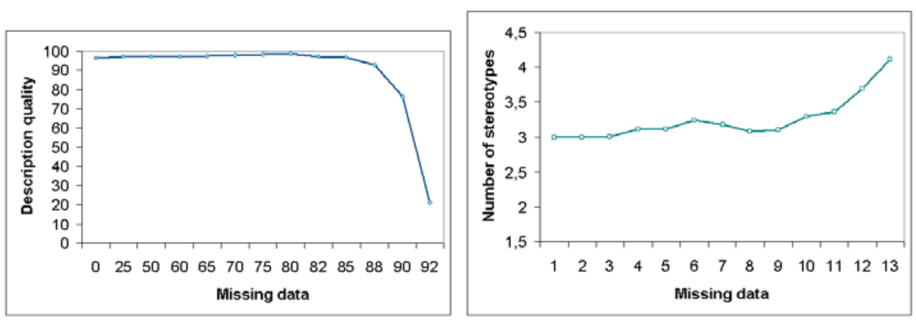

Fig. 2. Quality and number of stereotypes discovered.

The second test counts the classification error rate, i.e. the rate of degraded facts that are not covered by the right stereotype. We mean by "right stereotype" the discovered description that corresponds to the initial fact the degraded facts come from. Fig. 3 shows the results of our program P.R.E.S.S. relatively to three classic algorithm for classification: k-means, COBWEB and EM. These experiments clearly state that the results of P.R.E.S.S. are really good with a very stable learning process: up to 75% of degradation, the error rate is less than 10% and best as the three others.

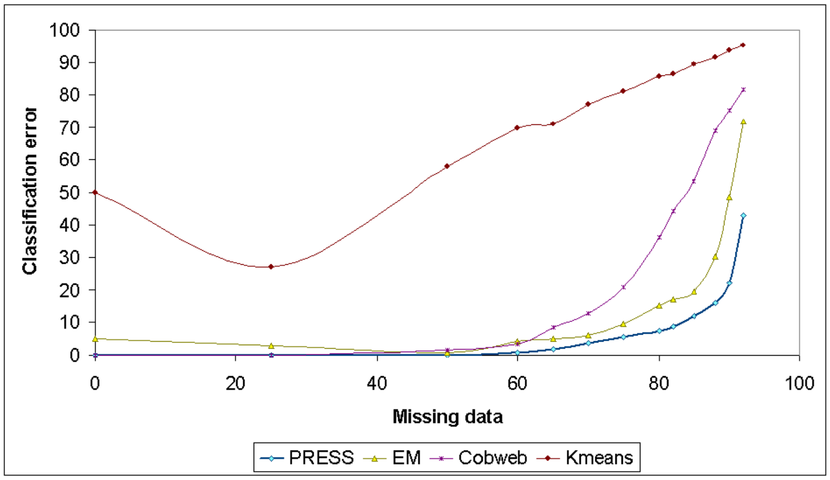

Fig. 3. Classification error of degraded examples.

3.2 Studying social misrepresentation

The application we propose deals with an historic event, the famous miscarriage of justice known as the Dreyfus Affair which occurred at the end of the 19th century in France. In 1894 Captain Alfred Dreyfus, an officer on the French general staff, was accused of spying for Germany, France's opponent in the previous war. There were many articles about this very complex affair, bringing different views depending on the date, recent events, the newspaper political leanings. Thus, the liberal pro-Dreyfus *Le Siècle* expressed opinions which were diametrically opposed to those of the conservative anti-Dreyfus *L'éclair*. The facts we considered have been taken from these articles and translated into conceptual graphs, in order to build automatically a simplified model of the affair. The objective is to understand the influence of the press on the mental representations during this period.

Type hierarchies including 399 concepts and 174 relations were built for this specific context. In addition, a typical graph was proposed in order to translate the articles into facts. Fig. 4 shows an example of a graph. It represents an article from the newspaper *L'éclair* using the CoGITaNT library implemented by D. Genest and E. Salvat [14]. This library in C++ manipulating conceptual graphs was chosen because of the gnu public licence, its great flexibility and the quality of the available documentation.

It could be summarized as follows : *the article taken from the newspaper* L'éclair *explicitly asserts that Alfred Dreyfus is guilty because Esterhazy was proved innocent by the courts.* Once several articles have been translated in this way, stereotypes can be discovered using the methods proposed earlier.

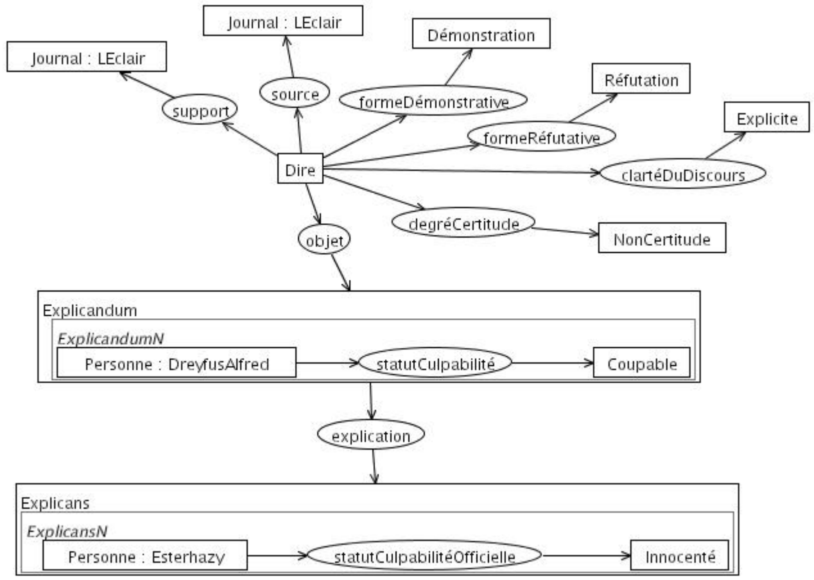

Fig. 4. A conceptual graph which translates a newspaper article.

4 Conclusion

Flows of information play a key role in today's society. However, the value of information depends on its being interpreted correctly, and implicit knowledge has a considerable influence on this interpretation. This is the case in many today's heterogeneous databases that are far to be complete and, consequently, need special techniques to be automatically completed. This is particularly true of the media, such as newspapers, radio and television, where the information given is always sparse.

In this context we propose a cognitive model based on sets of stereotypes which summarize facts by "guessing" the missing values. Stereotypes are an alternative to prototypes and are more suitable in the categorization of sparse descriptions. They rely on the notion of default subsumption which relaxes constraints and makes possible the manipulation of such descriptions. Descriptions are then completed according to the closest stereotypes, with respect to the dissimilarity measure M_D. Very good results have been found in the Attributs/Values formalism with artificial data sets. Our interest is now focused on a real application using conceptual graphs.

This work relates to the domain of social representations as introduced by Serge Moscovici in [15]. According to him, social representations are a sort of "common sense" knowledge which aims at inducing behaviors and allows communication between individuals. We think that social representations can be constructed with the help of sets of stereotypes. The way these representations change can be studied through the media over different periods and social groups in comparison with such sets. This represents an unexplored way for enriching historical and social analysis.

Finally, this paper emphasizes the danger related to a conceptualization not really adapted to a particular problem. Missing data might induce bad interpretations and lead to erroneous results. This is also what we show with the concept of stereotypes and the notion of default subsumption.

Acknowledgments

The authors would particularly like to thank David Genest for his patience and useful advice concerning the CoGITaNT library and Anne Remillieux for constructing the type hierarchies and graph examples.

References

1. Sowa, J.F.: Conceptual Structures: Information Processing in Mind and Machine, Addison-Wesley Publishing Company (1984), Massachusetts, The Systems Programming Series.
2. Reiter, R.: A logic for default reasoning. In: Artificial Intelligence, number 13 (1980) 81–132.
3. Corruble, V., Ganascia, J.-G.: Induction and the discovery of the causes of scurvy: a computational reconstruction. In: Artificial Intelligence Journal, Special issue on scientific discovery. Elsevier Press (1997) 205–223.
4. Velcin J.: Reconstruction rationnelle des mentalités collectives: deux études sur la xénophobie, DEA report, Internal Report University Paris VI, Paris (2002).
5. McDermott, D., Doyle, J.: Nonmonotonic logic 1. In: Artificial Intelligence, number 13 (1980) 41–72.
6. McCarthy, J.: Circumscription : a form of non-monotonic reasoning. In: Artificial Intelligence, number 13 (1980) 27–39, 171–172.
7. http://www.cs.uah.edu/ delugach/CG/.
8. Rosch, E.: Cognitive representations of semantic categories. In: Journal of Experimental Psychology: General, number 104 (1975) 192–232.
9. Lippman, W.: Public Opinion, Ed. MacMillan (1922), NYC.
10. Zhong, J., Zhu H., Li J., Yu Y.: Conceptual Graph Matching for Semantic Search. In: Proceedings of the 10th International Conference on Conceptual Structures: Integration and Interfaces, Spring-Verlag (2002) 92–106.
11. Glover, F., Laguna, M.: In: Tabu Search, Kluwer Academic Publishers (1997).
12. Corter, J.E., Gluck, M.A.: Information, uncertainty, and the utility of categories, In: Proceedings of the Seventh Annual Conference of the Cognitive Science Society. Lawrence Erlbaum Associates (1985) 283–287.
13. Fisher, D.H.: Knowledge Acquisition Via Incremental Conceptual Clustering. In: Machine Learning, number 2 (1987) 139–172.
14. Genest, D., Salvat, E.: A platform allowing typed nested graphs: How cogito became cogitant. In: Proceedings of the Sixth International Conference on Conceptual Structures, Springer-Verlag (1998), Berlin.
15. Moscovici, S.: La psychanalyse: son image et son public. PUF (1961), Paris.