Forgetting in Intelligent Systems

Martin Strecker
Laboratoire Artemis/IMAG, BP53X
F-38041 Grenoble, e-mail: martin@sherpa.imag.fr

April 21, 1993

Abstract

The aim of this work is to model the phenomenon of "forgetting". For this purpose, we characterize forgetting in terms of a change induced on a stable expansion in Autoepistemic Logic. Forgetting roughly consists in withdrawing a formula from the expansion, while leaving other formulas of the expansion unaffected, as far as possible. This approach is contrasted with another formalism of theory change. We then turn to the practical use of these considerations and describe how forgetting can be applied to Truth Maintenance Systems. It permits to eliminate obsolete propositions from the dependency graph, however without changing the belief status of the remaining ones. Thus, forgetting can be understood as a deletion operation that respects certain consistency constraints.

1 Motivation

Until recently, reasoning has mostly been characterized "statically", whereas little thought has been given to the change of theories as a function of changes the underlying sets of axioms. However, it becomes increasingly important to investigate how theories may change in the course of time. Modelling the dynamic evolution of theories permits to gain a more realistic understanding of how information processing works in "natural" systems.

A first step in this direction has been made by the introduction of non-monotonic logics. They enable a reasoner to derive propositions that, in a strict sense, are not sanctionned by the facts he knows; in consequence, the reasoner has to be disposed to withdrawing these propositions when confronted with new facts. In the following, we will adopt the framework of Moore's Autoepistemic Logic (AEL) [4, 5] for describing such reasoning processes.

Non-monotonic logics aim at permitting plausible reasoning in a context of incomplete knowledge, and it is tacitely assumed that the "rational agent" will in due course acquire more and more factual knowledge. With our discussion, we want to question this principle of ever growing knowledge, and we will examine how forgetting of facts can be modelled in AEL.

Why are we interested in forgetting? First of all, forgettig is a psychical phenomenon, and it is tempting to find a logical characterization on a purely descriptive level. Some results in Cognitive Science suggest that the principal reason why natural information processing systems are faced with forgetting is that they have only limited capacities. By means of forgetting, obsolete knowledge is withdrawn for the benefit of being able to store relevant knowledge. Logic has never felt a serious concern for this "bounded resource" problem, primarily because the abstract entities it deals with are not supposed to be materialized anywhere. In this paper, we do not advocate any changes to the logic itself; however, we want to model "rational agents" that might be obliged to come to terms with limited resources. In this light, we understand an ideally rational agent's theory as the set of beliefs he might *potentially* hold.

2 Formalization of Forgetting in Autoepistemic Logic

As mentioned before, we take the framework of AEL for our investigations, and we assume that the reader is familiar with its terminology. We only recall that AEL models an ideally rational agent who holds a set of "premisses" or "axioms", A, from which he constructs a stable expansion, E, which represents his theory of the world. The language of AEL is that of propositional logic, augmented by a belief operator, L. An AEL formula which does not contain this operator is called an objective formula; the subset of objective formulas of a set of formulas S is denoted by kernel(S).

When talking about forgetting of a proposition, we refer to an operation that transforms a set of axioms, A, and one of its stable expansions, E, into another set of axioms, A', with a stable expansion, E', such that the forgotten proposition cannot be deduced from E'. This can be achieved by many transformations, most of which cause the expansion obtained after forgetting to differ too much from the original one to be of any interest.

In order to make the expansions E and E' differ as little as possible, we demand that forgetting of a fact should not lead to the acquisition of new factual knowledge; thus, the set of objective formulas of E' should be contained in the set of objective formulas of E. This requirement alone might accept an expansion which forgets too much (for example the set of all tautologies). We are therefore only interested in an expansion E' whose kernel is maximal in this respect.

The above postulates motivate the following definition of forgetting, which we will call *autoepistemic consolidation.*

Definition:
Let A be a set of formulas and E = TH(A) an autoepistemic expansion of A. The autoepistemic consolidation of a formula f $\notin$ TH($\emptyset$) is the construction of the sets A' and E' such that

1. E' = TH(A') *(E' is stable expansion of A')*
2. E' $\not\models$ f *(Forgetting of the formula f)*
3. kernel(E') $\subseteq$ kernel(E) *(No new facts can be deduced)*
4. If E" satisfies conditions (1)..(3), and kernel(E') $\subseteq$ kernel(E"), then E' = E" *(Maximality)*

We will employ the notation (A', E') = $cons_{AEL}$(n, A, E).

Autoepistemic consolidation can be contrasted with the "contraction functions" that have been proposed in [2]. These contraction functions are defined in terms of a classical theory operator, which is monotonic, as opposed to the operator TH which we use here. Furthermore, it is possible to cite examples for which two different sets A' resp. E' fulfill the conditions of the definition of autoepistemic consolidation, and therefore it is not a "function". With these restrictions in mind, we can make a comparison between the postulates for a contraction function and the forgetting operator, and we obtain a far-reaching agreement.

3 Forgetting in Truth Maintenance Systems

Up to now, we have only characterized our forgetting operator in a non-constructive way. In this section, we will consider certain sets of autoepistemic formulas for which an effective transformation from A to A' can easily be carried out. This kind of set is obtained when translating dependency graphs of Truth Maintenance Systems (TMSs) into AEL. It will be seen that the results of this section are not only interesting from a theoretical viewpoint, but have important practical consequences because they permit to reduce the size of the dependency graph while respecting the consistency of the reasoning process. Forgetting in TMSs seems to have been introduced in [1], motivated by the need to reduce storage requirements in TMSs which process great amounts of data; a more detailed discussion is given in [7]. A dependency graph is the most important data structure of a TMS (for some of the terminology, consult [6]). It consists of elements that have been communicated to the TMS during its interaction with a problem solver, namely: nodes, representing propositions; and justifications, representing rules or implications. For a dependency graph D = (N,J), N is the set of nodes and J the set of justifications. In the following, a node will be represented by a lower-case letter; a justification j = <A|B→c> has the set of nodes A as its monotonic support, the set of nodes B as its nonmonotonic support, and the node c as its consequence. One or both of the sets can be empty.

A labelling assigns a belief status of IN or OUT to each n∈N; nodes which are IN are currently believed by the TMS; an admissible state of a dependency graph D is the set of all nodes that are IN under a labelling that is closed and globally grounded. As such, an admissible state can be understood as the current theory of the TMS. This intuitive notion is confirmed by results that have been established in [3, 6], who translate a dependency graph D into a set of autoepistemic formulas, F(D), and can prove that there exists a bijection between the admissible states of D and the set of stable expansions of F(D).

After these preliminaries, we will now give a definition of forgetting for a dependency graph of a TMS. Later on, the correspondance between TMSs and AEL will be exploited to show that forgetting in a dependency graph can be interpreted as a specific form of autoepistemic consolidation.

In the case of a TMS, the operation of forgetting (which will also be called "consolidation") eliminates a node from the dependency graph, and deletes or transforms the justifications in which the node appears.

Definition:
Let D=(N,J) be a dependency graph with an admissible state S. The consolidation of a node n in D for S is defined as the construction of a dependency graph D'=(N',J'), such that:

- N' = N$\setminus\{n\}$
- Let j=<A|B→c>∈J be a justification. J' is constructed in the following way:
  - If n=c, then j is deleted from J
  - If n is IN and n ∈B, or if n is OUT and n ∈A, then j is deleted from J.
  - If n is IN and n ∈A, or if n is OUT and n ∈B, then j is transformed into j'=<A$\setminus\{n\}$|B$\setminus\{n\}$→c>.
  - All other justifications remain unchanged

The following theorem shows that the belief status of the nodes remaining in the graph are not changed by consolidation:

Theorem:
Let D be a dependency graph, and D' the graph obtained by consolidating the node n in D for the admissible state S. Then S'=S$\setminus\{n\}$ is an admissible state of D'.

This result is advantageous for an implementation of consolidation, because it permits to reduce the size of the dependency graph without making a recomputation of an admissible state of the new graph necessary. But also for semantical reasons, this result is satisfactory. An operation whose purpose it is to free memory space should not change the belief in any of the propositions that are not involved in the operation.

An analogy between the consolidation in dependency graphs and consolidation in AEL seems plausible, and it can indeed be established formally by means of the above-mentioned correspondence between a dependency graph and its admissible states on the one hand and a set of autoepistemic formulas and its stable expansions on the other. The following theorem states that the consolidation of a node in a dependency graph is equivalent to the autoepistemic consolidation of the proposition represented by the node in the associated sets of AEL formulas:

Theorem:
Let D' (with admissible state S'=S$\setminus\{n\}$) be the result of consolidating the node n in the dependency graph D for the admissible state S. Further suppose that A=F(D) resp. E are the AEL sets corresponding to D and S, and A'=F(D') resp. E' the sets corresponding to D' and S'.
Then (A', E') = $cons_{AEL}$(n, A, E).

In this article, we have exposed the utility of forgetting for modelling a rational agent with limited spatial resources, and we have provided a formal definition in AEL that captures its intuitive properties. The interest in forgetting is not purely theoretical; we have presented an application of the concept in the domain of TMSs. A perspective for future investigations is to make use of forgetting in other fields, in particular if cross-references to AEL already exist (for example inheritance graphs).

References

[1] Euzenat, Jérôme; Maesano, Libero: *An architecture for selective forgetting.* In: Barbara Smith: Artificial Intelligence and Simulation of Behaviour. Springer Verlag, 1990.

[2] Alchourrón, Carlos E.; Gärdenfors, Peter; Makinson, David: *On the Logic of Theory Change: Partial Meet Contraction and Revision Functions.* The Journal of Symbolic Logic, Vol. 50, Nr. 2, June 1985.

[3] Fujiwara, Y.; Honiden, S.: *Relating the TMS to Autoepistemic Logic.* Proc. 11th IJCAI, pp. 1199-1205, 1989.

[4] Moore, Robert C.: *Possible-world semantics for Autoepistemic Logic.* Proc. of the 1984 Nonmonotonic Reasoning Workshop, pp. 344-354, AAAI, New Paltz, 1984.

[5] Moore, Robert C.: *Semantical Considerations on Nonmonotonic Logic.* Artificial Intelligence, Vol. 25, 1985.

[6] Reinfrank, Michael: *Fundamentals and Logical Foundations of Truth Maintenance* Dissertation, Department of Computer and Information Science, Linköping, Sweden, 1989.

[7] Strecker, Martin: *Formalisation de l'Oubli dans les Systèmes de Maintien du Raisonnement* (in French). Master Thesis, Institut National Polytechnique, Grenoble.