Discussion of: Forget It!
By Fangzhen Lin and Ray Reiter

Alon Y. Levy
AT&T Bell Laboratories
levy@research.att.com

The authors present an intuitive and well motivated model-theoretic definition of forgetting a sentence (or a set of sentences). They give a syntactic method for creating the theory resulting from forgetting a sentence, and describe several properties of their definition.

Forgetting
There are two questions about the authors' treatment of forgetting:

1. Although the definition given for the notion of forgetting is rather intuitive, are there any desiderata from any definition of forgetting? i.e., are there any intuitive properties that we would like any definition to satisfy? How does the notion of forgetting differ from other notions investigated in the context of belief revision? In particular, how does forgetting differ from belief erasure [Katsuno and Mendelzon, 1991] and belief contraction [Alchourron et al., 1985]?

2. Clearly, one motivation for investigating the notion of forgetting is to obtain a theory with which it is easier to reason. Can it be shown that theories resulting from forgetting sentences are (in some sense) simpler? (e.g., not much larger, or do not require a more expressive language).

Relevance and Forgetting
I have questions about the usefulness of the definitions of relevance based on forgetting. The key problem is that forgetting is a model-theoretic notion, while there is a strong intuition that (ir)relevance is more a proof-theoretic or a meta-theoretic notion. The problem is that when looking at models of a theory, all the logical conclusions become explicit, thereby losing the dependencies between sentences, and information about how a conclusion is derived from the finite axiomatization of the theory (i.e., the knowledge base). The example given in the paper illustrates this point. Consider the theory

\[ T = \{ \text{student}(John), \text{student}(John) \supset \text{young}(John) \}. \]

In this theory, the sentence \( \text{student}(John) \) is relevant to the query \( \text{young}(John) \), whereas in the logically equivalent theory:

\[ \{ \text{student}(John), \text{young}(John) \}, \]

\( \text{student}(John) \) is irrelevant to \( \text{young}(John) \). Therefore, even though the theories are equivalent, irrelevance is not preserved. In fact, the authors designed the forgetting operator such that it affects the truth value of as few sentences as possible in the resulting theory (specifically, it affects only the truth values of sentences containing the predicate in the forgotten sentence). Therefore, the resulting definition has a somewhat strong property that if a query does not contain the predicate of the forgotten sentences, then the forgotten sentence will be irrelevant. (The authors consider relevance w.r.t. a set of learnable sentences. However, even in that definition, if the query or the learnable sentences don’t contain the predicate of the forgotten sentence, irrelevance will hold). Model-based definitions of irrelevance are especially inadequate when one wants to use the notion of irrelevance to speed up inferences. In fact, work on using irrelevance to speed up inferences (e.g., [Levy, 1993; Levy et al., 1994]) has shown that syntactic approaches to irrelevance have several advantages. In particular, these approaches have led to novel (and practical) algorithms for automatically detecting irrelevant sentences, and has shed light on the utility of relevance reasoning.

References

