Belief Revision in Type Theory

Fairouz Kamareddine
Joint Work With
Tijn Borghuis and Rob Nederpelt

http://www.cee.hw.ac.uk/~fairouz/talks/talks2001/talkiat01.ps

23 October 2001

IAT'01, Maebashi, Japan

Summary

- We explore belief revision for belief states in which an agent's beliefs as well as his justifications for these beliefs are explicitly represented in the context of type theory.

- We make the justifications an agent has for his beliefs as first-class citizens.

- Since every belief is accompanied by its justification (and the rules operate on both), every inconsistency that surfaces in the agents belief state has its own justification containing the justifications of the beliefs that cause the inconsistency.

- This allows for a deductive perspective on belief revision which can be implemented using existing machinery for deductive reasoning.

Type theory for knowledge representation

- *Subjectivity:* Knowledge of an agent is *partial*: no one knows everything, and agents differ in what they know.

- *Justification:* Knowledge is justified: agents not only *know* things, but they have *reasons* for knowing them.

- *Incrementality:* The knowledge of an agent can be *extended* as new information becomes available.

- *Subjectivity is captured by types:* Each concept is formalized as a type, each instance of the concept is a term inhabiting this type.

- *Justification is captured by terms:* by the PAT-principle, justifications are first-class citizens, formalized in the type-theoretical syntax as terms.

- *Incrementality is captured by contexts:* An agent's knowledge state can be formalized as a type-theoretical context. Addition of new information to the knowledge state can be formalized by adding statements to the context, dismissing information amounts to reducing the context.

- 'Everything an agent knows' at a certain instant can be divided into:

  - *Explicit knowledge* expressed by the statements in context Γ. These are explicitly represented pieces of knowledge directly available to the agent.
  - *Implicit knowledge* expressed by statements *derivable* on context Γ. These are consequences (obtained by inference) of an agent's explicit knowledge.

Concluding remarks

- We explored the use of explicitly represented justifications in belief revision where beliefs and belief states were represented respectively as type theoretical statements and contexts

- Justifications make it easy to identify the beliefs that cause inconsistency of the belief state and greatly simplify the handling of dependencies between beliefs.

- Our approach is applicable to agents with limited computational resources because it is deductive and we do not require that our theory of belief revision itself selects which beliefs have to be removed.

- This holds independently of the strength of the logic in which the belief change operations are cast.

- Our approach is applicable to: a) a large family of type systems, and hence b) given the connections between type theory and logic, in a wide range of logics

- Our work has been implemented by Bunt on the basis of a standard type theoretic theorem prover where the agents belief state is represented as type theoretical contexts as described in this talk.

- Our framework is related to:
  - revision for belief bases and to Foundations Theory, but does not suffer from the drawbacks usually associated with foundations theory such as problems with disbelief propagation, circular justifications, and multiple justifications for the same belief;

– the work of Hansson on semi-revision, whose notion of consolidation can be simulated in our framework and where new information is not automatically completely trusted.