Representation and Reasoning

Proceedings of the Stuttgart Conference
Workshop on Discourse Representation,
Dialogue Tableaux and Logic Programming

Sonderdruck

Max Niemeyer Verlag
Tübingen 1988

Table of Contents

LOGIC-BASED OPEN SYSTEMS

Robert Kowalski
Imperial College
London

Carl Hewitt (*Hewitt, in BYTE 1985*) has argued that logic is inadequate for complex open systems. Open systems, he argues, involve cooperating and conflicting agents interacting in parallel with mutually inconsistent beliefs. Logic is inadequate because it is suitable only for static, passive systems which cannot tolerate inconsistency.

I shall argue that the methodology of logic and logic programming can be extended to meet the needs of open systems. The concept of knowledge assimilation (*Kowalski, 1979*), which was developed to deal with changing, possibly inconsistent belief systems represented in logic, goes part of the way, but needs to go further: Belief systems need to be embedded within agents which process their own, resident goals, plan courses of actions to achieve their goals, and have the ability to put actions into effect. Such agents need to interact, co-operate and compete with other agents within the same environment.

I shall argue that these requirements can best be achieved within a framework where individual agents employ logic for reasoning about the world, the accomplishment of goals and the execution of plans.

The Theorem-proving paradigm

Many of our attitudes towards the use of logic during the past century have been conditioned by the theorem-proving paradigm associated with the use of logic to formalise axioms, theorems and proofs in mathematics. The paradigm focuses attention upon the activity of proving a variety of theorems from a fixed and unchanged set of axioms. The axioms are generally required to be consistent and complete. Applied outside mathematic, the formalisation of common sense reasoning about the world, the theorem-proving paradigm provides a powerful model of problem-solving, which is neutral with respect to wether the problemsolver is a human or a machine (*Kowalski 1979*): Axioms represent beliefs about the world, theorems represent goals, and proofs represent solutions. Logic programming and logic databases are special cases, in which axioms represent programmes or databases and theorems represent computational tasks or queries.

Different theorem-proving strategies correspond to different, and sometimes quite powerful problem-solving techniques. The outstanding limitation of the theorem-proving paradigm as a problem-solving model, however, is its lack to attention to the process of successive refinement of belief and to reflect changes in the world itself.

Knowledge Assimilation

The notion of knowledge assimilation was developed to adress these limitations. The use of logic for databases provides one of the clearest examples of what is required: Just as important as the process of posing different queries to a given, fixed database is the need to update the databases to deal with changes in the world.

Updates in a conventional database consist of additions and deletions of sentences expressing relationships among individuals. An upate is performed on the database if it is logically consistent with the integrity constraints; otherwise it is rejected.

In the more general case where the "database" represents beliefs about the world, it is appropriate to restrict updates to the additon of new beliefs. The effect of **deletion** can be obtained by representing **explicitly** the time period for which a relationship holds and **adding** a statement that the time period has ended. Such an additive approach to databases has been developed as a calculus of events (*Kowalski 1984/85*) in a logic programming framework.

Deduction can be used to detect violation of integrity constraints in conventional database systems; but it plays a more central role in additive knowledge assimilation. Four particular deductive relationships between an existing belief system, including integrity constraints, and new additional belief stand out:

I) The new belief can be derived from existing beliefs. In this case the new belief is redundant and the new belief system is identical to the old.

II) Part of the existing beliefs can be derived from the new belief together with the rest of the beliefs. In this case the derived beliefs are redundant and can be replaced by the new belief.

II) The new belief is inconsistent with the existing beliefs. In this case the detection of inconsistency identifies some subcollection of beliefs which contribute to the inconsistency. Any such belief is a candidate for rejection or alteration.

IV) None of above relationships can be determined within the resources allocated to assimilation of the new belief. In this case the new belief is added to the existing beliefs.

In practise, these deductive relationships can be determined by means of proofs and their detection depends upon the resources which can be allocated to the proof process. Most logic theories of belief, however, assume that beliefs are determined by means of model-theoretic semantics. Critics of logic are quick to point out that according to such semantics an inconsistent set of beliefs implies any conclusion. But the complex nature of many real-world tasks, they argue, is such that we can rarely expect a useful collection of beliefs to be consistent. So using logic to represent beliefs about the world forces us to accept all possible beliefs as conclusions and is therefore worthless in practise.

This argument has led many critics to reject the use of logic as a model of common sense reasoning. But the same argument can be used to reject the model-theoretic account of implication in favour of resource-constrained proof-theoretic explication of belief, it might be possible to derive both **"A"** and **"if A then B"**, but not the further conclusion **"B"** within the resources which are available.

Moreover, since some proofs are easier to find than others, the resource-constrained proof-theoretic account of belief associated different degrees of accessibility with different logical consequences. Furthermore, since the resources available for finding proofs may vary from occasion to occasion, the beliefs derivable on one occasion may differ from those derivable from another. No such distinction between the logical consequences of a set of assumptions is possible with the model-theoretic account of logical implication.

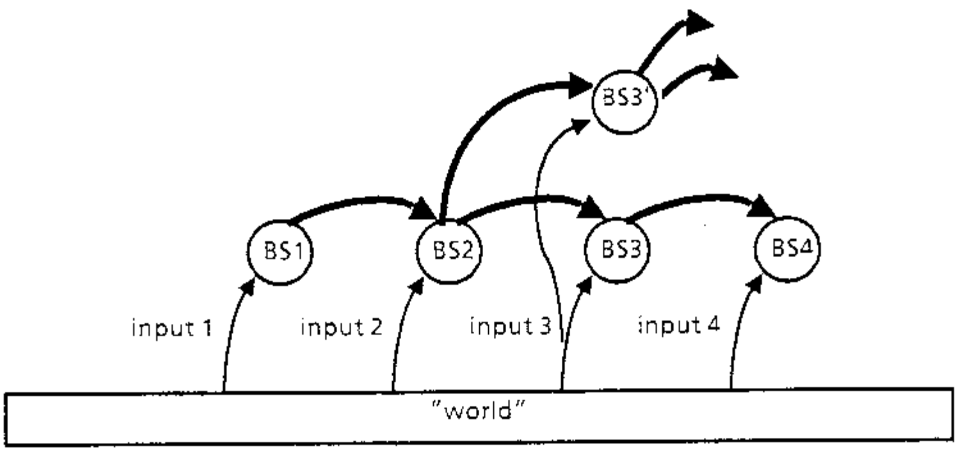
The proof-theoretic and the model-theoretic notions of implication do not necessarily conflict with one another. As long as the completeness theorem holds, the model-theoretic notion of implication coincides with the proof-theoretic notion with unconstrained resources. Thus resource-constrained, proof-theoretic implication converges to model-theoretic implication in the limit.

Other pragmatically important differences exist between the model-theoretic and the proof-theoretic treatments of implication. For the sake of efficiency, all good proof procedures constrain the form of deductions so as to avoid as much as possible the generation of useless consequences. Resolution, for example, deliberately avoids the **"thinning rule"** which makes it possible to derive conclusion **"A or B"** from assumption **"A"**, thereby loosing information. It is precisely this thinning rule which also leads to the derivation of arbitrary conclusions **"B"** from proofs of contraction. (**"Falsity"** implies **"Falsity or B"**, which implies **"B"**.)

Thus the theorem-proving paradigm provides an alternative account to relevant implication: An implication of a conclusion from a set of assumptions is a relevant implication for a given proof procedure if the conclusion can be derived without introducing "obvious and eliminatible irrelevancies". Such a notion of relevance is wholly proof-theoretic and has no corresponding model-theoretic analogue. Moreover, depending on the efficiency of the proof procedure and the resources available, an implication which may seem to be relevant in one circumstance may prove to be irrelevant in another.

It can even be argued that on the proof-theoretic account the detection of inconsistency, rather than leading to the derivation of arbitrary and therefore useless conclusions, can be a positive force for change. Theorem-provers which attempt to avoid irrelevant deductions for the sake of efficiency also tend to avoid introducing irrelevant assumptions into proof of contradiction. Therefore, a derivation of inconsistency by such a theorem-prover has the positive merit of identifying a relatively small set of beliefs which contribute to the derivation and therefore candidates for belief revision.

The process of knowledge assimilation by a single agent can be pictured as a search tree, each node of which represents a system of beliefs. Inputs from "the world" trigger transitions from one state of belief to another. Alternative states of belief can arise either because of ambiguities in the interpretation of the input or because of alternative ways beliefs can be revised when an inconsistency is detected.



Such a search space can be explored depth-first, one belief system at a time; or it can be explored in parallel, alternative belief systems at the same time.

The notion of knowledge assimilation has been dicussed in greater detail elsewhere (*Kowalski 1979*). It attempts to adress some of the needs Hewitt has identified, but it doesn' t go far enough. Even if we allow multiple belief systems and organise them into communities of problem-solvers, they would still be essentially passive. They can be made active, however, by giving them their own goals and the ability to perform actions on the enviroment.

Active belief systems

The ascription of goals to belief systems is a natural extension of the use of logic in the theorem-proving paradigm. In its full generality, logic is concerned with the relationship which holds when assumptions imply conclusions. The theorem-proving paradigm focuses on only one of the important ways in which this implication relationship can be used: to derive different theorems from a given fixed set of axioms. In the problem-solving interpretation of theorem-proving this corresponds to the solving different goals from a given fixed set of beliefs.

The implication relationship can be used in other ways. Its uses for consequence generation and for induction have been mentioned elsewhere (*Kowalski 1979*).

For active belief systems, however, we need to use it for beliefs which change over time and resident goals which persist from one state of belief to another.

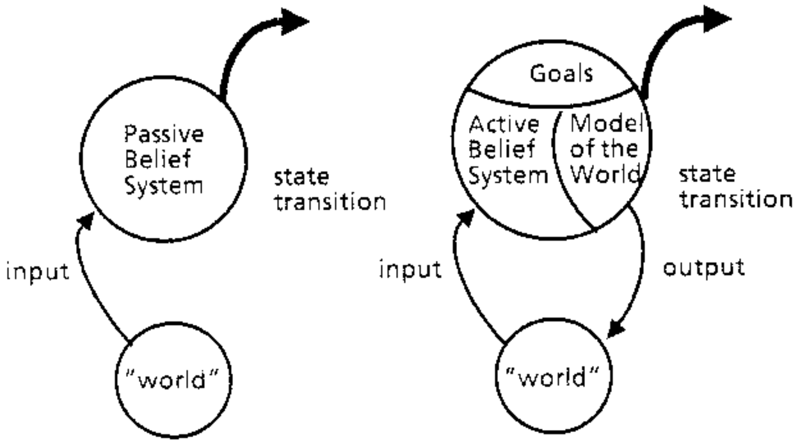
The existence of resident goals adds further dimensions to the process of knowledge assimilation. The assimilation of a new statement can be assessed not only for its deductive relationship with the belief system as a whole but for its deductive relationship with resident goals in particular.

The addition of a new belief might lead to the solution of an existing goal or collection of goals. Everything else being equal, such a belief would be preferred to one, which is logically unrelated to any goals at all. Such discrimination between beliefs can be used to favour one belief over another when both contribute to the derivation of an inconsistency or when both are alternative interpretations of the same ambiguous input.

A new belief might lead instead to the partial solution of existing goals. Such a partial solution determines subgoals whose future accomplishment would result in the complete solution of those goals.

A belief system with resident goals could recognise and reward the accomplishment of its own internal goals as well as respond to externally posed goals and querries. However, such a system would still be essentially passive. An active belief system requires the ability to satisfy goals by performing actions which produce changes in the world. To do so, it needs to have a model of the world and of the effect its actions have upon it. The possession of such a model would give the further ability to plan and execute courses of actions which achieve goal satisfying states of the world. Since reasoning about the world and about the effect of actions on the world can be performed by means of deductive logic, the only extension of the notion of passive belief system which is reqired is the ability to perform actions.

The difference between an active and a passive system, therefore, is the difference between the possession and the absence of an output channel for communicating actions to the world. Both kinds of systems can be pictured as processus in a dataflow network where the world itself is regarded as a process.



The interface between the active belief system (actor) and the "real world" consists of variable-free atomic sentences which report events and states of affairs to the actor and record the actor' s action in the world. For the actor to fulifil its plans and achieve its goals, its model of the world needs to correspond as closely as possible to the world itself. According to the proof-theoretic account, this correspondence between the actor' s beliefs and reality is tested deductively by the actor processing input reports and comparing them with its own predictions derived from general beliefs about the world and specific statements about the past. This contrasts with the standard model theoretic account of the relationship between belief and reality.

According to the model-theoretic and proof-theoretic accounts, the interface between the real world and the actor's theory about the world consists of "reports" which are variable-free atomic sentences.

Whereas on the model-theoretic account such sentences validate or invaliate the theory, on the proof-theoretic account they are processed deductively like any other belief which is to be assimilated. Truth and falsity of sentences in the theory is replaced by consistency and inconsistency of reports about the world. A report which is consistent can be analysed further according to wether it is already implied by the theory, implies part of theory or is logically independent from it. A report which is inconsistent, on the other hand, identifies a subcollection of beliefs which participate in the proof of inconsistency and which therefore are candidates for revision. If the report is relevant to the proof of inconsistency, then these candidates include the report itself. Thus, whereas reality is the arbiter of truth in the model-theoretic account, it is as much a possible subject of belief revision as any other belief on the proof-theoretic account.

Multi-actor systems

The greater the number of actors, the greater the likelihood of a mismatch between the intended effects of an individual's actions and the perceived, resulting state of affairs. Different actors with different and possibly competing goals can formulate conflicting plans and perform conflicting actions independently and concurrently.

Actors can reduce the likelihood of their actions not accomplishing their goals by taking another actor into account during the plan formation process. By including a model of another actors' beliefs and goals, they can anticipate and attempt to avoid the occurence of actions which compete and conflict with the achievement of their own goals. Alternatively, they might negotiate and agree a revised and non-conflicting set of goals in advance. They may even recognise that they share common or mutually supportive goals and agree to co-operate for their mutual advantage.

To participate in the negotiation process, an actor needs to have an understanding of its own goals. Such "self-knowledge" may or may not be accurate, in the same way that an actor' s belief about another actor may also be more or less complete or correct. Moreover, an actor may confuse the image it wishes to project to other actors with its own "true-self".

The appearence of irrationality

Such inaccuracy in an actors´s model of the world can give the appearance of irrationality. Explicitly stated goals and beliefs may contradict external behaviour.

Critics of logic point to such contradictions to support their thesis that human are not logical and that logic is inappropriate and irrelevant for the design of complex, open systems in general. Our contrary view postulate a need for logic at two levels. At the object-level, an actors needs to use its beliefs about the world to derive actions which contribute to the achievement of its internal goals. The use of logic to perform the reasoning that is needed is largely subconscious and inescapable.

It is the conscious exercise of logic at the metalevel to reason explicitly about the world and other actors that is more controversial. To the extent that the actor´s beliefs are correct and complete such conscious use of logic can only augment and improve the exercise of logic at the subconscious level. To the extent that is inaccurate and incomplete, however, it can magnify errors, lead to in consistent conclusions and result in self-destructive behaviour.

This line of argument leads to an interesting and possibly unexpected conclusion. Given a conflict between an individual´s consciously reasoned argument and his subconsciously supported intuition, it may be better to reject the argument than to quarrel with the intuition. There is greater scope for error when logic is applied consciously to models of one´s own beliefs and of other actors beliefs than there is when it is applied to one´s own beliefs directly and subconsciously.

Conclusions

Hewitt draws attention to the "challenge" of open systems. His discussion of the phenomena which need to be dealt with and of the problems which need to be solved is compelling. However, his arguments against logic are not.

I have argued that logic is highly suited for open systems. The notion of knowledge assimilation already goes part on the way, by adressing some of the issues involved in dealing with belief system that change over time.

However, the notion of knowledge assimilation needs to be extended. Greater attention needs to be given to the belief system´s own internal goals and to the process of formulating plans or actions to achieve those goals. More significantly, belief systems need to be given the power to perform actions. The likelihood of such actions achieving their desired effect can be increased by improving the actors´model of the world, including its beliefs about other actors and their goals. An actor can anticipate other actors´ behaviour and can plan to deal with conflict or to avoid it. Conflict can be avoided by negotiating with other actors and revising goals and subgoals. Mutually supportive goals can be identified and actors can co-operate to achieve them.

I have considered two of the arguments against logic. I have argued that one of them is an argument against the model theoretic semantics of logic rather than against logic itself. The practical application of logic involves resource constrained deduction rather than model theoretic application of truth definitions. Moreover, for the sake of efficiency, deductions need to be restricted so that they are as relevant as possible. These restrictions block the derivation of arbitrary conclusions from inconsistent assumptions in particular.

The other argument against logic is based upon the appearence of irrationality.I have argued that logic needs to be exercised at both the conscious and the subconscious levels. At the subconscious level it is inescapable, though subject to resource limitations. At the conscious level it is prone to error and produces conclusions which are no better than the assumptions upon which they are based. In some cases the assumptions may be so inadequate that the conscious exercise of logic may do more harm than good.