ELECTRONIC NEWSLETTER ON
REASONING ABOUT ACTIONS AND CHANGE

Issue 99011 Editor: Erik Sandewall 27.4.1999

Today

Today we have the pleasure of announcing an additional submitted article: "Causality in Action Theories", by Camilla Schwind. The present issue also contains Graham White's concluding answer to the referees of his article, and an answer by Pat Hayes to Graham's earlier comment in the review discussion of the Levesque/Pirri/Reiter reference article.

Murray Shanahan has written that he leaves his present article as it is, and does not propose to update it further.

On the practical side, our web page structure has been revised in a way that speeds up the interactions. The new design only updates individual frames (one or more) instead of the whole window. The graphical design has also been improved. As usual, the webpage is the best way of accessing the past discussions in this electronic colloquium, and it is found at http://www.etaij.org/rac/

ETAI Publications

Received research articles

Causality in Action Theories

by **Camilla Schwind**

Abstract: Causal reasoning has become very important in the last few years in logic based action theories. In this article, we study how causal inference has been integrated and used in action theories. After presenting and discussing criteria considered as the most meaningful, we present four formalisms which are approaches to actions and causality and we try to situate each formalism in relation to the criteria.

Discussion about received articles

Simulation, Ramification, and Linear Logic

by **Graham White**

Comment from: **The Author**

Here's my reply to the referee's remarks.

> But, I am talking about a case with ramification, so that Reiter's methods do not apply; if we allow the effect axioms to talk about all

> *But Reiter's approach has been extended to dealing with state constraints and ramifications [2,3]. Several other successful approaches to the Ramification Problem emerged during the past five years or so; I mention [4,5,6]. Repeating myself, I am missing convincing arguments why your approach is substantially different and superior in comparison.*
>
> *[2] F. Lin and R. Reiter, State constraints revisited, J. of Logic and Computation 4(5), 1995.*
>
> *F. Lin, Embracing causality in specifying the indirect effects of actions, IJCAI 1995.*
>
> *N. McCain and H. Turner, Causal theories of action and change, AAAI 1997.*
>
> *M. Thielscher, Ramification and causality, AIJ 89, 1997.*
>
> *M. Denecker etal, An inductive definition approach to ramifications, ETAI 3, 1999.*

Let's start with what appears to be a misconception. The referee writes "state constraints and ramifications", as if the two were synonymous. They are not, and one of the articles he refers to (by Deneker et al.) explains why not ([6], pp. 3f.):

> *we consider causal laws to be representations of how changes propagate. As such, causal laws may be but do not need to be related to state constraints. Below is an example.*
>
> *Assume a network with many input signals which can be directly modified through actions, an arbitrary internal structure, and one or more output lines. A counter on one of these output lines counts the number of times the output voltage changes from 0 to 1. [rules follow]*
>
> *The indirect effects described by the above rules are in no way related to a state constraint: there is no relation between the contents of the counter and the current state of the network."*

So there are examples which formalisms based on state constraints cannot cope with: this rules out the referee's reference [2].

Of the other approaches, [3], [4] and [6] suffer from the limitation (which Deneker et al., in [6], admit to) that multiple changes of the same fluent at the same time cannot occur. One can give well motivated examples which exhibit this behaviour: consider, for example, a modification of Deneker et al.'s counter example in which the counter is now connected to two of the output lines, and in which the input action may cause changes in either or both of them. The propagation of the effects may well result in the occurence of intermediate states of the counter when one of the output lines has had its effect but not the other.

Similarly, the electrical circuit example in Thielscher's [5] (p. 341, Fig. 5) has intermediate states of fluents (namely, the light bulb being on) which are caused during the ramification process but which do not persist into the final state. Deneker et al. do describe such examples as "not well defined", but they do not argue for this restriction.

So my approach, which correctly handles examples such as those given above, is more general than all of the cited approaches apart from Thielscher's [5], which also handles these approaches correctly.

And regardless of the success or failure of my own approach, what we have done so far is to eliminate all but one of these supposedly "successful approaches to the Ramification Problem" by referring to arguments and examples found in the other ones. This makes me a little suspicious of the rules of this refereeing game as it is played in the AI community. After all, if I say "The sky is blue" and the referee says "No, you're wrong, the sky is green", then I may at least want to verify that the sky is, in fact, blue, just to check. But if the referee says "you're wrong, the sky is green AND the sky is red", then what do I reply? And if the referee cites supposedly "succesful approaches to the ramification problem", and not all of them can be simultaneously succesful because the later ones argue against the earlier ones (so that either the earlier ones are wrong, or the arguments and examples in the later ones are wrong); well, what can be concluded from that, other than that there are papers in the literature which disagree with mine (no surprise) and also with each other (also no suprise).

What I want to insist on is the argument of my paper, and particularly the argument in Section 2.4. What I argue there is this: that the standard approach to these problems is unduly sensitive to vocabulary, because we can make formalisations, which seem to be adequate, of the same problem in vocabularies which seem to be logically equivalent (because they are equivalent in first order logic), and we find that when we apply the circumscription procedure to them, that one is successful and the other is not. And we seem to have no way of selecting one vocabulary rather than the other, other than that one gives a successful circumscription and the other does not. So this circumscription procedure does not seem to be anything that we can apply to a problem that we do not know the answer to: regardless of how adequate our formalisation might seem to be, we always have to face the possibility that there might be another formalisation (equivalent in first order logic, but using a different vocabulary) which gave different results under circumscription. We seem to have no way of differentiating between good vocabularies and bad, other than by looking at the results.

Almost all of the cited approaches are, at least prima facie, susceptible to this sort of problem; they all rely, to some extent, on a preferred vocabulary – explicitly in [2] and [3], because we have fluent-based circumscription; implicitly in [4], because of the definition of interpretations there as sets of fluents, and in [6] because of the use of reification in the definition of the fluent calculus. My own approach is not: I refer, in the article, to the fact that "in [the technical report] we show that we have a good notion of interpretation for theories of this sort, and that, in our example, we can change from the [one vocabulary to the other] without breaking anything". (p. 12 of the article). Furthermore, I have a semantics for my logic (also in the technical report); all of the above I am willing to expound at greater length if required (though I am still somewhat attached to the principle of separating the more philosophical and expository arguments from the mathematical details, and that was what I had hoped to accomplish in the separation between article and technical report).

One could, finally, talk about the great computational advantages of approaches based on linear logic when compared to those based on circumscription or other of the conventional non-monotonic logics.

So, finally, if you want a summary of what I think the advantages of my approach are, they would be:

- Correctness on a wider range of examples than any other approach, apart from Thielscher's [6];
- Extensionality: that is, the formalism admits a robust notion of change of vocabulary (what is technically known as interpretation), and this notion of interpretation is supported by a semantics;
- Computational efficiency.

Foundations for the Situation Calculus

by **Hector Levesque, Fiora Pirri, and Ray Reiter**

Comment from: **Pat Hayes**

Graham White's extended comments (Enrac 13.4) are interesting, but not for the reason I suspect he had in mind. White, as he says, is approaching the entire discussion from a mathematical perspective. If things like situations (whatever they are), world-states, actions and histories were indeed mathematical structures, then his remarks would be apposite: it would be good professional practice to ask what the appropriate morphisms were between them, for example, and what would count as isomorphism. But they aren't mathematical structures, and this entire game we are playing here isn't mathematics. We are using logic not to describe mathematical structures, but to describe aspects of nature.

This isnt a matter of looking on mathematics with 'extreme suspicion', by the way. (I have a mathematical training myself, and find the logical security of mathematics comforting, if anything.) The point is that what we are wanting to do is more like science than mathematics. There is a real world out there, and our job is to try to describe it adequately. Many people in action reasoning, in fact, see their task as rigorously empirical, and advocate testing their logical theories by incorporating them into actual physical robots with measurable behaviour. Having made that analogy, however, let me immediately modify it. What we often find ourselves doing isnt science so much as a kind of pre-science: the kind of conceptual analysis that precedes an exact science. We are still sometimes arguing about what our words mean. But whether it is scientific or pre-scientific, it is always directed towards, and inspired by, something in the actual world, not by mathematical criteria.

To see the difference, take a mathematical concept such as 'group', and ask how one could know a group if one met it. Something is a group just when it satisfies the group axioms. The axioms are all there is: *anything* that satisfies the group axioms is a group, by definition. One can find examples of many famous finite groups all over the place: the Klein group, for example, is revealed in the possible orientations of a rectangular card. Now, contrast this with a concept like 'banana'. Even an optimistic mathematician would surely not expect the be able to reconstruct bananas within set theory (even up to isomorphism), or give a sensible account of banana-morphisms. Actions and world-states are more like bananas than groups.

White's examples make the contrast very clear. He uses the example of the natural numbers:

> *For example (and this example is taken from Benacerraf's article What Numbers Cannot Be) one can define the natural numbers both as a sequence of sets thus:*
>
> $\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\{\{\emptyset\}\}\}, ...$
>
> *or as a sequence of sets thus:*
>
> $\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} ...$
>
> *Now clearly nothing hangs on this: the two definitions are clearly equivalent in any sense which matters. They can be regarded as two different implementations of the natural numbers.*

First, note that the idea of natural number preceded both of these set-theoretical constructions. In fact, it preceded the idea of 'set' by several centuries, if not millenia. (Its easy to forget that the great success of logicism in modern foundations of mathematics tends to produce a historical revision of ideas that go back a very long way into human history; modern accounts of mathematical structure are written from a perspective which *identifies* mathematical domains with their modern reconstructions in set theory from the 1920s, or (still more sophisticated) in category theory from the 1950s.) So even in mathematics, I think that the formalisations tend to be driven by pre-formal intuitions. But why would we expect that actions would form a category, or be definable (even up to isomorphism) in set theory? I see no reason to suppose they would, and several good reasons to suppose not.

But even if they did, there is still a further question to ask, which is how that mathematical structure fits onto the actual world we are aiming to describe.

An example. Consider a ball bouncing on a table. We could say that each period between two bounces is a 'state' of the ball, and that each bounce is a state-transition (one where it might lose a little energy, say, producing a lower bounce next time.) Or, we could say that each moment of contact with the table was a 'state', and each trajectory between bounces a state-transition. These might well be isomorphic structures (finite total orders) in White's mathematical sense, but they kinds of axioms one will write about the 'states' and 'actions' will be very different in the two cases; and if someone is thinking in one way when talking to someone else who is thinking the other way, mutual incomprehension may well result. (Ray Reiter and I are conducting a discussion offline to see if we are misunderstanding each other in exactly this way.)

> *... So what I'm really worried about is this: do the "snapshot" and "history" definitions of situations express essentially different (i.e. non-isomorphic, or more exactly not naturally isomorphic) mathematical objects, or is the difference only notational, that is, analogous to the two definitions of the natural numbers above.*

Answer: neither of these. The difference between a state of a world and a sequence of actions is an ontological distinction, like that between a banana and an orange (or, a better analogy, between a positions and a velocity.)

Pat Hayes

PS a minor note:

> *Now Reiter here talks of constructors, and says, quite correctly, that his approach has constructors: for each action α, we get a constructor $do(a, s)$ which maps*
>
> $$Actions \times Situations \mapsto Situations$$
>
> *But can't we also define similar things for the snapshot view? If there are a number of actions that lead from a snaphsot-situation to its successor, then we also have a map*
>
> $$Actions \times Situations \mapsto Situations$$
>
> *(and maybe it's only a partial map, but a partial map is still a morphism in some appropriately defined category, and maybe you can still establish suitable equivalences...) So I'm not convinced that there's anything in this distinction.*

Of course there is nothing in this as a *distinction.* The $do(a, s)$ notation was used to refer to the world-state which results from doing the action a in the world-state s since the very first formulation of the situation calculus ontology.

This newsletter is issued whenever there is new news, and is sent by automatic E-mail and without charge to a list of subscribers, in plaintext format. The present postscript issue is a bona fide rendering of the same information in alternative typography. Instruction for contributors and other additional information is found at the following URL:

`http://www.ida.liu.se/ext/etai/actions/njl/`