Electronic News Journal on Reasoning about Actions and Change 0
Vol. 2: *1–1000*, 31.1.1998
http://www.ida.liu.se/ext/etai/actions/nj/9801-3/

Protocol of the discussion during January, 1998 about the following article:

Wolfgang Bibel

Let's Plan It Deductively!

Edited by: Erik Sandewall
Linköping University, Sweden

The discussion about Wolfgang Bibel's invited lecture at IJCAI 1997 started in the September, 1997 issue of the present News Journal. After a pause of several months, the discussion continued in January as follows.

From: Erik Sandewall on 25.1.1998

Dear Wolfgang,

With respect to your IJCAI article, I have a number of questions several of which relate to the research methodology or paradigm being used. I observe that when Marc Friedman ended one of his questions with "why prefer one solution to the other", your answer was

> *To the best of my knowledge Michael Thielscher (...) was the first who gave a solution to the ramification problem which overcomes deficiencies in any previous solution (...) no previous solution would model reality in a correct way. A better solution in this sense must be preferred to a deficient one. (...)*

In this way, you appeal to the traditional method of validating approaches to common-sense reasoning by ways of counterexamples: a method is accepted until disproved by an example. Unfortunately, this research methodology can not provide any reliable conclusions. As has been recognized in core computer science since a long time, the lack of knowledge of a counterexample does not prove that a proposed solution (or a program) is correct.

On this background, I have the following questions or observations:

1. What is your perspective on current research in reasoning about actions and change where we provide *assessments* or *validation results* for new theories, instead of merely proposing new logics based on a combination of intuition and toy examples? The new approach is being used by several of us, including Shoham (who started this trend), Lifschitz, and myself. In spite of being a survey paper, your article does not mention this development at all.

2. In the section on qualification, you describe a solution using transition logic: "TL opens a new way to deal with this problem". The solution is exemplified using McCarthy's old "potato in the tailpipe" problem, but not

motivated in any other way. In particular, there is no statement of when the method is correct or is not correct, nor any proof or reference to a proof of its correctness under some precisely stated assumptions.

Unfortunately, however, it is well known that simple solutions to the qualification problem will easily fail if the naive potato and tailpipe example is modified ever so slightly. For example, suppose there are two cars, A and B, it is known that a potato is put into the tailpipe of one of them, and one asks whether car B will start properly. In such a case, the absence of positive knowledge that the tailpipe of car B has been plugged, does not allow one to draw the default conclusion. Or, suppose one does not know whether the plugging of the pipe preceded or succeeded the attempt to start the car; the same difficulty arises again.

Given that previously proposed solutions have failed so easily, one would like to have some evidence that the method proposed in your article is not going to encounter those problems, or other ones which may come up. In other words, here is a case where the shortcomings of the older, example-driven methodology are evident.

My question is what can be said in general about the method you propose for qualification: when is it known to work, and when not?

3. In section 4, you write "LCM has been the first method which actually solved these aspects of the frame problem and did so in the optimally possible way". Since there is no obvious definition for optimality in this respect, I wonder which quality measure do you use, and what is the proof that the LCM method is optimal with respect to it?

4. Another example of an incompletely substantiated claim occurs in your section on ramification. You write "Lifschitz' categorization of fluents does *not* work in this example. We need to *categorize the actions* into primary and secondary ones (rather than the fluents) as done in the solution presented in this section". However, the second sentence does not follow from the first one. The fact that the particular variety of fluent categorization that was proposed by Lifschitz doesn't work for the example, does not prove that **all** fluent categorization methods fail for it, or for some reasonable class of examples. So in what sense do we "need to" categorize the actions?

5. Also with respect to your formalization of Thielscher's example with three switches and a relay, it is remarkable that the electric circuit in question can easily be understood in terms of dependencies and persistence, but the proposed formalization requires the axioms to represent the propagation of *changes*: "if this fluent changes in such-and-such a way, then that fluent also changes in such a way". This seems clumsy and counterintuitive. Do you claim that it is the best possible representation in the present state of the art?

6. As you correctly observe, a rational agent must be capable of reasoning about the timing of actions and about changes within the duration of an action. In section 6.4 of the paper you describe in outline how to introduce timing of actions into TL. However, the most obvious way of arranging this is by approaches that use explicit, metric time, as in Shoham's work in the 1980's, the Features and fluents approach, and the modern event calculus. The PMON(RCs) logic presented by Doherty and Gustavsson at KR-96 is an example of such a logic. To summarize,

- PMON(RCs) uses explicit time, which allows one to characterize the duration of actions and the intermediate changes within its duration;

- PMON(RCs) also allows one to characterize the timing of concurrent actions with overlapping duration periods;
- PMON(RCs) allows delayed effects of actions;
- The circumscription policy used in PMON(RCs) has been shown to be reducible to the first-order case, which insures that standard theorem-proving techniques and their optimizations can be used to compute entailment.

In other words, it does in concrete detail what you only describe in gross outline. My questions are:

1. What are the advantages of your approach over the one I just referred to? (This is a valid question in view of your statement "A better solution in this sense must be preferred to a deficient one").
2. In PMON(RCs) and other approaches that use explicit metric time, it is straightforward to make statements about durations, comparing durations of actions, and so on, basically because each interpretation represent an entire history of the world. How can this be done in a logic like TL where the => operator takes one from one state to the next, and updating the current time in the process?

7. You write "section 6 shows how the various aspects involved in reasoning about actions and causality can be taken into account within TL". However, nothing in section 6 or elsewhere in the article presents any concrete results about how spontaneous change in the world can be represented - it is as though the world were entirely static when no actions are taken. The resulting concept of causality is quite meagre.

Here again, there *exists* in fact a body of results in this area, ranging from the Situated Action Theory of Morgenstern and Stein and my own early work on integrating differential equations into logic (presented at KR-89 and IJCAI-89) to Murray Shanahan's earlier and recent work which uses a very similar approach, and which has also been used for implementing a simulated robot. When you write "another issue concerns the integration of differential equations and their computation within a logic such as TL", the uninformed reader will not easily guess that the integration occurred nine years ago, although not for TL.

8. In section 4, on the topic of how to obtain a specific plan as a solution to a planning problem, you write "[Bibel 1986a] introduced state literals, *S(x)*, which keep track of the states passed through while executing a plan. (...) By unification the variable denoting the goal situation will then along with a successful proof always provide a term that expresses the linearized sequence of actions." Yes, but how is this different from the use of the Answer predicate which was proposed and used by Cordell Green in the late 1960's?

9. In subsection 6.3 you refer to an "easy solution" for allowing action laws for special cases to take precedence over a more general law for the same action. However, in the example you quote, the specialized case only contains an *additional* effect besides the effect in the general case. Does the same solution also apply if some of the effects of the general case *do not* arise in the special case?

10. If the answer to the previous question is positive, does the planning method that you propose using unification against the state literal (answer literal) still work when such specificity is allowed? The question arises since unification against answer literals results in planning backwards from goals, whereas the solution for specificity is defined in terms of forward simulation of the plan.

11. A final observation: The title and the introduction of your article focusses on planning, as shown when you write: "In this paper I review the state of the art in deductive planning...", but the major part of the contents (after introducing the TL logic as such) deals with problems in reasoning about actions and change: the frame problem, ramification, and so on. May I suggest that some additional coverage of modern results in the latter area would be appropriate, in particular since this is claimed to be a survey article.

Sincerely
Erik Sandewall

From: Wolfgang Bibel on 27.1.1998

Dear Erik,

Thank you for your continuing interest in the paper underlying my invited IJCAI-97 talk. Since the main focus of your questions concerns "the research methodology or paradigm being used" I need first to clarify the nature of this paper.

As you correctly state the paper is meant to be "a survey paper"; however you failed to add (as clearly stated in the paper in the same sentence) that it is meant to be one "with an emphasis on the contributions from research groups influenced by the author's own work" (p2). As an invited talk this is an absolutely appropriate focus. For instance, your own invited talk at the ECAI conference in Budapest was of exactly the same nature. In fact you may recall that I then for fun complained about YOUR ignorance of all of OUR work (although NOT in public as you now decided to do). Reiter's Chambery IJCAI address did again the same representing the Toronto school. And so forth.

In consequence of this purposely chosen and clearly stated nature of the paper there are of course substantial omissions wrt planning, action and causality approaches in general. In fact I would not even consider myself capable of giving a fair survey of the current state of the art in these areas (perhaps not even in deduction any more where I feel myself still better "at home"). And of course in such a paper there is no need to justify each of the contributions from scratch (even for invited papers there are at least unwritten space limits). If such a need would be, one should not have invited me to give this talk in the first place. 9 out of 11 of your questions are essentially answered by these statements, as I will briefly show you shortly, and have been disappointing in this regard.

Apart from the (in more than one respect incomplete) survey character there is one novel contribution in this paper which is the extension of reasoning about change and causality within the LCM (the action part of TL) frame to include (or to be embedded into) also classical reasoning. In other words the paper marries two complementary areas which belong together. Given this fact the focus of any questions wrt this particular paper should actually be directed to this particular contribution rather than to the contributions which have been done many years ago (as far back as 1985) and to a large extent by students and close colleagues of mine rather than to me. Now to your questions in (boring but time-consuming) detail. As to the paper's section numbering I refer to the version in the web (not to the one in the proceedings).

ad 1.

> *provide* assessments *or* validation results *for new theories*

Since there is no new theory I do not think this particular paper is in lack of such an addition, nor is this the case for the old theories surveyed in the paper. Classical deduction is so established that I am sure you agree wrt this part. As to the action part I know from you personally that you (as well as others) appreciate Thielscher's work especially for his contributions in this respect (relation of the fluent calculus with Lifschitz' $\mathcal{A}$ etc). In my Section 6.3 I survey the equivalence results between the fluent calculus and the action part of TL. On the basis of these theorems his results apply to TL in the same way.

ad 2.

> *My question is what can be said in general about the method you propose for qualification: when is it known to work, and when not?*

In line with the paper's survey character I properly state in Section 7.2 to which you refer here the source of this approach: "The solution is again adapted to TL from [Thi96]'s FC using the example discussed there in great detail." In that paper Thielscher indeed gives theorems and proofs addressing your question. He is the right person to answer specific questions in this regard. But I can add a general comment to issues like

> *For example, suppose there are two cars, A and B, it is known that a potato is put into the tailpipe of one of them, and one asks whether car B will start properly. In such a case, the absence of positive knowledge that the tailpipe of car B has been plugged, does not allow one to draw the default conclusion.*

The advantage of staying within the framework of logic is that you have a long experience on your side (and this is my research methodology which I followed for more than 3 decades). TL can of course handle your example (and also the other one you mention) easily since it features (a resource sensitive) disjunction.

ad 3.

> *Since there is no obvious definition for optimality in this respect, I wonder which quality measure do you use, and what is the proof that the LCM method is optimal with respect to it?*

Assuming the method is semantically correct (discussed in 1) there is of course an "obvious definition of optimality" namely the ones familiar from AD. Any explicit frame axiom increases the search space for the deductive mechanism (and even the proof lengths). Since LCM, TL, FC, LL have no frame axioms at all they are optimal in this respect.

ad 4.

> *So in what sense do we "need to" categorize the actions?*

Your question refers to the paper's Section 7.1 in which I again clearly stated: "The discussion in this section closely follows [Thi97b] ...". So, given the equivalence results mentioned above, your question should again better be addressed to Thielscher directly who in his paper and his habilitation thesis (also cited in this part) provides excellent answers to your question.

ad 5.

> *Also with respect to your formalization of Thielscher's example with three switches and a relay, it is remarkable that the electric circuit in question can easily be understood in terms of dependencies and persistence, but the proposed formalization requires the axioms to represent the propagation of* changes*: "if this fluent changes in such-and-such a way, then that fluent also changes in such a way". This seems clumsy and counterintuitive. Do you claim that it is the best possible representation in the present state of the art?*

If you know a less "clumsy and counterintuitive" formalization of the example then use it. TL as any logic is a neutral formalism which does not bother about the way particular scenarios are represented in them (as Bob Kowalsky convincingly argued in the memorable Crete workshop in 1985 where, in your presence and bombarded by your criticisms, I first presented LCM - as you can still read in the proceedings transcript of the discussions then).

ad 6.

> *My questions are:*
>
> - *What are the advantages of your approach over the one I just referred to? (This is a valid question in view of your statement "A better solution in this sense must be preferred to a deficient one").*

Again this four paragraph Section 7.4 summarizes the work reported [BT97] and [Gro96] so that the answers to your question are better to be looked up in the original sources rather than in my paper and addressed to the authors rather than to me.

> - *In PMON(RCs) and other approaches that use explicit metric time, it is straightforward to make statements about durations, comparing durations of actions, and so on, basically because each interpretation represents an entire history of the world. How can this be done in a logic like TL where the => operator takes one from one state to the next, and updating the current time in the process?*

It is obvious how to state metric time durations in TL (and in fact illustrated in the lifting example of that section) by having Tt before => and $T(t+d)$, d being the duration, after it (the example specializes to $d = 1$). I am not sure at this point whether Bornscheuer and the other people at Dresden working on time did already consider this more detail. My survey is definitely not a complete one.

ad 7.

> *You write "section 6 shows how the various aspects involved in reasoning about actions and causality can be taken into account within TL". However, nothing in section 6 or elsewhere in the article presents any concrete results about how spontaneous change in the world can be represented - it is as though the world were entirely static when no actions are taken. The resulting concept of causality is quite meagre.*

Again, my survey is definitely not (meant to be) a complete one even within the LCM family of approaches let alone all the rest. Spontaneous change has been formalized again by Thielscher in the references given in Section 7. I will include a pointer to that issue to avoid the (false) impression of "meagreness". Your other point concerning differential equations does not tell me anything new and my remark in the paper remains anyway correct.

ad 8.

> *Yes, but how is this different from the use of the Answer predicate which was proposed and used by Cordell Green in the late 1960's?*

The technique is of course the same and well established in logic. So what is the point?

ad 9.

> *Does the same solution also apply if some of the effects of the general case* do not *arise in the special case?*

No, it does not. It only applies in "occasions of a similar nature" as stated there. Again Thielscher is the expert on these issues.

ad 10.

> *If the answer to the previous question is positive ...*

Since it is negative I skip this point.

ad 11.

> *I believe that some additional coverage of modern results in the latter area would be appropriate*

I disagree as explained at the outset.

Best regards, Wolfgang

From: Erik Sandewall on 28.1.1998

Dear Wolfgang,

A question can serve as a means of obtaining information, but it can also be the syntactic form that's used for questioning a proposition. For those of my questions that were of the former kind, you have provided valuable answers - thank you. However, for those that are of the latter kind, I want to pursue our discussion a bit further on two specific issues: methodology of research in our area, and citation principles. Both of them have a general interest, I believe, and for simplicity let's address them one at a time.

I'll start with the citation principles. The following are two cases containing a quotation from your article, my question or comment, and your answer:

1. Concerning the answer literals or state predicate:

> *For that purpose, [Bibel, 1986a] introduced state literals,* S(x), *which keep trace of the states passed through while executing a plan. ... By unification the variable denoting the goal situation will then ... always provide a term which expresses the linearized sequence of actions.*
>
> > Yes, but how is this different from the use of the Answer predicate which was proposed by Cordell Green in the late 1960's?
>
> *The technique is of course the same and well established in logic. So what is the point?*

2. Concerning the combination of differential equations and logic:

> *Modelling continuous processes within a logic has become an active area of research though. ... [Herrmann and Thielscher, 1996]. Another issue concerns the integration of differential equations and their computation within a logic such as TL.*
>
> > When you write "another issue concerns the integration of differential equations and their computation within a logic such as TL", the uninformed reader will not easily guess that the integration occurred nine years ago, although not for TL.
>
> *Your other point concerning differential equations does not tell me anything new and my remark in the paper remains anyway correct.*

I must say that I find your reactions exceptional. It is true that it's always difficult to produce correct and complete references to earlier work; it is difficult for example to give adequate coverage of, and comparisons with approaches different from one's own. In the beginning of your answer, you gave an eloquent description of the problem that we all encounter: you are frustrated by a lack of reference to your own work (even in a case where it was not relevant, actually), and you don't have the space to give an adequate presentation of the work of others.

There is no easy solution to this problem, but would it be possible to agree on a few elementary principles, such as:

- Don't quote a well established result as your own innovation
- Don't try to make earlier work invisible, for example by quoting a topic as a current area of investigation and entirely omitting references to earlier work on that topic that you do know about
- Most importantly, if by accident you get out of line wrt principles such as these, agree to make a correction.

These points are fairly obvious, I would think, although I'm afraid they're not always practiced. There may be a number of other principles that ought to be added as well; the present ones resulted from reflecting on the two concrete cases that I quoted above.

Besides a clarification of policy, I think we need additional mechanisms for feedback and correction of citations. The traditional mechanism, namely, the confidential peer reviews, is clearly not sufficient. One obvious way of proceeding is to use ETAI's review mechanism using public discussion, which allows anyone to contribute questions or suggestions to the author in a transparent but undramatical way. This may be used for pointing out additional references that are relevant for the article.

Using a public forum for this feedback to the author has the additional advantage that it can help to resolve disagreements. If an author is unwilling to accept a suggestion for additional citations, it becomes possible for others to listen to the arguments and to form their own opinion. Then at least, if there is a citation fault it's less likely to propagate to later articles written by others.

I am sure that you have also thought about these issues, in particular as the Research Notes Editor of the Artificial Intelligence Journal, and I look forward to your comments on this important topic.

\- Erik