ELECTRONIC NEWSLETTER ON
REASONING ABOUT ACTIONS AND CHANGE

Issue 98083 Editor: Erik Sandewall 25.11.1998

Today

Today we send out David Poole's answer to Hector Geffner, in the discussion about David's ETAI submitted article. This discussion is now closed since David has indicated that he wishes to revise the paper and then it will be sent to the referees. We reopen the discussion for possible additional comments after the referees have returned their verdict.

We can also announce a contribution today by Hector Levesque, Fiora Pirri, and Ray Reiter to the collection of reference articles for approaches to actions and change. The collection now has a very good coverage, and we also hope of course to have an article about the event calculus.

Additional contributions are of course welcome as well. Remember, however, that the purpose of these reference articles is to provide the basic definitions for a well established approach, so that it shall not be necessary to repeat those definitions each time in later articles. The purpose is *not* to present new results. On the contrary, the reference articles should only contain material that is sufficiently stable that it can be expected to remain relevant for a while.

Contributions to the review discussion about these reference articles is also welcome, and we suggest that the articles ought to be discussed from the point of view of their intended purpose. That is, the major question is whether the proposed text characterizes sufficiently well the current best practice in the approach being presented, and whether it is clear, understandable, and on the right level of detail.

A second use of these reference articles may be to provide a more solid basis for the perennial comparative discussion about advantages and limitations of current approaches. During earlier periods of activity, this discussion was hampered by the fuzzyness about exactly what was included in each approach. If this discussion should be reignited after the appearance of the reference articles, then the new contributions will be included as continuations of the existing panel discussion on ontologies of actions and change.

ETAI Publications

Received research articles

Foundations for the Situation Calculus

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

Abstract: This article gives the logical foundations for the situations-as-histories variant of the situation calculus, focusing on the following items:

- The language of the situation calculus.

- Foundational axioms for the domain of situations.
- Axioms for an underlying domain theory.
- The syntax and semantics of the logic programming language GOLOG.
- Axioms for knowledge and sensing actions.
- Essential metatheoretic results about the situation calculus.

Discussion about received articles

Decision Theory, the Situation Calculus and Conditional Plans

by David Poole

Comment from: David Poole

To:**Hector Geffner**

I've enjoyed reading your paper "Decision Theory, the Situation Calculus and Conditional Plans". I find it hard to disagree on your general point that a general model of action must combine elements from logic and probability/utility theory, so my questions and comments are addressed more to the question of how this combination is achieved in your proposal. I'll try to comment first on the more general issues.

1. In page 24, you say "The representation in this paper can be seen as a representation for POMDPs".

If so, I wonder, why not present the representation in that way from the very beginning? Wouldn't that make things much clearer?

One of the problems with writing a paper that is trying to bridge different areas is to try to explain it for all readers (or to keep them all equally unhappy). I suppose I have succeeded when POMDP researchers say "this is just a representation for POMDPs" and when situation calculus researchers claim "this is just the situation calculus with some some probability added" or those studying Bayesian networks claim that "this is just a rule-based representation for dynamic Bayesian networks with actions and observables". Describing it explicitly in terms of POMDPs may have made it easier for POMDP researchers, but not necessarily for everyone else.

Namely, a POMDP is described by (see [1]):

A - state space
B - actions
C - transition probabilities
D - cost function
E - observations
F - observation probabilities

Then the constructions in your language could be seen as a convenient/concise/modular way for specifying these components.

I thought I did do this. (At least I was trying to do this, and I knew when I was finished when I had defined all of the components of a POMDP). The actions and observables of Definition 2.6 are the same as B and E. I specify C, D & F using rules. In particular I represent the cost function in terms of rules that imply the *utility* predicate (Section 2.6), and observation probabilities in terms of rules that imply the *sense* predicate (Section 2.7).

> *Actually, once the mapping from the language to these components is determined, no further semantics would be needed (i.e., the probabilities of trajectories, the expected utility of policies, etc., are all defined as usual in POMDPs). This is actually the way we do things in [2].*

That is essentially what I did. There are some problems with just relying on the definition of a POMDP. This has to do with what is a state. In the formal work, the notion of state is taken as primitive. However, for a well-defined semantics we need to be clearer. A state could either mean:

(a) everything that is true at a stage. This would include the actual time (if it is relevant), the values of sensors, the accumulated reward, etc.

(b) a sufficient statistic about the history to render the future independent of the past given the state (which, by definition is Markovian). This typically doesn't include the time, the output of the sensors, or the accumulated reward.

I have taken the approach of (a). Most POMDP researchers take the approach of (b) because their algorithms reason in terms of the state space. They therefore want to keep the state space as small as possible. I don't want to reason at the level of the state space, but instead at the level of the propositions.

> *2. I can see some reasons why not to present your representation as a "front end" to POMDPs.*

In general I don't want to do this, because I don't think that reasoning in terms of the state space is the right thing to do. What we can learn from AI is that we want to reason in terms of propositions or random variables, not in terms of the state space. One of the themes of my work is to find compact representations and to exploit the compactness for computational gains. Thus, I don't want to present this as a "front end" to POMDPs. If you mean following the semantic intuitions of POMDPS, then I think I do. But POMDPs are just one of the formalisms I am building on. They have nothing to say about useful independence assumptions or about concise descriptions of actions (state transition functions), which is what this paper is mainly about.

> *you want to be more general; e.g., be able to deal with cost functions that depend on the system history. If this is the case, I'd suggest to introduce "generalized" POMDPS where (cumulative) cost functions do not have to be additive (btw, in such POMDPS belief states are not necessarily sufficient statistics ..)*

(unless you include the accumulated reward in the state.)

> *you want to accommodate multiple agents. Yet this is not done in this paper, but even in that case, multi-agent POMDPs could be defined as well, and have been defined in the completely observable setting (e.g., see Littman).*

This was done in my AIJ paper, "The Independent Choice Logic for Modelling Multiple Agents Under Uncertainty". I actually don't think that the ICL_SC formulation is as good for this as the formalism in that paper.

> *you are not interested in determining an optimal or near-optimal solution of the POMDP but are interested in analyzing the expected cost of a policy supplied by a user in the form of a contingent plan. Again, this is no problem from the point of view of POMDPs, as long as the contingent plan determines a (stationary or non-stationary)* function *of belief states into actions (see [4]).*

First let me point out that POMDPs are not defined in terms of belief states. It is a theorem about POMDPs (Astrom 1965) that says that an optimal policy can be represented in terms of a function from belief states (probability distributions over ordinary states) into actions. However, "optimal" here doesn't take into account computation time. It is pretty obvious that this isn't true for bounded optimal agents (where the computation time and the space are taken into account), which is where I would like to take this.

I *am* interested in determining an optimal or near-optimal solution (but one that takes computation time and space onto account). I want to be able to define the expected utility of any strategy the agent may be carrying out, whether it maintains a belief state or not. No matter what program the agent is following, it can be described in terms of a conditional plan that describes what the agent will do based on the alternate sequences of observations (i.e., its function from observation sequences into actions). Note that this doesn't refer to anything inside the agent. The agent could be reasoning with a probability distribution over states, remembering a few previous observations, just reacting to its current observations, or actually following a conditional plan. The important thing is what it does is based on what it observes.

What I am advocating is good hacks; the best agent won't do exact probabilistic reasoning. We need good strategies, perhaps like your RTDP-BEL. However, it is not obvious (to me anyway) that the bounded optimal solution will necessarily be an approximation to the (unbounded) optimal solution. The best real (bounded time and bounded space) agent may do something quite different than approximating belief states. Unless we have some representation that lets us model the expected utility of an agent following its function from observation and action history into actions (i.e., its transduction), we won't be able to specify when one agent is better than another. The conditional plans (policy trees) let us model this, without any commitment to the internal representation of the agent.

> *Indeed, you require* more *than this when you demand (page 18), that the tests in conditions of the plan, be* directly observable *at the time when the conditions need to be evaluated. This is not required in POMDPs [4] or in possible world accounts [3], where the test may be known indirectly through other observations.*

NO! I mean that the conditions are those things that are observable. I mean whatever evidence (the "other observations") that the POMDPs or possible worlds accounts take into account. This is a standard technique in Bayesian networks/ influence diagrams; we create a variable that represents the actual value sensed. I am quite happy to have this as a (state) variable. You don't want it as a part of your state because it isn't necessary to make the system Markovian. This is fine, but I can model these "other observations" using the "sense" predicate. Is the sense predicate part of the state? I don't know; that is one reason why I didn't blindly accept the notion of state.

> *3. A final comment about a non-trivial problem that results from the decision of removing all probabilities from state transitions transferring them into suitable priors in "choice space".*
>
> *Consider the effect of an action "turn_off" on a single boolean fluent "on" with transition probabilities:*
>
> $$P(on = false \mid on = true \;;\; a = turn\_off) = .5$$
> $$P(on = false \mid on = false \;;\; a = turn\_off) = 1$$

Let's say that initially $on = true$ and then we perform a sequence of N consecutive *turn_off*s. Then the Probability of $on = false$ is 0.5^N, which clearly depends on N.

I don't see how this behavior could be captured by priors over choice space. This seems to be a strong limitation. It looks as if probabilistic actions cannot be handled after all. Is this right?

Here is how I would represent this: Facts:

$$on(do(turn\_off, S)) limp on(S) \wedge off\_failed(S) dot$$
$$on(init) dot$$

C_0 contains $\{off\_failed(S), off\_succeeded(S)\}$ for all situations S $P_{0(off\_failed(S))} =$ 0.5 for all situations S [Note that the completion of the facts for on gives your two conditional probabilities.]

We can then derive

$$P(on(do(turn\_off, do(turn\_off, do(turn\_off, init))))) = 0.5^3$$

just as you want (similarly for any N). We need an assumption for each time step. The minimal explanation for

$$on(do(turn\_off, do(turn\_off, do(turn\_off, init))))$$

is the set of the following:

$$off\_failed(init)$$
$$off\_failed(do(turn\_off, init))$$
$$off\_failed(do(turn\_off, do(turn\_off, init)))$$

In all worlds with this explanation true, the light is on in the state after three *turn_off*s. The probability of these worlds sum to 0.5^3.

Any probabilistic actions that you can represent in a dynamic Bayesian network, I can represent in the ICL_SC. Moreover I can represent it at least as succinctly as the dynamic Bayesian network (up to a constant factor) and sometimes exponentially (in the number of state variables) more succinctly (when there is context-specific independence). The mapping is local and modular.

Thanks for taking the time to read and comment on the paper. I hope my comments help makes the paper and what I am trying to do clearer. Let me stress that this paper has nothing to say about computation. Whereas your work was motivated by being a front end to an actual POMDP algorithm; the motivation here relies on being able to deliver on POMDP algorithms that can exploit the conditional and contextual independencies of the representation. Unfortunately, I can't show you such algorithms now. But we are working on it.

David

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/`