ELECTRONIC NEWSLETTER ON
REASONING ABOUT ACTIONS AND CHANGE

Issue 98007 Editor: Erik Sandewall 23.1.1998

Today

Hector Geffner, Judea Pearl, and David Poole have reacted independently but in similar ways against one of Pat Hayes's statements yesterday; their answers follow. Then, Erik Sandewall answers Hector Geffner concerning nonmonotonicity with respect to observations.

Today's contributions are accumulated to the continuing panel debate on "ontologies for actions and change", and can be found under that heading in the on-line panel-debate structure.

Debates

Ontologies for actions and change

From: Hector Geffner

Some brief comments about Pat's last comments.

> *In any AI the same (causality) principle makes perfect sense when actions are* exogenous*; such actions, I think, we can agree, should never affect your beliefs about the past ..*
>
> *But actions - or external events - do change ones beliefs about the past. They do not change the past itself, of course: that is the causality principle. But consider for example coming into a room in an empty house and finding a hot cup of coffee resting on a table. One immediately infers that someone else has been present there recently.*

It's important to distinguish *observations* from *actions*. In dynamic systems the first are usually expressed as "initial conditions" e.g., $x(0) = 5$, $loaded(0) = false$, etc; while the latter are the inputs to the system.

An *observation* at time i ("cup in the table") of course should have an effect on your beliefs at times $j \leq i$ (the past) Basically the effect of an observation is to prune state trajectories (this is explicit in Sandewall's filtered entailment, in $\mathcal{A}$, etc)

On the other hand, what I'm saying (and of course, many other people) is that *exogenous actions* at time i, unlike observations, should *not* have an effect on your beliefs at times $j \leq i$.

You may say that if you drop the cup to the floor and it breaks, then you should change your beliefs about the past, infering things like that the cup *was* made of glass, etc.

Yet, it's not the *action* that is having an influence on your beliefs about the past; it is the *observation* that it breaks (you may say that "breaks" is not an observation but an action; yet in that case, it's definitely *not* an *exogenous* action as it depends on variables in the model)

If you remove that observation, and simply drop the cup, you will learn nothing about the past.

BTW, the causality principle is *not* about physics; I believe it's about *models of physics.* Whether such (dynamic) models are specified by means of mathematical equations or rules in a logical language, I don't think it's relevant (for compliance with the principle).

A final point before this gets to long. Pat says

> *.... such simulations are often unreliable precisely because we don't have sufficiently complete knowledge; and when this is so, we cannot cleave to the strict causality principle, but are obliged to use techniques such as nonmonotonic reasoning which allow us to recover gracefully from observed facts which contradict our predictions, which would otherwise enmesh us in contradictory beliefs. Nonmonotonicity is a good example of the need to revise ones beliefs about the past in the light of unexpected outcomes in the present, in fact, which gets us back to the YSP:*

Actually, I think *none* of the received models of actions in AI (say $\mathcal{A}$, Toronto Sit Calc, Basic Features and Fluents, etc) does that.

I believe they are all *monotonic* in the set of observations. In other words, if they predict F at time i, nothing that they observe is going to affect that prediction. At most they can make the theories inconsistent. (one of the few exceptions that I'm aware of, is a proposal in a paper of mine [c-aaai-94-245] in which state trajectories are ordered by a plausibility measure)

They are *non-monotonic* however in the set of *actions.* That is, you can affect the prediction "F at time i" by performing an action *before* i.

It's basically like in the standard dynamic models either deterministic or probabilistic.

So I think our models are not so different from more standard models of action. Of course, they are very *different* in the description languages; but that's the type of difference that you have between Strips, and transition-functions. The first is much more convenient, but it's basically a "front-end". It's not a *new* model of action; it's a new language for describing the most basic one (deterministic models). Of course, I'm completely convinced that this is very important, and I think it's precisely there where KR/action fits in.

- Hector Geffner

References:

[c-aaai-94-245] H. Geffner. *Causal Default Reasoning: Principles and Algorithms.* Proc. AAAI National Conference on Artificial Intelligence, 1994, pp. 245-250.

From: Judea Pearl

Correcting a statement by Pat Hayes.

Hector Geffner said, about the principle of causality:

> In any AI the same principle makes perfect sense when actions are *exogenous*; such actions, I think, we can agree, should never affect

> your beliefs about the past (indeed, as long as you cannot predict exogenous actions from your past beliefs, you shouldn't change your past beliefs when such actions occur).

To which Pat Hayes replied:

> *But actions - or external events - do change ones beliefs about the past. They do not change the past itself, of course: that is the causality principle. But consider for example coming into a room in an empty house and finding a hot cup of coffee resting on a table. One immediately infers that someone else has been present there recently. We constantly make inferences about the past on the basis of present knowledge, even in such technical areas as military planning.*

Correction

The principle of causality is in fact stronger than Pat's interpretation of it. Not only the past, but also our beliefs about the past, do not change as a result of actions, *unless* the acting agent is part of our model. Moreover, if the agent *is* part of our model, then actions cease to be interesting and problematic as they are today (at least in some AI approaches to actions).

To explain: The action of putting a cup of coffee on the table does not change the state of the coffee or the table before the action, and does not change our beliefs about the state of the coffee before the action. *But*, Pat will argue: seeing the coffee on the table allows us to infer that "someone else has been present there recently." True, but only if we are concerned about the actor's whereabout and if the limitations or motivations of the action-producing agents are **in** the model (e.g., that someone must be in a house to do it, and will probably do it if he/she is thirsty, etc.) Once this action are perceived as produced by a modeled agent, it is no different than any other events, say the agent tripping or being hit by a car, because then it is licensed to trigger the usual inferences of predictions and abduction that observations trigger.

Hector said all this by using the term "exogenous". Indeed, the problematic aspects of actions surface when we try to treat actions as exogenous, namely produced by external agents **not** modeled in our system. Only by considering an action exogenous can we talk about the world "changing" (anyone ever saw a world truly changing?.)

Put another way, the peculiarities of actions vis a vis observations stem from the boundaries we put around our models and our audacity to call our models "worlds".

B. Russel (1910) noted (my words, not his) that if we model the entire universe, there is no meaning to causality or to actions, because the manipulator and the manipulated lose their distinction.

One problematic aspect of dealing with external actions is that ordinary abduction, from the immediate consequences of those actions, must be suppressed. For example, we do not want to abduce that it rained when we decide to pour water on the driveway and get it wet. This suppression is easily enforced in systems such as logic programming (Lifschitz et al) in which inference has preferred directionality, and where abduction is not a built-in feature (e.g., contraposition is not sanctioned) but must be crafted as a special feature. However, the suppression is problematic in standard logic and probability where inference is symmetric, and no differentiation is made between abduction and prediction. Such differentiation is one role played by causally ordered relations.

Well, enough said for one tiny correction.

From: David Poole

Pat Hayes wrote:

> *But actions - or external events - do change ones beliefs about the past. They do not change the past itself, of course: that is the causality principle. But consider for example coming into a room in an empty house and finding a hot cup of coffee resting on a table. One immediately infers that someone else has been present there recently. We constantly make inferences about the past on the basis of present knowledge, even in such technical areas as military planning.*

No. No. No. *Observations* can change one's beliefs about the past. Doing an action doesn't change beliefs about the past. Observing the hot coffee made you update your beliefs about the past. The action of coming into the room didn't change your beliefs about the past (unless you include the observation that the action succeeded).

To make everything clearer it is crucial to distinguish sensing (making observations) from motor control (doing/attempting actions). The only role of sensing is to update your beliefs. Sensing in this sense doesn't change the world (of course, acting in order that you can sense can, and often does, change the world). Acting (i.e., sending a command to your motor controllers) doesn't change your beliefs about your present or your past, but only about the future. Acting can change the world.

If we fail to distinguish these, confusion will reign supreme. Surely in order to make some progress we could agree on such a distinction, then get back to the argument with at least one confusion cleared up!

David Poole

P.S. I am wondering why such confusion between observing and acting may have arisen in the first place.

I conjecture is that it has to do with the preconditions of actions. For example, the $pickup(x)$ action has the precondition that there is nothing on x. Then by carrying out the action, can you infer that there was nothing on x? But this doesn't make sense. What happens if I had attempted to pickup x when there was something on it? What if I didn't know whether there was something on x when I tried to pick it up? It seems that the only sensible interpretation of the precondition is that if there was nothing on x and I carried out $pickup(x)$, then the expected thing would happen. If something was on x and I carried out $pickup(x)$ then who knows what may happen. The role of the precondition is that it is only sensible to attempt to carry out the action when the preconditions hold.

An alternative explanation of the confusion may be that the action $pickup(x)$ is an observation of the effect of my motor control on a particular state of the world. If I carry out a particular motor control when there is nothing on x, then a $pickup(x)$ action arises. When I do the same motor control when there is something on x then some other action arises. Then I do not choose the action, but I only choose the motor control (consider the case when I don't know whether there is something on x or not, and I try to pick it up). Is this what people mean by an action? Surely then it is imperative to distinguish that motor control (for a better description) that I get to choose, from the observation of the effect of that motor control.

From: Erik Sandewall

Do all current approaches behave monotonically with respect to observations? On 23.1, Hector Geffner wrote:

> *Actually, I think* none *of the received models of actions in AI (say A, Toronto Sit Calc, Basic Features and Fluents, etc) does that.*
>
> *I believe they are all* monotonic *in the set of observations. In other words, if they predict F at time i, nothing that they observe is going to affect that prediction. At most they can make the theories inconsistent.*

With respect to features and fluents, this is true for those cases where proven assessments exist at present, but not in general. The full catalogue of ontological characteristics includes such phenomena as "surprises" and "normality", both of which are nonmonotonic with respect to observations. A simple example is for the stolen car scenario: the car is left in the parking lot, three nights pass, it is for sure that the car is not removed during the day, the default is for it to stay where it is at night as well. The default conclusion is that it's still there after the three nights. Then add the observation that it's gone at the end of the period. Without admitting surprises, this scenario is inconsistent, as Hector writes. If surprises are allowed for, then you may conclude that it was removed during one of those three nights.

If you add in addition the fact that the parking lot was supervised the first and the third night, making removal impossible, then one is entitled to infer that it was stolen during the second night. Thus, the initial default conclusion regarding the presence of the car in the parking lot during the day after the second night goes from "yes" via "don't know" to "no" as these successive observations are added.

From the point of view of diagnostic reasoning these are familiar problems, but I can't think of any work in mainstream actions and change that has addressed nonmonotonicity with respect to observations in a serious way. Except, without knowing the details, I would imagine that the people who do probabilistic or possibilistic approaches might have something to say about this. Judea, or the folks in Toulouse, any input?

Do all current approaches comply with the causality principle? On 21.1, Hector Geffner wrote:

> *Most recent models of action comply with the causality principle. In some it comes for free (e.g., language $\mathcal{A}$ due to the semantic structures used (transition functions); in others (Reiter, Sandewall, etc), I'm sure it can be proved.*

Yes, with respect to the features and fluents approach, it also "comes for free". The underlying semantics that's used there is essentially a simulation of the world, using non-deterministic transition from state to state or from state to state sequence (the latter in order to account for actions with extended duration). In either case the simulation proceeds forward in time, so it complies with the causality principle. The same applies for the extensions of the approach to deal with concurrency, ramification, and continuous change.

Erik Sandewall

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/`