ELECTRONIC NEWSLETTER ON
REASONING ABOUT ACTIONS AND CHANGE

Issue 00008 Editor: Erik Sandewall 30.12.2000

Today

Several articles have been accepted since the latest newsletter.

For inclusion in Section A of ETAI Volume 3 (1999):

Tom Costello and John McCarthy: Useful Counterfactuals. Accepted on December 20.

For inclusion in Section E of ETAI Volume 3 (1999), Special Section on Non-Monotonic Reasoning, Actions and Change (Section Editor: Michael Thielscher):

Giuseppe De Giacomo and Riccardo Rosati: Minimal Knowledge Approach to Reasoning about Actions and Sensing. Accepted on December 5.

Fiora Pirri and Alberto Finzi: An Approach to Perception in Theory of Actions: Part I. Accepted on December 5.

Congratulations to all these authors! The referee reports follow below for those cases where the referees have additional comments besides plain 'yes' and 'no' to the refereeing questions.

Today we also have a set of questions by Graham White to Tom Costello and John McCarthy for their article. Because of the length of Graham's message, it will be set up as an ENRAC note. It follows here in full.

ETAI Publications

Questions by Referees

An Approach to Perception in Theory of Actions: Part I

by **Fiora Pirri and Alberto Finzi**

Comment from: **Anonymous Referee 1**

1. are the results of the article, as specified in the summary, of significant interest for researchers in our area?

yes

2. does the full text substantiate the claims made in the summary?

yes

3. is the text intelligible and concise?

yes

4. do you know of any previous publication of the same result?

no

5. are there any other considerations that may affect the decision?

I just found the remark on p.11 irritating: The successor state axiom for Percept is not per se nondeterministic, say, in the sense of [Lin, 96]. In fact the axiom can be rewritten to

```
Percept(isf,pr,do(a,s)) <=> ( (pr=1 & delta+) v (pr=0 & delta-) )
                            v
                            Percept(isf,pr,s) &
                               not ( (pr=1 & delta-) v (pr=0 & delta+) )
```

which is the standard form of a successor state axiom.

Formula (7) should be added to the conditions of Lemma 1. What is the role of the the second argument for isInTrolley on p.31?

[Lin, 96] Fangzhen Lin. Embracing Causality in Specifying the Indeterminate Effects of Actions. AAAI'96.

Minimal Knowledge Approach to Reasoning about Actions and Sensing

by **Giuseppe De Giacomo and Riccardo Rosati**

Comment from: **Anonymous Referee 2**

1. are the results of the article, as specified in the summary, of significant interest for researchers in our area?

yes

2. does the full text substantiate the claims made in the summary?

yes

3. is the text intelligible and concise?

yes

4. do you know of any previous publication of the same result?

no

5. are there any other considerations that may affect the decision?

The authors may wish to comment on the following observation: It seems that the inference algorithm requires instances of the frame axiom schema for disjunctions of fluents in case of incomplete states. Are there good reasons why one is not forced in practice to feed your algorithm with exponentially many frame axioms?

Questions to Tom Costello and John McCarthy

Question 1.

Costello and McCarthy, 10.1: "Theories given by differential equations give some clearcut examples.

The solutions are determined by boundary conditions. If the theory includes both the differential equations and the boundary conditions, the most obvious counterfactual is to keep the differential equations and change the boundary conditions to a different set of admissible boundary conditions.

A simple example is provided by the equations of celestial mechanics regarding the planets as point masses."

Here is a question. Consider a classical, deterministic dynamical system. In such a system, the future evolution is given by the positions, and velocities, of its components at a particular time. The space of all position-velocity pairs – usually known as phase space – will be our "space of possible states" (Costello and MacCarthy 3.2). So the question is: can we set up coordinates on phase space, which are coordinates in MacCarthy and Costello's sense (that is, which allow for their sort of counterfactual reasoning) and for which time is one of the coordinates?

Suppose so. Then time would be one of the coordinates, and would thus vary along trajectories of the system. The other coordinates would, by the orthogonality requirement, differentiate between trajectories. MacCarthy and Costello's orthogonality requirements would thus amount to the following: i) if we change the time coordinate, but keep the other coordinates fixed, then we move along a fixed trajectory, and ii) if we fix the time coordinate, but change the others, then we vary the trajectory.

Let us call the other coordinates the trajectory coordinates. We can regard the trajectory coordinates as functions on phase space. The trajectory coordinates, by requirement i), would have to be constant along trajectories: by requirement ii) – and by the fact that they are genuinely coordinates – there would have to be enough of them so that any two distinct trajectories differed in at least one of the trajectory coordinates.

Now functions like these trajectory coordinates are well known in classical mechanics: they arose in the study of a similarly natural problem in dynamics, namely the problem of finding closed-form solutions to systems of differential equations. Functions on the phase space which satisfy requirement i) are known as integrals of a dynamical system ([1], pp. 79ff.; cf.

[2]). Integrals are, however, very rare: most systems do not have enough of them to differentiate between trajectories. Consider the example which Costello and MacCarthy use, that of "the equations of celestial mechanics regarding the planets as point masses." Even with only three particles, the equations do not have enough integrals: the only constants of the motion are the momentum, and angular momentum, of the system, together with the total energy. Trajectories with fixed values of these quantities lie in seven-dimensional subspaces of the phase space ([1], pp. 98ff.): there is still far more variation possible than can be described by values of the integrals.

(The fact that integrals are hard to find can be explained by Noether's theorem, which relates integrals to symmetries of the dynamical system: most systems are not symmetrical enough to have many integrals, and it is only those which are exceptionally symmetrical which are completely integrable.)

So, for most dynamical systems, we cannot find a set of Costello-MacCarthy coordinates in which time is a coordinate. Which means that (strangely enough) we cannot apply their approach to the semantics of assertions like "if the ball is here now, it will be there in a second": that is, assertions which, on the basis of observations of a state of the system, predict another state of the system which differs only in distance along the trajectory, all other things being fixed.

Question 2a.

Does Costello and MacCarthy's theory have a semantics, in Quine's sense?

Quine, in [3], outlines an approach to semantics which is presented as a story about language acquisition (either by an infant, learning from the adults around her, or by an anthropologist, learning from a strange tribe); the semantics, however, tells us about more than merely language acquisition.

Quine distinguishes between two sorts of terms: there are those, such as 'water' and 'red', which are purely phenomenal and do not involve any grasp of conditions of identity. Correct application involves merely being able to respond to the appropriate set of stimuli [3, p. 7]. On the other hand, there are terms, such as 'apple', the use of which involves a grasp of conditions of identity: "To learn 'apple' it is not sufficient to learn how much of what goes on counts as apple; we must learn how much counts as an apple, and how much as another" [3, p. 8]. Terms such as these, says Quine, are those "whose ontological involvement runs deep" [3, p. 8], the terms from the use of which a speaker's implicit ontological commitments can be recovered. Mastery of these terms involves being able to "ask whether this is the same so-and-so as that, and whether one so-and-so is present or two." [3, p. 2]

Crucial among these abilities is being able to re-identify the same object under two different descriptions. How does this apply to Costello and MacCarthy's talk of the possible states in their "space of possible states"? Well, if we were to encounter the same state again, under another description, we may well do so using another descriptive vocabulary, which would put another coordinate system on the space of possible states. So – if we follow Quine's line of reasoning – we seem to be asking whether there are any allowable coordinate transformations on the space of possible states. Costello and MacCarthy argue, it is true, that some coordinate transformations on this space are inadmissible (because they alter the counterfactual properties of the states in this space): but, if we are to follow Quine, there must also be admissible transformations. For if not, then we cannot really be talking of objects when we talk of these states: they would have no coherent conditions of identity.

Question 2b.

Related to Question 2a is the folowing, more practical question: it uses an example of Davidson [6], who gets it from Austin [7]. It is best given by an extended quote.

" 'I didn't know that it was loaded' belongs to one standard pattern of excuse. I do not deny that I pointed the gun and pulled the trigger, nor that I shot the victim. My ignorance explains how it happens that I pointed the gun and pulled the trigger intentionally, but did not shoot the victim intentionally. ... What is the relation between my pointing the gun and pulling the trigger, and my shooting the victim? The natural ... answer is that the relation is that of identity. The logic of this sort of excuse includes, it seems, at least this much structure: I am accused of doing b, which is deplorable. I admit I did a, which is excusable. My excuse for doing b rests upon my claim that I did not know that a = b." [6, p. 109]

Thus: an important part of commonsense reasoning is to be able to describe the same events using different vocabularies. Our descriptions of those events may, of course, apprehend certain features or other of them: but there is a notion of identity of events, which we seem to use in everyday reasoning, which transcends difference of vocabulary. So how can we accomodate this sort of reasoning in Costello and MacCarthy's framework?

Question 3a.

Einstein writes "If, relative to K, K' is a uniformly moving coordinate system devoid of rotation, then natural phenomena run their course with respect to K' according to exactly the same general laws as with respect to K". [5, p. 18] Principles like these – which assert that the laws of physics are invariant under transformations of coordinates – are, according to the story which Einstein tells, fundamental to modern physics: progress in physical understanding is accompanied by the enlargement of the group of transformations which the laws of physics are invariant under (first galilean transformations, then Lorenz transformations, then, with general relativity, curvilinear coordinate transformations which preserve the Lorenz metric).

Is a similar outlook possible with Costello and MacCarthy's framework? Are we allowed to speak of coordinate transformations? Would an increase in our understanding of the commonsense laws of nature be accompanied by a parallel increase in the number of allowed coordinate transformations?

Question 3b.

This same point has practical consequences: in physics, we want to change coordinates, either because we can solve the equations more easily thereby, or because the new coordinate system illuminates some feature of the situation which we want to concentrate on. Thus, the study of celestial mechanics was accompanied by the intensive use of coordinate transformations: partly these were to permit solutions (one would try to find coordinate systems in which the motions could be expressed as sums of trigonometric functions) and partly they were for conceptual reasons (one wanted to investigate the stability of the solar system, and thus one would try to find coordinate systems which separated periodic changes from the so-called "secular", non-periodic changes). [8, pp. 15ff.]

If Costello and MacCarthy's counterfactuals are to be "useful", are we allowed, in searching for solutions, to change our vocabulary to one which is useful for us – either because it permits an easier solution, or because it lays bare some feature of the situation which we want to concentrate on?

[1] V.I. Arnol'd, V.V. Kozlov, and A.I. Neistadt, "Dynamical Systems III: Mathematical Aspects of Classical and Celestial Mechanics", vol. 3 of the Encyclopaedia of Mathematical Sciences, Springer 1988.

[2] G. White, "Lewis, Causality and Possible Worlds", Dialectica 54 (2000), pp. 133-137.

[3] W.v.O. Quine, "Speaking of Objects", in [4, pp. 1-25].

[4] W.v.O. Quine, Ontological Relativity and Other Essays (Columbia 1969).

[5] A. Einstein, Relativity: The Special and the General Theory (London 1962)

[6] D. Davidson, "The Logical Form of Action Sentences", in D. Davidson, Essays on Actions and Events (Oxford 1980), pp. 105-122.

[7] J.L. Austin, "A Plea for Excuses".

[8] J. Barrow-Green, Poincare and the Three Body Problem (American Mathematical Society 1997)

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.etaij.org/rac/`