Electronic Transactions on Artificial Intelligence
Vol. 1 (1997): *105-128*
http://www.ep.liu.se/ej/etai/1997/005/

Refereed Article

Erik Sandewall:

Logic-Based Modelling of Goal-Directed Behavior

Author's affiliation: Linköping University, Sweden.

Abstract: *We address the problem of characterizing goal-directed robotic behavior using a logic of actions and change. Our approach is based on distinguishing two kinds of actions:* procedural *actions which are defined in a mechanistic way, and* goal-directed *actions which are performed through a process involving tries, possibly failures, and corrective action and new tries until the goal has been reached. (The definition of procedural actions may be done external to the logic, for example through differential equations, or through a conventional programming language). For both kinds of actions, the logic expresses explicitly whether the action* succeeds *or* fails. *Each execution of a goal-directed action is also characterized by a number of* breakpoints *where some sub-action has been completed and a new sub-action for getting to the desired goal is selected. The logic is used for characterizing the selection of sub-actions at breakpoints, and the success or failure of the goal-directed action in terms of the success or failure of the sub-actions.*

The article describes how goal-directed actions can be modelled by an extension of existing results on logics of actions and change.

Publication and review history

1. First version *published* by Linköping University Electronic Press on 19.12.1997 and permanently available at
   http://www.ep.liu.se/ea/cis/1997/019/

2. *Received* by the research area "Reasoning about Actions and Change" of the Electronic Transactions on Artificial Intelligence, posted on its web site, and advertised in its Newsletter on 19.12.1997 (The summary was included in the article).

3. Publicly *discussed* in the ETAI area of Reasoning about Actions and Change during January - April, 1998. The discussion protocol is available and will be retained in the Article Interaction Page at
   http://www.ida.liu.se/ext/etai/received/rac/006/aip.html

4. Revised version, with minor corrections compared to the first version, *published* by Linköping University Electronic Press on 15.4.1998 and permanently available at
   http://www.ep.liu.se/ea/cis/1997/019/
   under the label "Revised publication 1998-04-15".

5. Revised version *accepted* after due refereeing and according to scientific journal standards by the Electronic Transactions on Artificial Intelligence (ETAI) on 13.7.1998. Comments by the referees are included in the Article Interaction Page.

6. Final version, with minor corrections according to the suggestions of the referees, *published* by Linköping University Electronic Press on 29.7.1998 and permanently available at
   http://www.ep.liu.se/ea/cis/1997/019/
   under the label "2nd revised publication 1998-07-29".

7. Final version appears in Electronic Transactions on Artificial Intellligence, Volume 1 (1997), Issue 4, pages 105–128. The issue and the annual volume are made permanently available at
   http://www.ep.liu.se/ej/etai/1997/

8. Paper editions of the ETAI issue and the ETAI volume containing this artice *republished* by the Royal Swedish Academy of Sciences. For purchase of this edition, please refer to the URL mentioned in the previous item.

The review policy and the quality requirements for acceptance are documented at
   http://www.ida.liu.se/ext/etai/info/

Article maintenance

The Electronic Transactions on Artificial Intelligence maintains an Article Interaction Page for this article at
   http://www.ida.liu.se/ext/etai/received/rac/006/aip.html
Besides the publication and review history, it contains links for contacting the author(s), as well as to amendments and other related information for the article. It is intended to keep these links up-to-date in the future.

Conference publication

The first revised version (as of 15.4.1998) of this article was also published in the conference proceedings of KR-98, the Sixth International Conference on Principles of Knowledge Representation and Reasoning, pages 304-315. These proceedings were edited by Anthony G. Cohn, Lenhart Schubert, and Stuart C. Shapiro, and published by Morgan Kaufmann Publishers Inc.

Copyright conditions

Summary

ETAI authors are recommended that each article be accompanied by a summary. Longer and more specific than a conventional abstract, it should specify in concrete terms what are the new results in the article. If present, the summary also plays a role in the refereeing process: referees are asked to judge whether the results as specified in the summary *are of importance to the field, and whether the body of the article* gives sufficient evidence *for the claims made in the summary. – The Editor.*

We address the problem of characterizing goal-directed robotic behavior using a logic of actions and change. The logicist formalization is required to be such that the set of models is exactly the set of acceptable goal-directed behaviors in the application at hand, each model being a representation of a possible history of the world. It is not required from the logic that it should *represent* psychologically related concepts, such as goals or intentions, but merely that the resulting behavior obtained from the formalization should have the characteristics of being goal-directed.

Frame of reference and motivation

We view intelligent robotic behavior as something which has both a reactive level and a deliberative level. The deliberative level is in charge of such things as receiving instructions from the commander, pursuing goals, predicting the possible outcomes of intended actions, avoiding dangers, diagnosing faults, planning means of acquiring knowledge, and so on. The underlying, reactive level may be characterized by stimulus-response behaviors and by program-like composition of simpler behaviors into more complex ones.

Logic is clearly relevant for the design and analysis of the deliberative level of such a design. However, the logic being used then must be able to characterize its different functionalities, including its goal-directedness. It must also be able to characterize the underlying, reactive level, not in every detail, but with sufficient precision for the needs of the deliberative layer.

Procedural vs goal-directed actions

The present article focuses on the characterization of goal-directed behavior as one aspect of an intelligent robot. However, we do not introduce "goals" as a separate logical construct. Instead, our approach is to distinguish two kinds of actions: *procedural* actions which are defined using a subroutine, and *goal-directed* actions which are performed through a process involving tries, possibly failures, and corrective action and new tries until the goal has been reached. Procedural actions may therefore be used for representing reactive-level behaviors to the deliberative layer; goal-directed actions characterize processes within the deliberative layer itself.

Success vs failure of actions

For both kinds of actions (goal-directed and procedural ones), the logic expresses explicitly whether the action *succeeds* or *fails*. Each execution of a goal-directed action is also characterized by a number of *breakpoints* where some sub-action has been failed, and another sub-action is selected for one more attempt to arrive at the desired goal. The logic is used for

characterizing the selection of sub-actions at breakpoints, and the success or failure of the goal-directed action in terms of the success or failure of the sub-actions.

The term "goal" is used in the sense used in AI planning, that is, as the concrete goal (goal state) that is to be achieved by a plan = a sequence of actions. We do not use the word in the sense of "general goal" (goal in life).

Formalism for actions and change

We use a logic of time and action, that is, a multi-sorted first-order logic where points in time is one of the sorts. The following predicates are used:

$\mathsf{H}(t, p)$: the "propositional fluent" p holds at time t.

$\mathsf{X}(t, f)$: the fluent f is occluded at time t

$\mathsf{G}(s, a)$: the action a is invoked ("go") at time s

$\mathsf{A}(s, a)$: the action a is applicable at time s

$\mathsf{D}_{\mathrm{s}}([s, t], a)$: the action a is executed successfully over the time interval $[s, t]$: it starts at time s and terminates successfully at time t.

$\mathsf{D}_{\mathrm{f}}([s, t], a)$: the action a is executed but fails over the time interval $[s, t]$: it starts at time s and terminates with failure at time t.

In fact, these predicates are viewed as abbreviations for underlying, more elementary constructs indicating the applicability and the failure of actions.

Actions are composed using the conventional operators such as sequential composition (;), conditionals, etc.

The article proposes a set of axioms that characterize possible structures using these and related predicates.

Formalism for goal-directed behavior

A few specialized relations are introduced for characterizing the goal-directed behavior. The key relation is:

$\mathit{Option}(s', g, s, a)$: while pursuing the goal set by invoking the goal-directed action g at time s', one is now at a timepoint s (a breakpoint), and one considers performing the action a. This action may be a composite one, which means that it has the character of a plan. Several follow-up actions may be considered at the same point in time.

The article proposes a set of axioms that characterize goal-directed behavior in terms of this predicate and the ones mentioned under the previous heading.

Entailment methods

A logic characterizing goal-directed use of actions must deal with multiple knowledge sources, including action laws characterizing the effects of actions, observations of the state of the world, plans for achieving specific goals, and so forth. Each of these aspects of the system may call for some kind of nonmonotonicity in order to be sufficiently selective. It is not trivial how to combine the different types of nonmonotonicity that are required. The present article proposes and uses a general approach for dealing with this problem.