AAMAS 2010 TORONTO

Toronto, Canada
May 10 - 14, 2010

THE NINTH INTERNATIONAL CONFERENCE ON AUTONOMOUS AGENTS AND MULTIAGENT SYSTEMS

print

User: Lin Padgham (linpa)
- Start page
- Edit user data
- Logout

PCMember
- Reviewer Management
- Ballot and discussion

Author
- View own papers

Review Feedback and Response

Paper ID	170
Paper authors	Dhirendra Singh, Sebastian Sardina, Lin Padgham, Stephane Airiau
Paper title	Learning Context Conditions for BDI Plan Selection
Paper subtitle	
How are the claims of the paper evaluated?	Empirically (e.g., experiments)
Paper status	First Ballot
Keywords	Agent theories, Models and Architectures::BDI ** Agent-based system development::Agent programming languages ** Learning::Learning (single and multi-agent)
Abstract	An important drawback to the popular Belief, Desire, and Intentions (BDI) paradigm is that such systems include no element of learning from experience. In particular, the so-called \emph{context conditions} of plans, on which the whole model relies for plan selection, are restricted to be boolean formulas that are to be specified at design/implementation time. To address these limitations, we propose a novel BDI programming framework that, by suitably modeling context conditions as \emph{decision trees}, allows agents to \emph{learn} the probability of success for plans based on previous execution experiences. By using a probabilistic plan selection function, the agents can balance exploration and exploitation of their plans. We develop and empirically investigate two extreme approaches to learning the new context conditions and show that both can be advantageous in certain situations. Finally, we propose a generalization of the probabilistic plan selection function that yields a middle-ground between the two extreme approaches, and which we thus argue is the most flexible and simple approach.
Download	

Comments to author(s)

SUMMARY:

The paper proposes and evaluates an approach to learning context conditions for plans in a BDI agent program.

RELEVANCE:

The topic -- development of more robust agent programs -- is very relevant to AAMAS.

ORIGINALITY:

The paper extends previous work, but the contribution is clearly identified.

SIGNIFICANCE

While the work is interesting, I am not convinced of its significance. The approach presented is limited to environments in which unsupervised learning is feasible; if actions are hazardous, non-reversible, costly or simply time consuming, it would be difficult to apply. It's hard to think of application domains in which an agent could learn the context conditions of its plans in the manner proposed (no examples are given in the paper). Even if the environment is amenable to the approach (or we can simulate the effects of the agent's actions in the environment), the number of trials necessary to learn context conditions for relatively simple plans in very simple environments seems quite high.

TECHNICAL QUALITY

The work presented is technically sound. The first part of the evaluation is very clear, but I am less convinced by the evaluation in the final part of section 5 and in section 6. In particular, the assumption (on p. 6, col 1, line 6) that actions are deterministic would seem to make it impossible to directly compare the results given in figures 4 and 5. If these results are directly comparable, please say why.

Also, there are no complexity results for the informed plan selection approach presented in section 6. With non-trivial plans and environments, the number of trials necessary to ensure coverage can be very large. Even it is possible to arrange that the agent encounters all possible world states (for each plan), it may take a very long time (in general) to converge to the correct plan choice since the plan selection weight grows linearly with coverage. However I may be missing something.

QUALITY OF PRESENTATION

The paper is generally well organized and well written. The first part of the paper is particularly clear, and is a pleasure to read. However I found the material in the latter part of section 5 and section 6 more difficult to understand.

The relationship between the ACL and BUL approaches in the first part of section 5 and those in the rest of the paper is not very clear. Are the later approaches to be considered alternatives to those presented in the first part of section 5 (in which case when should they be used), or are they improvements? If they are improvements, why present the initial approaches at all?

I was also a bit confused by the relationship between stability and confidence. These two measures seem to be doing quite similar things -- would it not be possible to use the stability metric directly in choosing plans rather than add the extra requirement that a plan should be tried in a (significant fraction of) all world states? However I may have misunderstood the intent of section 6 and the results, which I found hard to understand.

There were also some more minor issues.

On page 2, in the discussion of plan rules, is it possible to extend the approach to rules without event goals (plans which are triggered by changes in the agent's beliefs)?

On page 4, it would be helpful to provide a definition of the rate of success of a plan. In section 5, it would be helpful to clarify possible structure of the decision trees -- is the set of possible tests (internal nodes) the set of propositional fluents used to represent the environment? If so, it would be helpful to say how this scales to large numbers of fluents (many of which could be assumed to be irrelevant to the success or otherwise of a given plan). Also what is meant by "posting the top-level goal repetitively under random world states": does this mean posting the goal repeatedly following plan success or failure, with a different initial world state on each trial?

On page 6, bottom of column 2, the argument that a plan is only executed in the subset of world states that are relevant to it seems circular, since the aim is to determine precisely for which world states a plan is relevant.

Finally, I was surprised to see no mention of learning in teleo-reactive programs in the related work section (or even chunking in soar).

Summary of review
An interesting paper. However the approach presented seems of limited applicability.

Comments to author(s)
SUMMARY:
The paper presents an extension to the BDI architecture to allow agents to learn which plan they should execute next. The agents lean based on the past execution of their plans in given world states. Each plan is associated with a decision tree that classifies world states into an expectation of whether the plan will succeed or fail.

RELEVANCE:
The paper presents a theoretical work that is relevant to the development of agents based on BDI and able to learn the most appropriated plan to be executed in the next state.

ORIGINALITY:
The authors should include a related work section moving part of the discussion section to this new section. In addition, it should be clarified the contribution of the paper faced the related work.

The authors should also look for approaches that, for instance, consider that agents can self-adapt their behavior due to failures in plans and for approaches that provide mechanisms that help agents on understanding the failures and on selecting new plans based on such feedback. Although such approaches are not directly related, those agents are able to learn that a given plan is not suitable in a given situation (world state).

SIGNIFICANCE
The work present is important because it addresses a basic problem found in the most used agent architecture.

TECHNICAL QUALITY
Section 3:
- I have not understood the need for assuming that there is not external change in the environment during execution other than those made by the agent. Fist of all, it is not realistic. Second, although knowing that a plan have succeed or failures in a given world state it does not mean that it will for such happen again. Thus you should be able to deal with such unpredictability.
- Why aren't you dealing with "automated" failure handling? When an agent faces a problem while executing a given plan, why must such failure be propagated to the root goal? (related to the comment below)

Section 4:
- Why can't the agent solve the problem before propagating it?
- In order to achieve a subgoal, only one plan must succeed, the others can fail. Thus, why is a goal considered stable for a state w only if ALL its relevant plans are table for w?
- It would be interested to see algorithms, one considering ACI and another BUL in order to clarify their difference.

QUALITY OF PRESENTATION
The paper is well written and presented.

Summary of review
The problem covered in the paper is relevant and well presented. Although the contributions are well described, the comparison with the related work should be clarified.

Comments to author(s)
SUMMARY:

The paper describes how context conditions of BDI plans can be modeled with decision trees. Decision trees are used to learn the applicability of plans at run-time. The authors introduce a probabilistic plan selection function and empirically evaluate different learning strategies.

RELEVANCE:

The paper is relevant.

ORIGINALITY:

The idea of using decision trees for learning context conditions has already been published earlier. However, the authors extend this work and their contribution is clear. The related work section should be extended (more on the relation to the previous work).

SIGNIFICANCE

In my opinion the paper is significant since the adaptation of plan selection during run-time allows the agents to re-/act and adapt more flexible to their environments. The paper extends state of the art regarding a probabilistic plan selection function and different learning strategies. Due to the limitations and assumption mentioned in the paper I am sceptical how this approach works out in real world systems. Nevertheless, the approach is current, interesting, and extends an existing approach.

TECHNICAL QUALITY

The paper is sound and the authors clearly point out and discuss the limitations, assumptions, and restrictions of their approach. In my opinion, the assumptions made in the paper are too strong for real world systems.

Section 4: write 2-3 sentences why k=3 and epsilon=0.3 has been chosen.

QUALITY OF PRESENTATION

The paper is well written and structured. The paper seems to be closely related to [1], so this alignment should be discussed in more depth.

Some minor things:

"Thus, for the case where ACL+O was performing reasonably well, the new ACL+O approach..." (prime is missing).

Categories, General Terms, and Keywords are missing.

Summary of review
The paper describes how context conditions of BDI plans can be modeled with decision trees (they extend previous work). The contribution is clear, however the made assumptions/restrictions are strong.