(dhirendra.singh@rmit.edu.au) | IJCAI 2011 --- User Logout

Twenty-Second International Joint Conference on Artificial Intelligence

IJCAI 2011

Author Response

Title: Integrating Learning into a BDI Agent for Environments with Changing Dynamics
Authors: Dhirendra Singh, Sebastian Sardina, Lin Padgham and Geoff James

Instructions

The author response period has begun. If you want to respond to the points raised in the reviews, you may do so in the box provided below.

You do not have to respond to the reviews. However, reviewers often treat failure to do so negatively if they have made any negative comments.

Your overall scores can be interpreted as follows:

10. This could win the best paper prize.
9. I would be willing to resign if this is not accepted.
8. This is a strong paper.
7. This is a good paper.
6. This would make an acceptable paper.
5. This would make an acceptable poster.
4. There are reasons to reject this.
3. There are many reasons to reject this.
2. This paper is not publishable.
1. If IJCAI had summary rejects, this would be a candidate.

Review #1

Relevance: 9
Significance: 8
Soundness: 8
Originality: 8
Evaluation and comparison: 6
Clarity: 8
OVERALL SCORE: 7

Comments

A BDI plan-selection learning mechanism is described that uses a dynamic confidence measure to judge an agent's current understanding (of how well plans solve goals) and the need for exploration to improve that understanding. The mechanism is designed for dynamic environments where the confidence in a plan does not increase monotonically. The work builds on the BDI learning approach of Singh, et al., and preserves the monotonic convergence of that approach in static environments.

I liked this paper. The extensions to BDI plan-selection learning are interesting and important, and the paper is generally clear and well written. The experiments section presents the behavior of using the extended mechanism in a modular battery controller setting. The presented results show the capabilities of the mechanism, but I would have liked to see a comparison with the performance obtained (or not) using a Singh-like monotonic learning mechanism in the same dynamic setting.

The abstract needs improvement/specificity (such as including some of the summary information in the first paragraph above as well as the contributions claimed).

Make clear that the "they" in the second sentence of the Introduction refers to the "classic" or "original" BDI paradigm (as when I first read the paper, I didn't realize until later in the Introduction that the authors were even aware of other BDI learning work).

I wanted more discussion of value(s) for "n" in the confidence measure (and indirectly in the final plan selection weight). For example, what "n" value(s) was used in your experiments? What adjusts "n" if the rate of dynamic change changes (what is the effect of using an inappropriate value of "n")? Was any thought given to using temporally reduced weighting (such that old experience carries less weight) rather than the "n" based cutoff with uniform weighting?

Minor nit: eliminate the space between "trace" and footnote reference 1.

The citations and discussion of other work are reasonable, as is the mention of limitations and possible work in the Conclusion. Again, I liked this paper.

Review #2

Relevance:	7
Significance:	4
Soundness:	6
Originality:	4
Evaluation and comparison:	4
Clarity:	7
OVERALL SCORE:	4

Comments

SUMMARY The work in this paper extends the existing BDI learning framework: it proposes a heuristic (the equation towards the end of Section 3.2) measuring the confidence of plans that automatically adjusts with the environment dynamics. The running examples are great for our understanding of the authors' work. I'm especially concerned by the significance and thoroughness of the experimental results presented in this paper. This work can be improved based on my feedback below:

DETAILED REVIEW

ORIGINALITY: The authors propose a minor fix to the existing BDI learning framework so that it can be used in a dynamically changing environment. The proposed simple idea does contribute incrementally towards agent programming research (but not towards machine learning research, as indicated by the authors). See the next few comments to see how this work can be improved.

SIGNIFICANCE, EVALUATION, AND COMPARISON: (1) The authors have mentioned the importance of addressing the exploration vs. exploitation dilemma in section 2.2 using the confidence measure. However, I notice that only the exploitation part is addressed in section 3.3. The exploration part seems to be missing, which, if included, will raise the significance of this work.

(2) I'm particularly concerned by the thoroughness of the experimental results, which can be significantly improved: (a) From Fig. 2, it appears that the learning time is long. I recommend that the authors conduct experiments or provide theoretical results analyzing the tradeoff(s) between the degree of change of environment dynamics and speed of learning and size of state space. (b) The authors did not provide experimental results comparing its proposed technique with the state-of-the-art BDI learning techniques. I recommend that the authors compare with the one-off learning techniques of Singh et al. (2010a), (2010b) by introducing a fixed re-learning rate.

TECHNICAL SOUNDNESS: (1) I understand from the paragraphs before Algorithm 1 that the stability degree is recorded in a plan only for failed execution traces. If this is the case, is it possible for the stability degree lambda to converge to 1 for fixed dynamics (as indicated in the paragraph after Algorithm 1) without taking into account the successful execution traces and the successful stable (possibly non-leaf) plans? I would like the authors to address this in the rebuttal.

(2) The authors have said that "When a plan finally succeeds, we take an optimistic view and record degree 1 against it". Does this apply to non-leaf plans that succeed (though they may fail in other execution traces) as well? I would like the authors to address this in the rebuttal.

(3) In footnote 3, I'm surprised to see that the value of alpha is not critical to performance. If I set alpha to either of the extremes (0 or 1), one of the confidence measures is dispensable. I would like the authors to address this in the rebuttal.

(4) It would be clearer if the authors elaborate on the relationship between the RecordDegreeStability function in Algorithm 1 and the average degree of stability in the text. Otherwise, a reader may be mistaken from Algorithm 1 that the current stability degree is entirely replaced by the newly-computed stability degree for plan P1 instead of computing their weighted average.

(5) The authors say that "NewStates(P, n) is the set of world states in the last n executions of P that have also been witnessed before that". Shouldn't this be corrected to "... have NOT been witnessed..."?

CLARITY: (1) This paper is clearly written except for some parts in Section 3.2.

(2) The last paragraph of the conclusion should be put in front as a separate section on related work.

Review #3

Relevance:	10
Significance:	9
Soundness:	9
Originality:	9
Evaluation and comparison:	10
Clarity:	10
OVERALL SCORE:	9

Comments

Relevance:

The paper is extremely relevant to the special track on Integrated and Embedded Artificial Intelligence. It presents an integration of machine learning techniques into a BDI Agent Model, which uses a measure of confidence to determine whether the learned plan trees for behaviour are appropriate (favoring exploitation), or not (favoring exploration of new strategies).

Significance and Originality:

One of the interesting aspects of the work presented is that the confidence measure does not converge monotonically to 1, making the agent always to perform exploitation. Departing from traditional approaches, authors assume that the learned behavior can sometimes become wrong and inappropriate in a dynamic environment where rules are not constant. To solve this, authors put forward a dynamic confidence measure that adapts accordingly with changes in performance.

Soundness:

The work presented, to the best of my knowledge seems to be sound and correct. I managed to follow and understand the formulas and formal definitions presented. However, since my background is not in machine learning, it is hard for me to properly evaluate their soundness.

Evaluation and comparison:

The evaluation performed seems very appropriate and I think it managed to prove the proposed plan-selection learning mechanism is able to adapt when the dynamics of the environment change over time.

Clarity:

The paper is clear and well written, with a clear description of the research goal and their contribution to the area. Although formal in some parts, authors manage to make the paper quite easy to read and follow by providing a set of illustrative and meaningful examples.

Submit Response

Use the following box to enter your response to the reviews. Please limit your comments to 750 words (longer responses will not be accepted by the system).

Submit