Learning Context Conditions for BDI Plan Selection

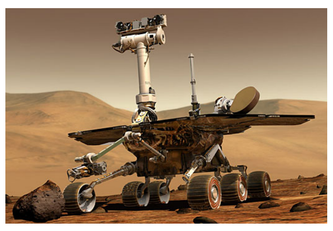
Dhirendra Singh[1] Sebastian Sardina[1] Lin Padgham[1] Stéphane Airiau[2]

[1]School of Computer Science & Information Technology, RMIT University, Australia
[2]Institute for Logic, Language and Computation, University of Amsterdam, The Netherlands

Summary

We address the plan selection problem in Belief, Desire, Intentions (BDI) Agent Systems.

Context conditions of plans determine applicability in given situations, and must be specified upfront. However, new environments often require learning changes to selection conditions.



Easing this constraint would allow conditions to be refined once deployed, improving adaptability.

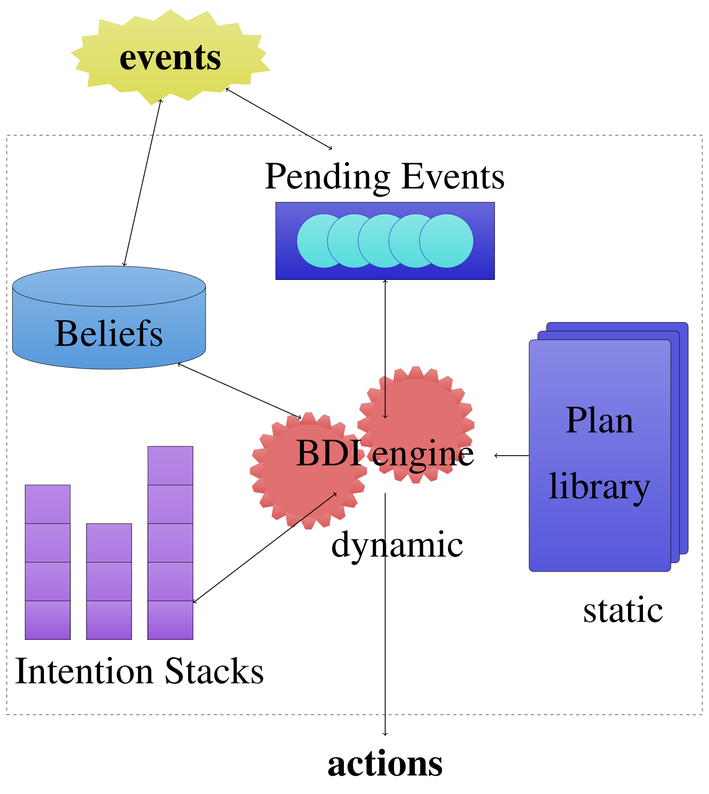
Our learning framework augments plan's context conditions with decision trees, allowing plan applicability to be learnt from experience.

Using a probabilistic plan selection function, the agent balances exploration and exploitation of plans, while learning online.

BDI Architecture

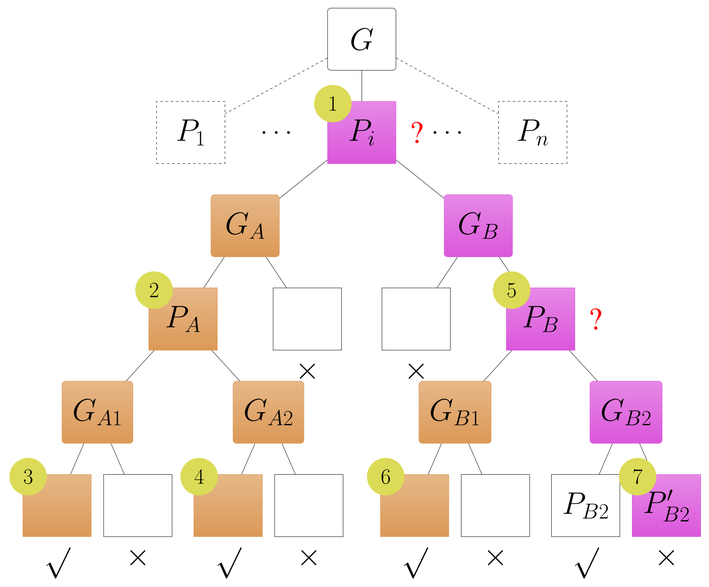
A plan is a rule $e : \psi \leftarrow \delta$; program δ is a strategy for goal e when condition ψ holds.

The burden for the programmer is to perfectly design the logical formula ψ.



Plans perform primitive actions or post subgoals to be handled in a hierarchical manner.

Learning Task



1. The imposed BDI hierarchy implies that high level plans may fail not because they were poor choices for the situation but due to poor choices further below.
2. Learning is performed online while acting in the environment, so care must be taken in how much confidence to put in each decision tree on an ongoing basis.

BDI Learning Framework

Each plan's logical formula context condition is augmented with a decision tree.

A probabilistic plan selection function balances exploitation of ongoing decision tree learning and further exploration of the state space.

Record outcomes for chosen plans to train decision trees

Probablistically select plans based on ongoing learning

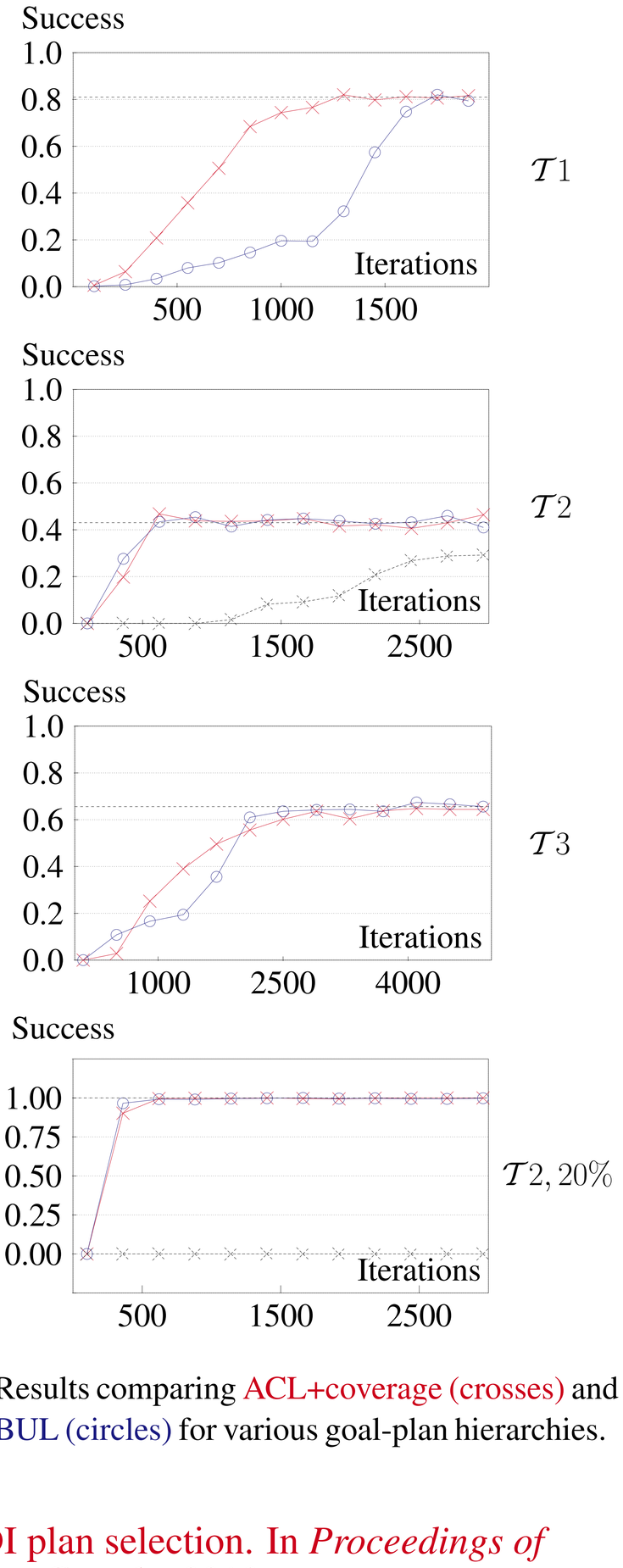
Acting and learning are interleaved. Ongoing learning impacts the choice of future actions that impact subsequent learning and whether a good solution is eventually found.

Experimentation

We study the impact of goal-plan structures on learning performance. We use synthetic hierarchies that model some features of real BDI programs.

How to record training set: We compare two approaches, a conservative one (BUL) that only records failures when all plan choices are considered well-informed, and an aggressive one (ACL) that records all outcomes.

How to use decision trees: A confidence measure is applied to the decision tree prediction to calculate plan selection weights. Confidence is related to the coverage of paths below a plan.



Results comparing ACL+coverage (crosses) and BUL (circles) for various goal-plan hierarchies.

D. Singh, S. Sardina, L. Padgham, S. Airiau, Learning context conditions for BDI plan selection. In *Proceedings of Autonomous Agents and Multi-Agent Systems (AAMAS)*, Toronto, Canada, 2010.