Learning Context Conditions for BDI Plan Selection

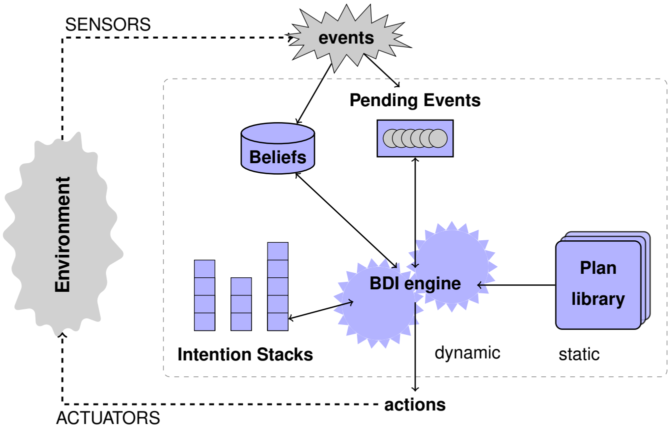
Dhirendra Singh[1] Sebastian Sardina[1] Lin Padgham[1]
Stéphane Airiau[2]

[1]School of Computer Science & Information Technology
RMIT University, Australia

[2]Institute for Logic, Language and Computation
University of Amsterdam, The Netherlands

Autonomous Agents and Multiagent Systems
May 2010

Learning BDI Plan Selection



Plan δ is a strategy to resolve event *e* whenever context ψ holds.
Our focus is the plan selection problem i.e. to learn ψ.

Motivation for Learning

The Belief-Desire-Intention (BDI) model of agency

- Is robust and well suited for dynamic environments.
- Has inspired several development platforms (PRS, AgentSpeak(L), JACK, JASON, SPARK, 3APL and others).
- Has been deployed in practical systems like UAVs.

Motivation for Learning

The Belief-Desire-Intention (BDI) model of agency

- Is robust and well suited for dynamic environments.
- Has inspired several development platforms (PRS, AgentSpeak(L), JACK, JASON, SPARK, 3APL and others).
- Has been deployed in practical systems like UAVs.

Motivation for Learning

The Belief-Desire-Intention (BDI) model of agency

- Is robust and well suited for dynamic environments.
- Has inspired several development platforms (PRS, AgentSpeak(L), JACK, JASON, SPARK, 3APL and others).
- Has been deployed in practical systems like UAVs.

Motivation for Learning

The Belief-Desire-Intention (BDI) model of agency

- Is robust and well suited for dynamic environments.
- Has inspired several development platforms
  (PRS, AgentSpeak(L), JACK, JASON, SPARK, 3APL and others).
- Has been deployed in practical systems like UAVs.

Motivation for Learning

The Belief-Desire-Intention (BDI) model of agency

- Is robust and well suited for dynamic environments.
- Has inspired several development platforms (PRS, AgentSpeak(L), JACK, JASON, SPARK, 3APL and others).
- Has been deployed in practical systems like UAVs.

Nonetheless

- Behaviours (plans) and the situations where they apply (context) are fixed at design time.
- For complex domain, it is difficult to specify complete context conditions upfront.
- Once deployed, the agent has no means to adapt to changes in the initial environment.

Motivation for Learning

The Belief-Desire-Intention (BDI) model of agency

- Is robust and well suited for dynamic environments.
- Has inspired several development platforms (PRS, AgentSpeak(L), JACK, JASON, SPARK, 3APL and others).
- Has been deployed in practical systems like UAVs.

Nonetheless

- Behaviours (plans) and the situations where they apply (context) are fixed at design time.
- For complex domain, it is difficult to specify complete context conditions upfront.
- Once deployed, the agent has no means to adapt to changes in the initial environment.

Motivation for Learning

The Belief-Desire-Intention (BDI) model of agency

- Is robust and well suited for dynamic environments.
- Has inspired several development platforms (PRS, AgentSpeak(L), JACK, JASON, SPARK, 3APL and others).
- Has been deployed in practical systems like UAVs.

Nonetheless

- Behaviours (plans) and the situations where they apply (context) are fixed at design time.
- For complex domain, it is difficult to specify complete context conditions upfront.
- Once deployed, the agent has no means to adapt to changes in the initial environment.

Motivation for Learning

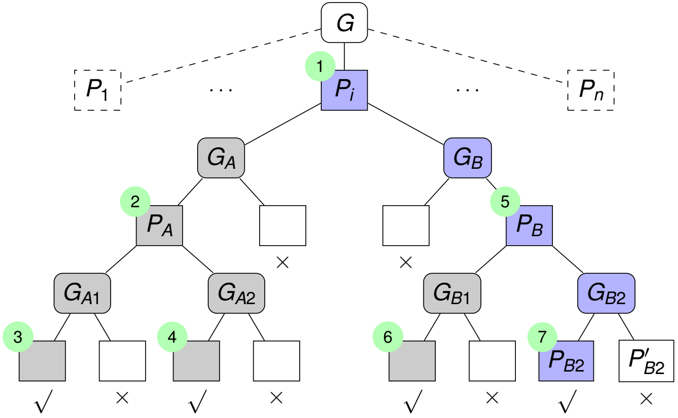
The Belief-Desire-Intention (BDI) model of agency

- Is robust and well suited for dynamic environments.
- Has inspired several development platforms (PRS, AgentSpeak(L), JACK, JASON, SPARK, 3APL and others).
- Has been deployed in practical systems like UAVs.

Nonetheless

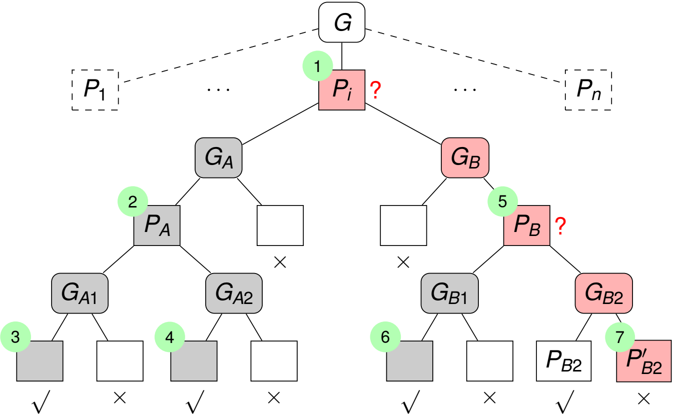
- Behaviours (plans) and the situations where they apply (context) are fixed at design time.
- For complex domain, it is difficult to specify complete context conditions upfront.
- Once deployed, the agent has no means to adapt to changes in the initial environment.

Learning From Plan Choices



Execution trace for successful resolution of goal G given world state w.
Success means that all correct choices were made.

Learning From Plan Choices



Possible execution trace where goal G is not resolved for w.
Should non-leaf plans consider this failure meaningful?

Learning Considerations

1. Collecting training data for learning

- ACL: Aggressive approach that considers all failures as meaningful.
- BUL: Conservative approach that records failures only when choices below are considered to be well-informed.
- Success is always recorded for both approaches.

Learning Considerations

1. Collecting training data for learning

- ACL: Aggressive approach that considers all failures as meaningful.
- BUL: Conservative approach that records failures only when choices below are considered to be well-informed.
- Success is always recorded for both approaches.

Learning Considerations

1. Collecting training data for learning

- ACL: Aggressive approach that considers all failures as meaningful.
- BUL: Conservative approach that records failures only when choices below are considered to be well-informed.
- Success is always recorded for both approaches.

Learning Considerations

1. Collecting training data for learning

- ACL: Aggressive approach that considers all failures as meaningful.
- BUL: Conservative approach that records failures only when choices below are considered to be well-informed.
- Success is always recorded for both approaches.

Learning Considerations

1. Collecting training data for learning

- ACL: Aggressive approach that considers all failures as meaningful.
- BUL: Conservative approach that records failures only when choices below are considered to be well-informed.
- Success is always recorded for both approaches.

2. Using ongoing learning for plan selection

- Obtain a numeric measure of confidence in the ongoing learning output (i.e. a plan's likelihood of success in the situation).
- Use the confidence measure to adjust selection weights during probabilistic plan selection.

Learning Considerations

1. Collecting training data for learning

- ACL: Aggressive approach that considers all failures as meaningful.
- BUL: Conservative approach that records failures only when choices below are considered to be well-informed.
- Success is always recorded for both approaches.

2. Using ongoing learning for plan selection

- Obtain a numeric measure of confidence in the ongoing learning output (i.e. a plan's likelihood of success in the situation).
- Use the confidence measure to adjust selection weights during probabilistic plan selection.

Learning Considerations

1. Collecting training data for learning

- ACL: Aggressive approach that considers all failures as meaningful.
- BUL: Conservative approach that records failures only when choices below are considered to be well-informed.
- Success is always recorded for both approaches.

2. Using ongoing learning for plan selection

- Obtain a numeric measure of confidence in the ongoing learning output (i.e. a plan's likelihood of success in the situation).
- Use the confidence measure to adjust selection weights during probabilistic plan selection.

BDI Learning Framework

Previous work (Airiau et al. 2009)

- Augment static logical context conditions of plans with dynamic decision trees.
- Select plans probabilistically based on their ongoing expectation of success.
- Learn context conditions over time by training decision trees using success/failure outcomes under various situations.

BDI Learning Framework

Previous work (Airiau et al. 2009)

- Augment static logical context conditions of plans with dynamic decision trees.
- Select plans probabilistically based on their ongoing expectation of success.
- Learn context conditions over time by training decision trees using success/failure outcomes under various situations.

BDI Learning Framework

Previous work (Airiau et al. 2009)

- Augment static logical context conditions of plans with dynamic decision trees.
- Select plans probabilistically based on their ongoing expectation of success.
- Learn context conditions over time by training decision trees using success/failure outcomes under various situations.

BDI Learning Framework

Previous work (Airiau et al. 2009)

- Augment static logical context conditions of plans with dynamic decision trees.
- Select plans probabilistically based on their ongoing expectation of success.
- Learn context conditions over time by training decision trees using success/failure outcomes under various situations.

BDI Learning Framework

Previous work (Airiau et al. 2009)

- Augment static logical context conditions of plans with dynamic decision trees.
- Select plans probabilistically based on their ongoing expectation of success.
- Learn context conditions over time by training decision trees using success/failure outcomes under various situations.

Contributions of this paper

- A more principled analysis of the work in [Airiau et al. 2009].
- Learning with applicability filtering (using thresholds to filter plans that do not apply in a given situation).

BDI Learning Framework

Previous work (Airiau et al. 2009)

- Augment static logical context conditions of plans with dynamic decision trees.
- Select plans probabilistically based on their ongoing expectation of success.
- Learn context conditions over time by training decision trees using success/failure outcomes under various situations.

Contributions of this paper

- A more principled analysis of the work in [Airiau et al. 2009].
- Learning with applicability filtering (using thresholds to filter plans that do not apply in a given situation).

BDI Learning Framework

Previous work (Airiau et al. 2009)

- Augment static logical context conditions of plans with dynamic decision trees.
- Select plans probabilistically based on their ongoing expectation of success.
- Learn context conditions over time by training decision trees using success/failure outcomes under various situations.

Contributions of this paper

- A more principled analysis of the work in [Airiau et al. 2009].
- Learning with applicability filtering (using thresholds to filter plans that do not apply in a given situation).

Assumptions

Aim is to understand the nuances of learning under different goal-plan hierarchies using a simplified setting:

- Recursive/parameterised events or relational beliefsets not addressed.
- BDI failure recovery mechanism disabled during learning.
- Synthetic plan library with empty initial context conditions used.
- Simple account of non-determinism: successful actions have a 10% probability of failure.

Assumptions

Aim is to understand the nuances of learning under different goal-plan hierarchies using a simplified setting:

- Recursive/parameterised events or relational beliefsets not addressed.
- BDI failure recovery mechanism disabled during learning.
- Synthetic plan library with empty initial context conditions used.
- Simple account of non-determinism: successful actions have a 10% probability of failure.

Assumptions

Aim is to understand the nuances of learning under different goal-plan hierarchies using a simplified setting:

- Recursive/parameterised events or relational beliefsets not addressed.
- BDI failure recovery mechanism disabled during learning.
- Synthetic plan library with empty initial context conditions used.
- Simple account of non-determinism: successful actions have a 10% probability of failure.

Assumptions

Aim is to understand the nuances of learning under different goal-plan hierarchies using a simplified setting:

- Recursive/parameterised events or relational beliefsets not addressed.
- BDI failure recovery mechanism disabled during learning.
- Synthetic plan library with empty initial context conditions used.
- Simple account of non-determinism: successful actions have a 10% probability of failure.

Assumptions

Aim is to understand the nuances of learning under different goal-plan hierarchies using a simplified setting:

- Recursive/parameterised events or relational beliefsets not addressed.
- BDI failure recovery mechanism disabled during learning.
- Synthetic plan library with empty initial context conditions used.
- Simple account of non-determinism: successful actions have a 10% probability of failure.

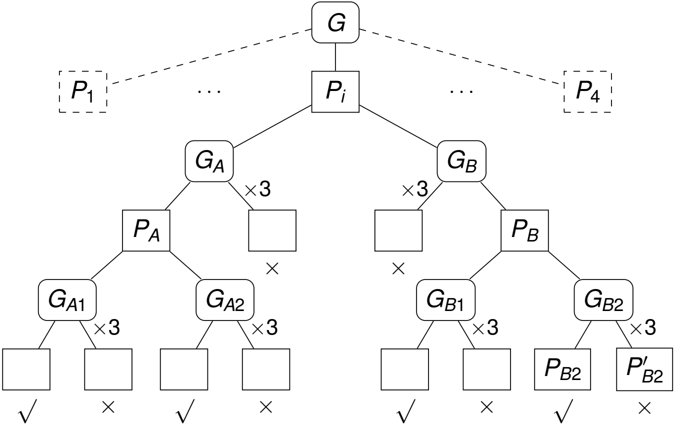
Assumptions

Aim is to understand the nuances of learning under different goal-plan hierarchies using a simplified setting:

- Recursive/parameterised events or relational beliefsets not addressed.
- BDI failure recovery mechanism disabled during learning.
- Synthetic plan library with empty initial context conditions used.
- Simple account of non-determinism: successful actions have a 10% probability of failure.

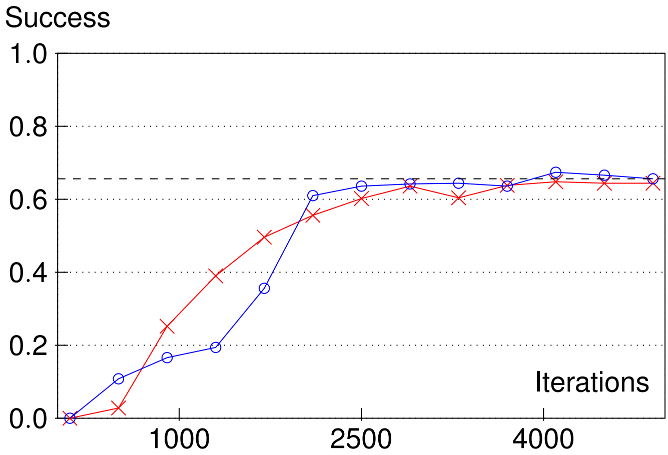
Ongoing work aims to relax these constraints towards a more practical system.

Results: Does Selective Recording Matter?



Structure where both schemes show comparable performance.

Results: Does Selective Recording Matter? (cont.)



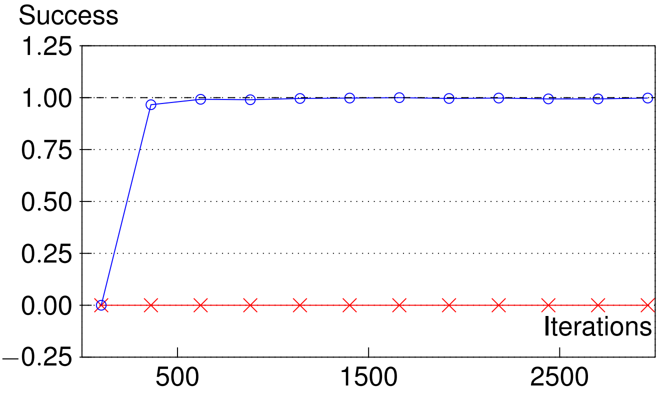
Performance of ACL (crosses) vs. BUL (circles).
Dashed line shows optimal performance.

Results: Learning with Applicability Filtering

Plan execution is generally not cost-free, so agent may fail a goal without even trying if it is unlikely to succeed.

Results: Learning with Applicability Filtering

Plan execution is generally not cost-free, so agent may fail a goal without even trying if it is unlikely to succeed.

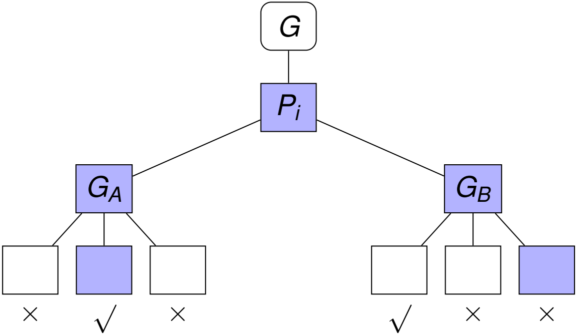


Performance of ACL (crosses) vs. BUL (circles).
Dashed line shows optimal performance.

Improving Plan Selection

Coverage-based confidence measure

Idea is that confidence in a plan's decision tree increases as more choices below the plan are covered.



Highlighted path shows $1/9$ possible choices under P_i.

Improving Plan Selection (cont.)

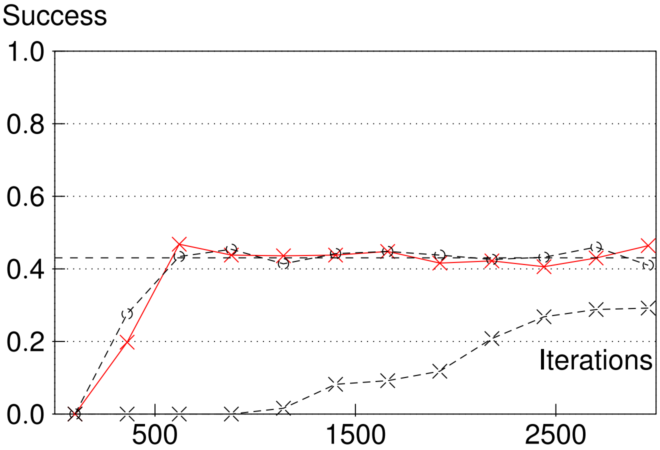
How confidence influences plan selection

- When the plan has not been tried before (zero coverage) we bias towards the default weight of 0.5.
- As more options are tried (approaching full coverage), we progressively bias towards the decision tree probability $p_T(w)$.

Plan selection weight calculation

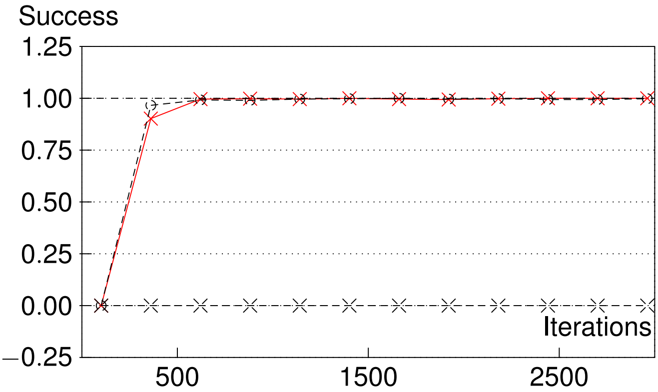
$$\Omega'_T(w) = 0.5 + [c_T(w) * (p_T(w) - 0.5)] \,.$$

Results: Goal-Plan Hierarchy B



Performance of ACL+Ω'_T (red crosses) vs. previous results in structure that suits the conservative BUL approach. Dashed line shows optimal performance.

Results: Learning with Applicability Filtering



Performance of ACL+Ω'_T (red crosses) vs. previous results

Learning Context Conditions for BDI Plan Selection

- Learning BDI plan selection is desirable since designing exact context conditions for practical systems is non-trivial.
- Our approach uses decision trees to learn the context condition of plans.
- We suggest that an aggressive sampling scheme combined with a coverage-based confidence measure is a good candidate approach for the general hierarchical setting.

Learning Context Conditions for BDI Plan Selection

- Learning BDI plan selection is desirable since designing exact context conditions for practical systems is non-trivial.
- Our approach uses decision trees to learn the context condition of plans.
- We suggest that an aggressive sampling scheme combined with a coverage-based confidence measure is a good candidate approach for the general hierarchical setting.

Learning Context Conditions for BDI Plan Selection

- Learning BDI plan selection is desirable since designing exact context conditions for practical systems is non-trivial.
- Our approach uses decision trees to learn the context condition of plans.
- We suggest that an aggressive sampling scheme combined with a coverage-based confidence measure is a good candidate approach for the general hierarchical setting.

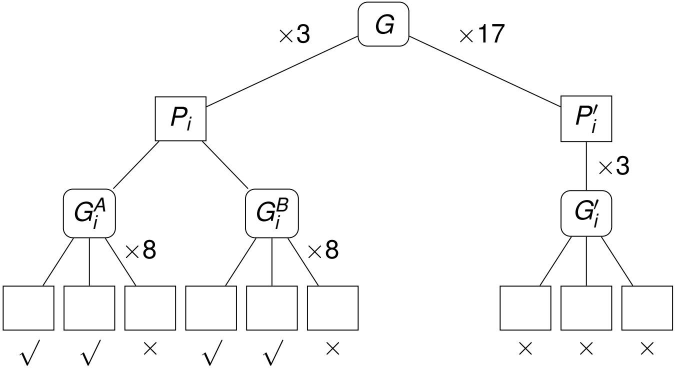
References

M. Bratman, D. Israel, and M. Pollack.
Plans and resource-bounded practical reasoning.
Computational Intelligence, 4(4):349–355, 1988.

A.S. Rao
AgentSpeak (L): BDI agents speak out in a logical computable language.
Lecture Notes in Computer Science, 1038:42–55, 1996.

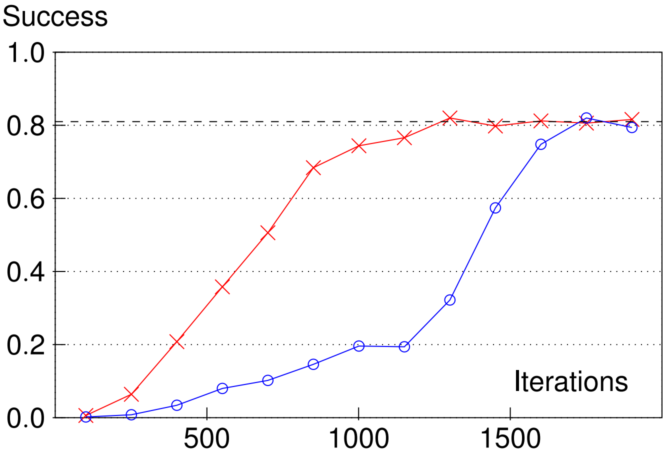
S. Airiau, L. Padgham, S. Sardina, and S. Sen.
Enhancing Adaptation in BDI Agents Using Learning Techniques.
International Journal of Agent Technologies and Systems, 2009.

Goal-Plan Structure T1



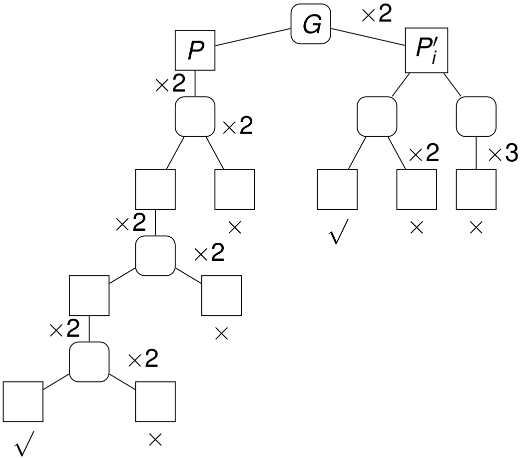
Structure where one of many complex options has a solution. This configuration suits the aggressive ACL approach.

Results: Goal-Plan Structure T1



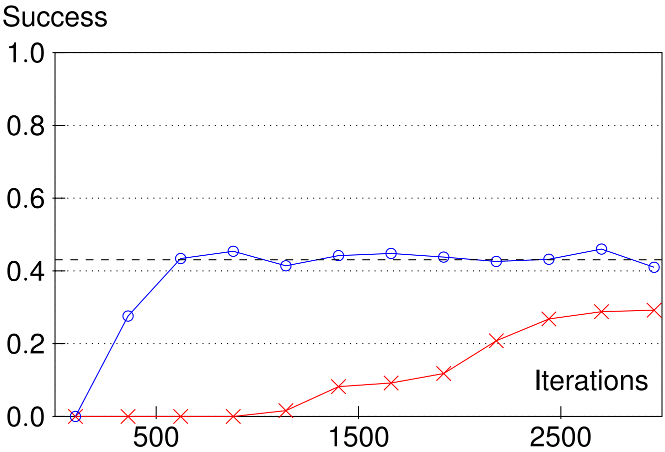
Performance of ACL (crosses) vs. BUL (circles).
Dashed line shows optimal performance.

Goal-Plan Structure T2



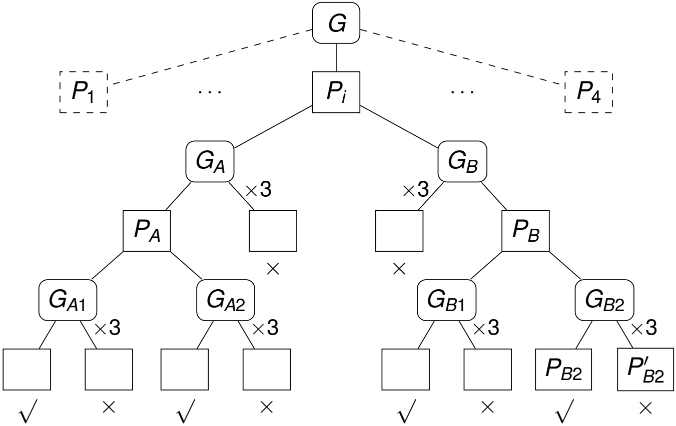
Structure has solution in one complex option.
This configuration suits the conservative BUL approach.

Results: Goal-Plan Structure T2



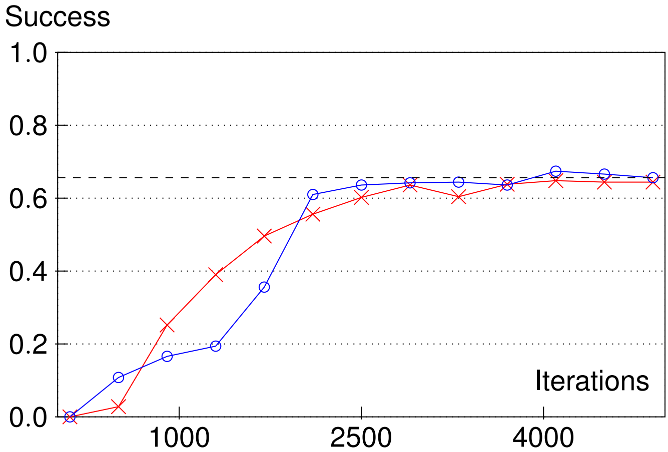
Performance of ACL (crosses) vs. BUL (circles).
Dashed line shows optimal performance.

Goal-Plan Structure T3



Structure where both schemes show comparable performance.

Results: Goal-Plan Structure T3



Performance of ACL (crosses) vs. BUL (circles).
Dashed line shows optimal performance.