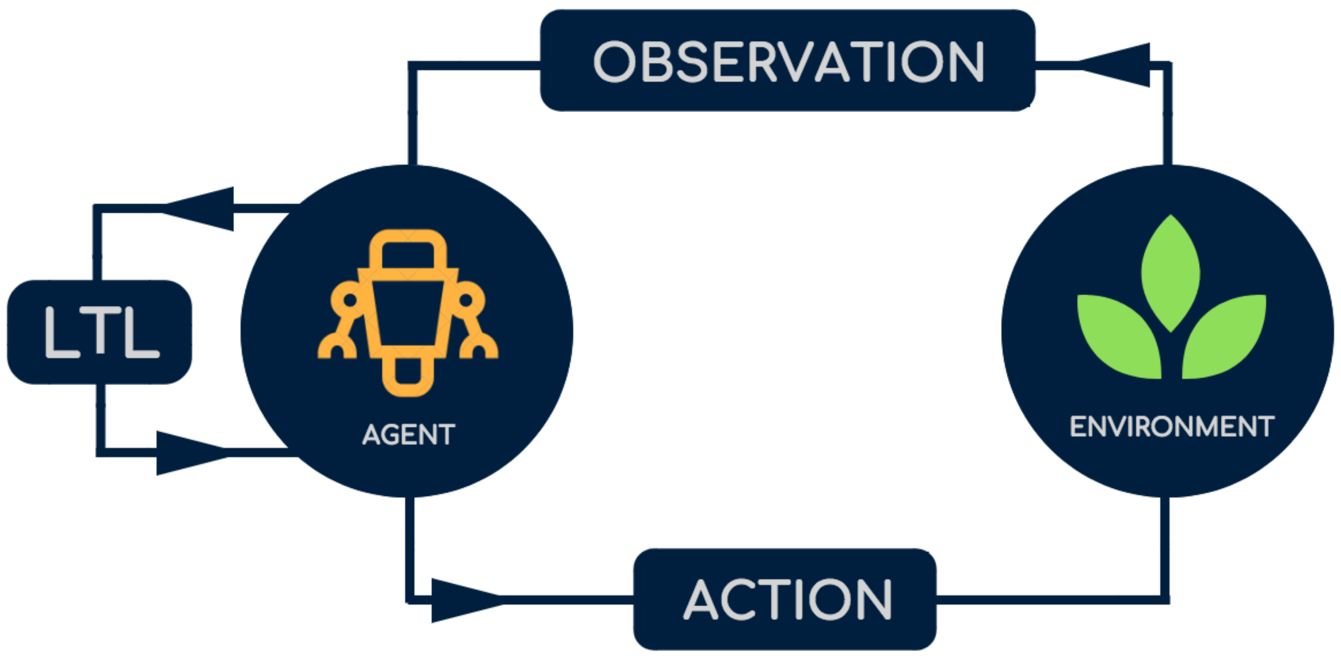
Logically-Constrained Reinforcement Learning

Mohammadhosein Hasanbeig, Alessandro Abate, Daniel Kroening

hosein.hasanbeig@cs.ox.ac.uk, alessandro.abate@cs.ox.ac.uk, kroening@cs.ox.ac.uk

Introduction

We propose **Logically-Constrained Reinforcement Learning (LCRL)** algorithm to synthesize policies for Markov Decision Processes **(MDPs)**, such that a linear time property is satisfied. Additionally, we show that LCRL sets up an online Asynchronous Value Iteration **(AVI)** method to calculate the maximum probability of satisfying the given property, at any given state of the MDP – a convergence proof for the procedure is provided.


OBSERVATION

LTL

AGENT

ENVIRONMENT

ACTION

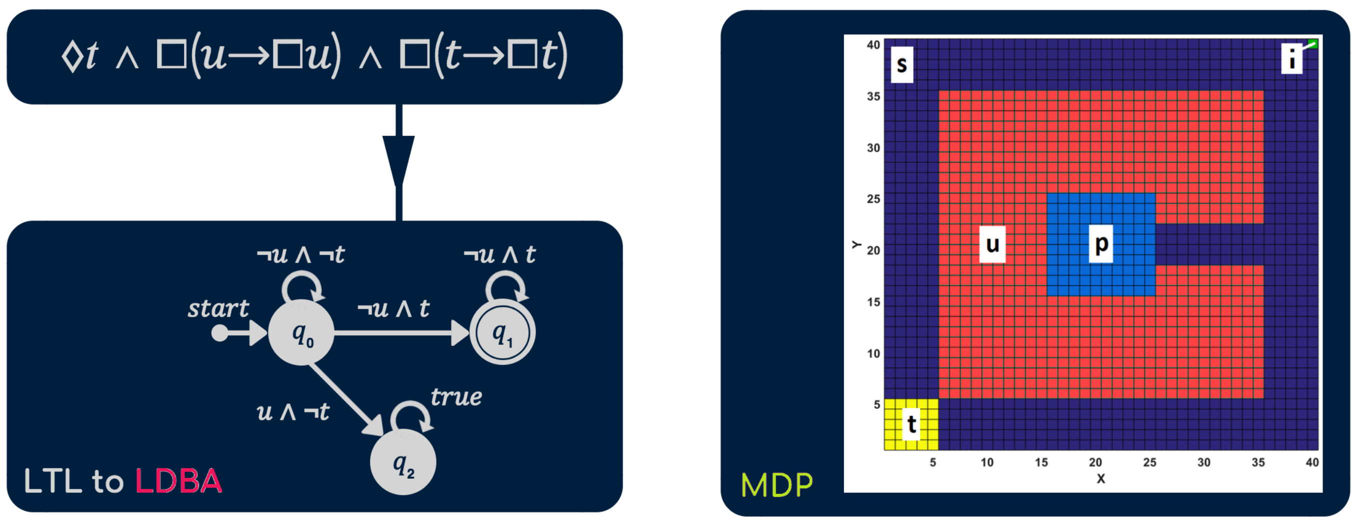
Reinforcement Learning with Logical Constraints

LCRL aims to

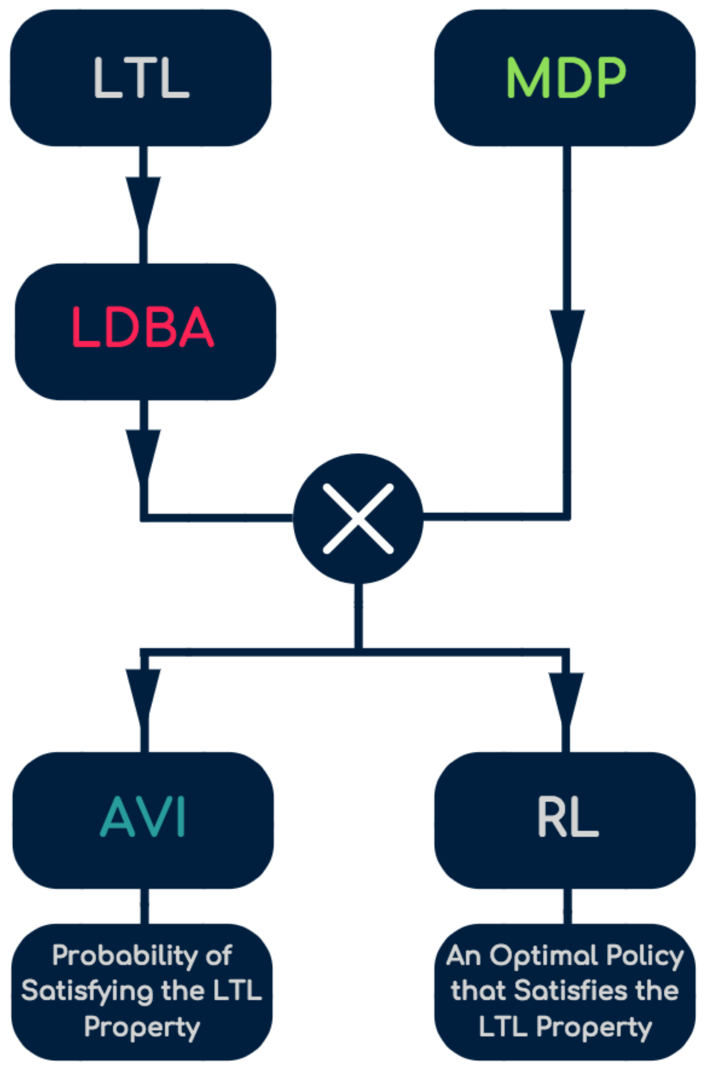
- synthesize a control policy for a stochastic model such that the resulting traces satisfy a given temporal logic property
- calculate the maximum probability of satisfying the property
- increase the scalability of conventional model checkers
- leverage machine learning techniques in formal methods

Algorithm Flow

- Limit Deterministic Büchi Automaton **(LDBA)** [1]
- Synchronizing LDBA with MDP, i.e. product MDP


$\Diamond t \wedge \Box(u \rightarrow \Box u) \wedge \Box(t \rightarrow \Box t)$

Left: LTL formula to LDBA conversion – Right: MDP


LTL

MDP

LDBA

AVI

RL

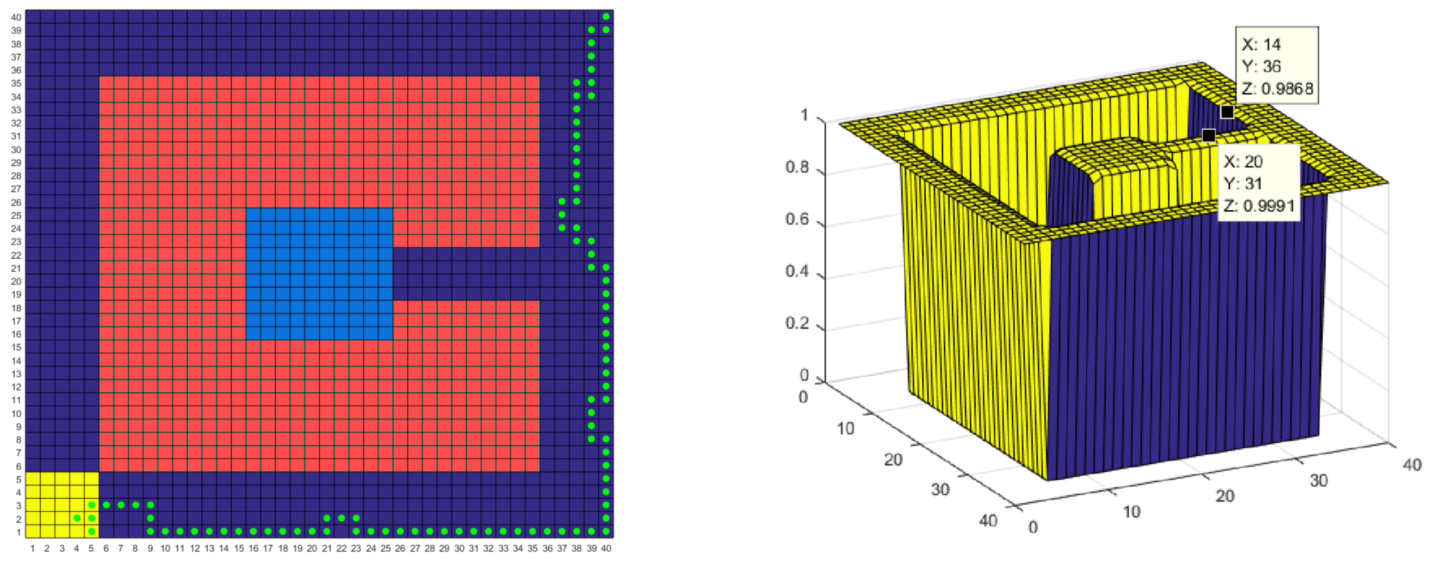
Probability of Satisfying the LTL Property

An Optimal Policy that Satisfies the LTL Property

Algorithm flow

Results

- Policy generates a trace that satisfies the LTL property
- Probabilities are accurately calculated comparing to conventional DP-based methods



Left: Satisfying policy – Right: Calculated probabilities [2]

Future Work

- Infinite-state space MDPs
- Multi-agent systems
- Empirical experiments

References

[1] S. Sickert, J. Esparza, S. Jaax, and J. Křetínskỳ, "Limit-deterministic Büchi Automata for Linear Temporal Logic," in *International Conference on Computer Aided Verification*, pp. 312–332, Springer, 2016.

[2] M. Hasanbeig, A. Abate, and D. Kroening, "Logically-Constrained Reinforcement Learning," *arXiv preprint arXiv:1801.08099*, 2018.