SENIOR THESIS 2021-22

Prospectus

Samar Rahmouni
srahmoun@andrew.cmu.edu

Advisor: Prof. Giselle Reis
giselle@cmu.edu

1 Description of the Problem

Take from proposal, adapt.

2 Significance

Take from proposal, adapt.

3 Background and Related Work

Take from proposal, adapt.

- Hybrid architecture becoming more popular [DIFFERENTIATION OF BLACKBOX COMBINATORIAL SOLVERS]
- Explains how most approaches so far only consider Neural networks [A Review of Formal Methods applied to Machine Learning] [why is it always the french in formal methods]
- Kurd and Kelly verifications goals G4 and G5, robustness to disturbance to inputs and ensuring that outputs are not hazardous. Both are ensured in RL, a change to the reward scheme can change all conclusions.
- SMT-based formal methods [CSP problems]
- MILP-based [mixed integer linear problem]
- Abstract Interpretation based.
- Interesting note: very little work in formal methods for data preparation. One cool work that deals with this uses static analysis to infer assumptions on the input data.
- Talk about developing approaches to determine constraints on the training process that enforce a certain behavior.

Artificial intelligence has moved away in the past years from symbolic to data-driven. Data-driven AI have proven both efficient and their algorithms easy to reuse. However, we are facing the limitations of such AI today whether in its incapability of developing abstract relations between components, or in the consistent problem of meta-learning in extrapolating learned knowledge in one environment to another, or finally, the safety properties that are harder to reason about and furthermore ensure. In the light of such limitations, hybrid architectures are becoming increasingly popular. [Add some of the examples].

Most approaches so far only considered neural networks. In [Add ref], they present multiple formal methods that have been investigate and applied to Machine Learning, these include SMT-based and MILP-based formal methods which use constraint satisfaction and mixed-integer linear problems to solve a subproblem of ML. Add further how they use them. Our approach is novel as it (1) focuses on tabular reinforcement learning rather than neural networks and (2) investigates a translation from symbolic compositional state representation in RL to logic-based inference.

[Add what verification goals are from Kurd and Kelly and why RL does not uphold them, and how our proposed approach can help.]

4 Research Contribution

- Provides a framework to incorporate a safety controller into a RL architecture.
- Compositional modeling of the states of a RL controller.
- Incorporate rule knowledge into decision making of a tabular RL in a deterministic environment.
- Investigate balance between optimization and safety in (1) Basic Reinforcement Learning, (2) Reinforcement Learning with a safe controller and (3) Logic-based inference Reinforcement Learning.

5 Evaluation Plan Timeline

The project will proceed in two phases. The first focuses on the implementation of the reinforcement learning controller along with connecting the safe controller. Precisely, we proceed with the following.

1. Formalizing the car platooning scenario as an optimization problem for the RL agent. [7th October]
2. Concretizing the architecture and describe the relations between components. [8th October]
3. Implementation of the RL controller independently of the safe controller. [20th October]
4. Modeling the states as a propositional formula in the lines of classical planning in order to match the safe controller specifications. [26th October]
5. Implementation of the framework to map from Maude to Python, after feeding the state representation into the safe controller. [10th November]
6. Experimenting and testing the optimization results of the RL controller with and without the safety guarantees of the safe controller. [20th November]

The second phase focuses on the possible inferences of safety properties in a deterministic environment given a compositional propositional formulas state representation.

1. Investigate learning inference rules [state given action] in a deterministic environment.
2. Incorporate rule knowledge into decision making of a tabular RL precisely by looking into a hybrid architecture that makes use of symbolic AI to adapt the Q-values table and/or the epsilon-greedy algorithm.
3. Develop a mapping from compositional state representation to reward schemes given safe controller outputs. To work out.
4. Experimenting and testing the optimization and safety results of the RL controller with modified rewards.

Given these two implementations, we want to compare both approaches into balancing between safety guarantees and optimization. Both phases will use the car platooning as a main scenario for implementation and testing.