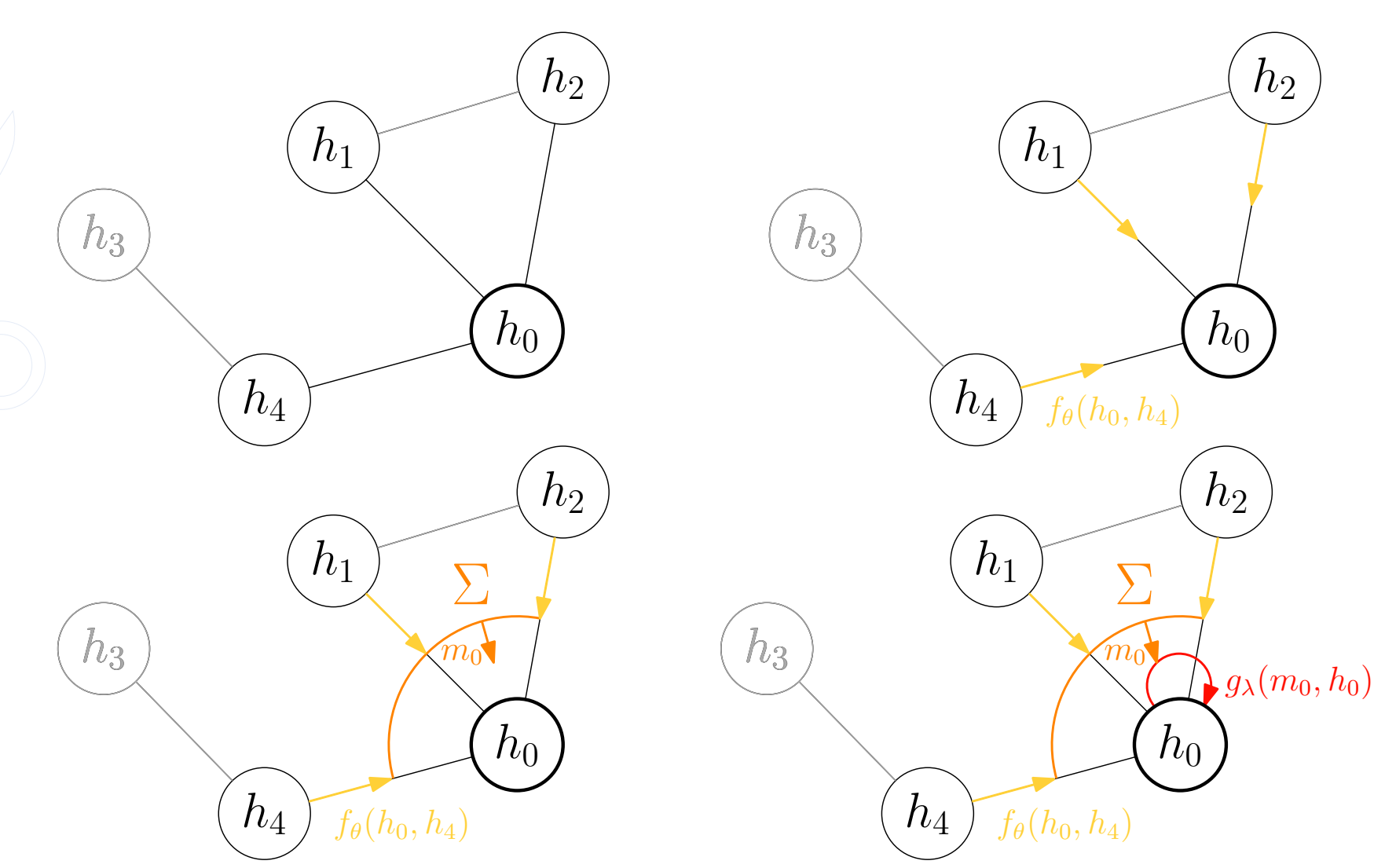


Learning to select cuts brings new challenges to deep reinforcement learning.



Node update in a graph neural network.

Learning to cut

Antoine Prouvost, Aleksandr M. Kazachkov, Andrea Lodi
antoine.prouvost@polymtl.ca

BACKGROUND

- Cutting planes are highly effective to solve Mixed Integer Linear Programs (MILPs);
- In large quantities, they burden the Linear Programming (LP) relaxation;
- Cuts interfere with each other and with branch-and-bound;
- There still lacks satisfying heuristics to evaluate cuts contributions.

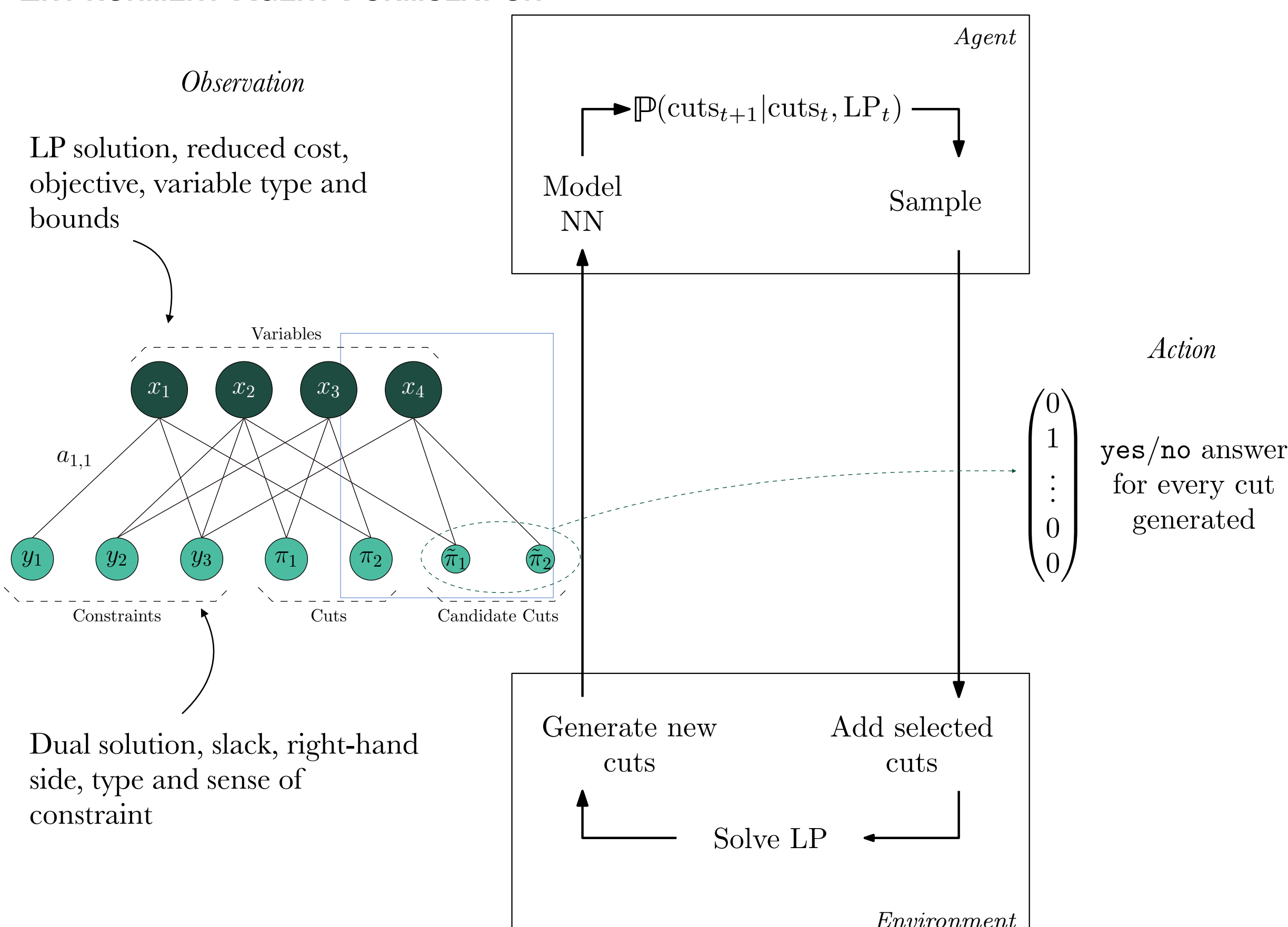
OBJECTIVE

We aim to learn a cut selection policy that **generalizes** to unseen instances.

REINFORCEMENT LEARNING (RL) METHODS

- Policy gradients actor critic (proximal policy optimization);
- Combine the loss with supervised signal: predicting the LP improvement for every cut generated;
- Graph neural network to parametrize the policy.

ENVIRONMENT-AGENT FORMULATION



EXPERIMENTAL SETUP

- Context: case study on one 2-matching instance (TSP lib gr48);
- Goal: find optimal integer solution with fewest number of cuts;
- Zero-half and Gomory mixed integer cuts;
- Agent select one cut per round;
- Reward of -1 per cut selected and terminate when integrality is reached.

EARLY RESULTS

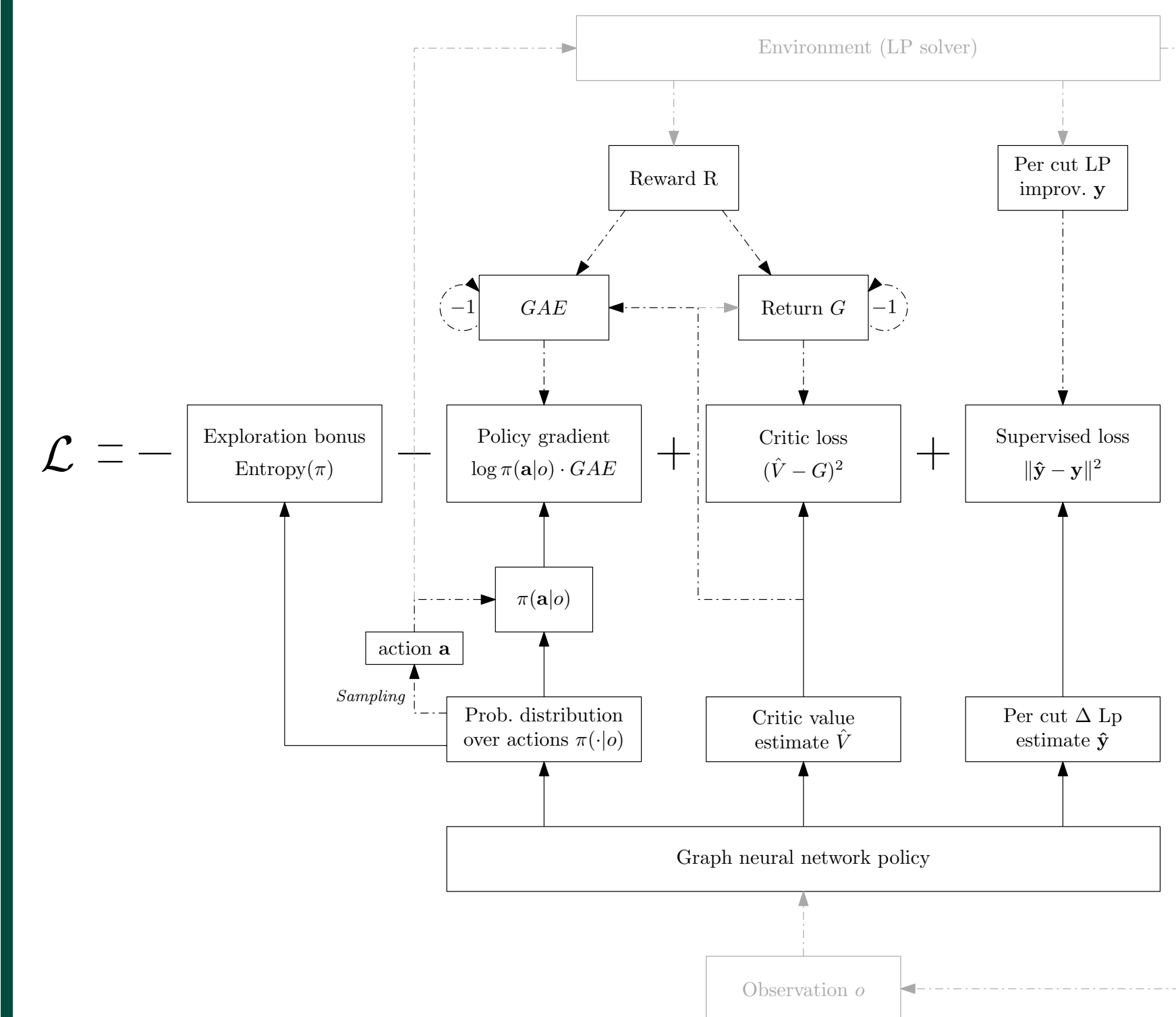


Two training runs over gr48, one cuts at the time.

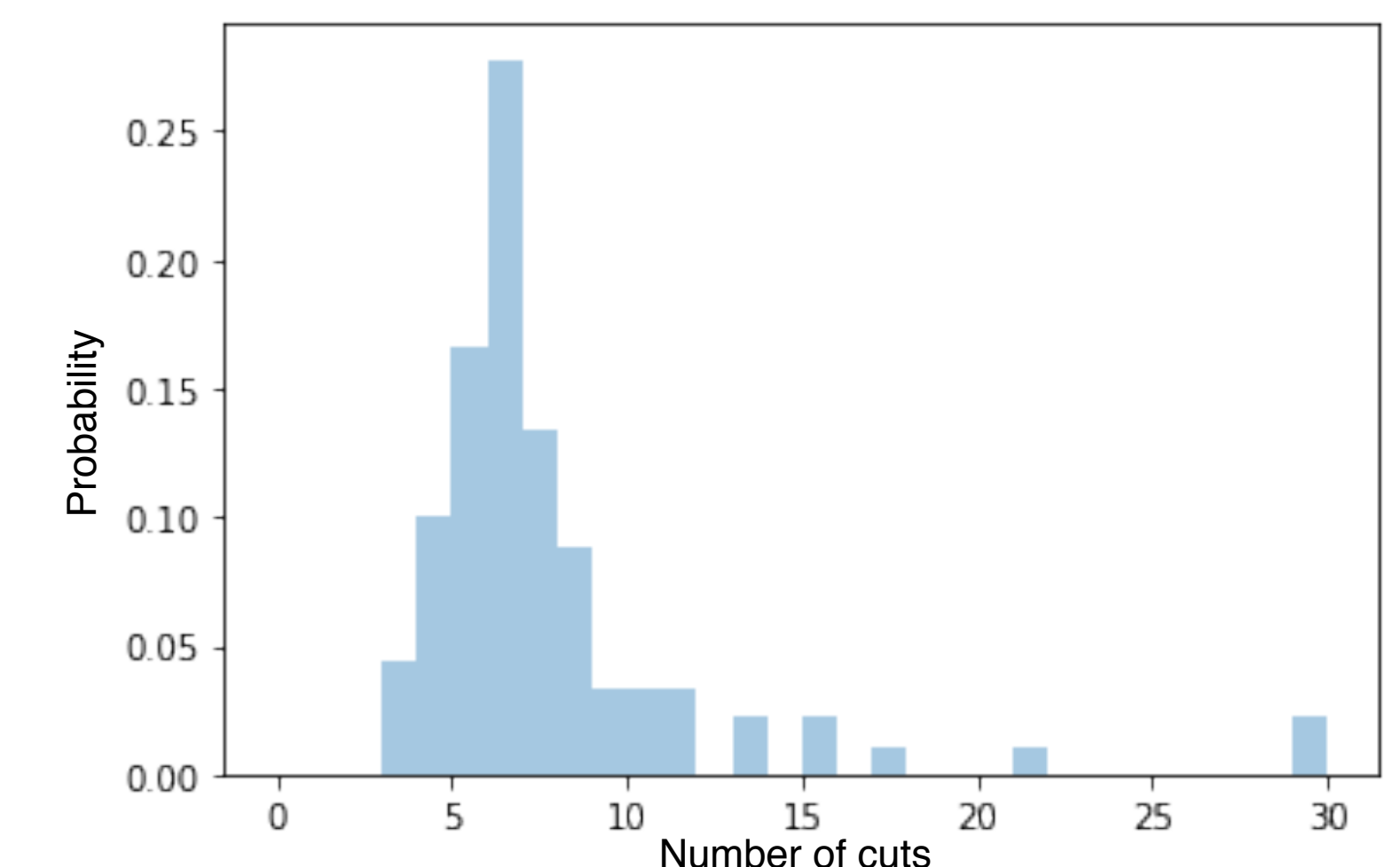
- Learning curves are unstable across different hyper-parameters;
- Agent is able to improve over random: “successful” solves the instance in **5 cuts** when random does in 8.8 ± 11.4 on average;
- This still doesn't account for any generalization.

DISCUSSION

- Cut selection is a novel, complex, RL problem and presents large structured state and action spaces;
- The multiple cuts setting is highly more challenging;
- We are moving toward more advanced RL algorithms to solve the problem.



Details of the supervised actor critic loss.



Distribution of the number of cuts needed to solve gr48 using random selection.