PAC Statistical Model Checking for Markov Decision Processes and Stochastic Games[1]

Pranav Ashok, Jan Křetínský, Maximilian Weininger

Technical University of Munich

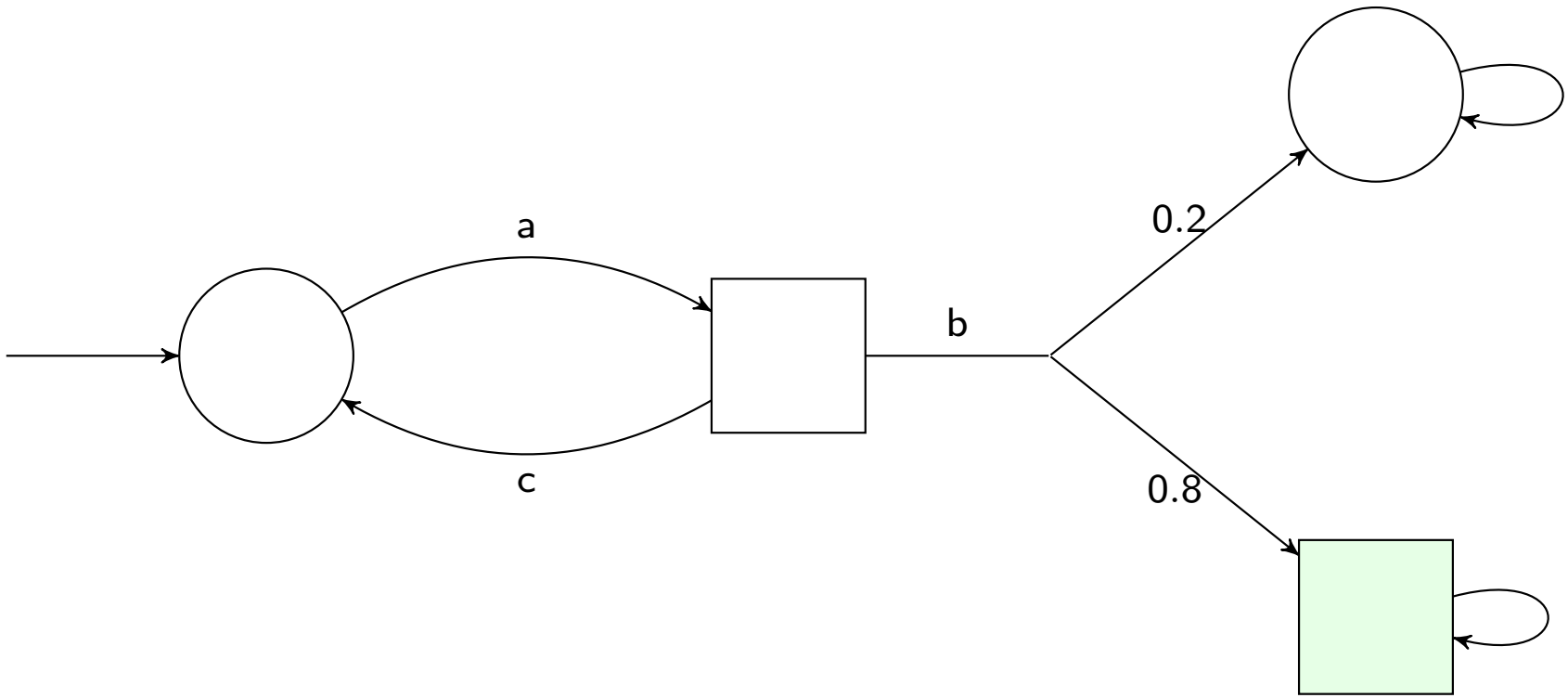
Highlights of Logic, Automata and Games
Warsaw, Poland

September 19, 2019

[1] based on paper presented at CAV 2019

Stochastic Game

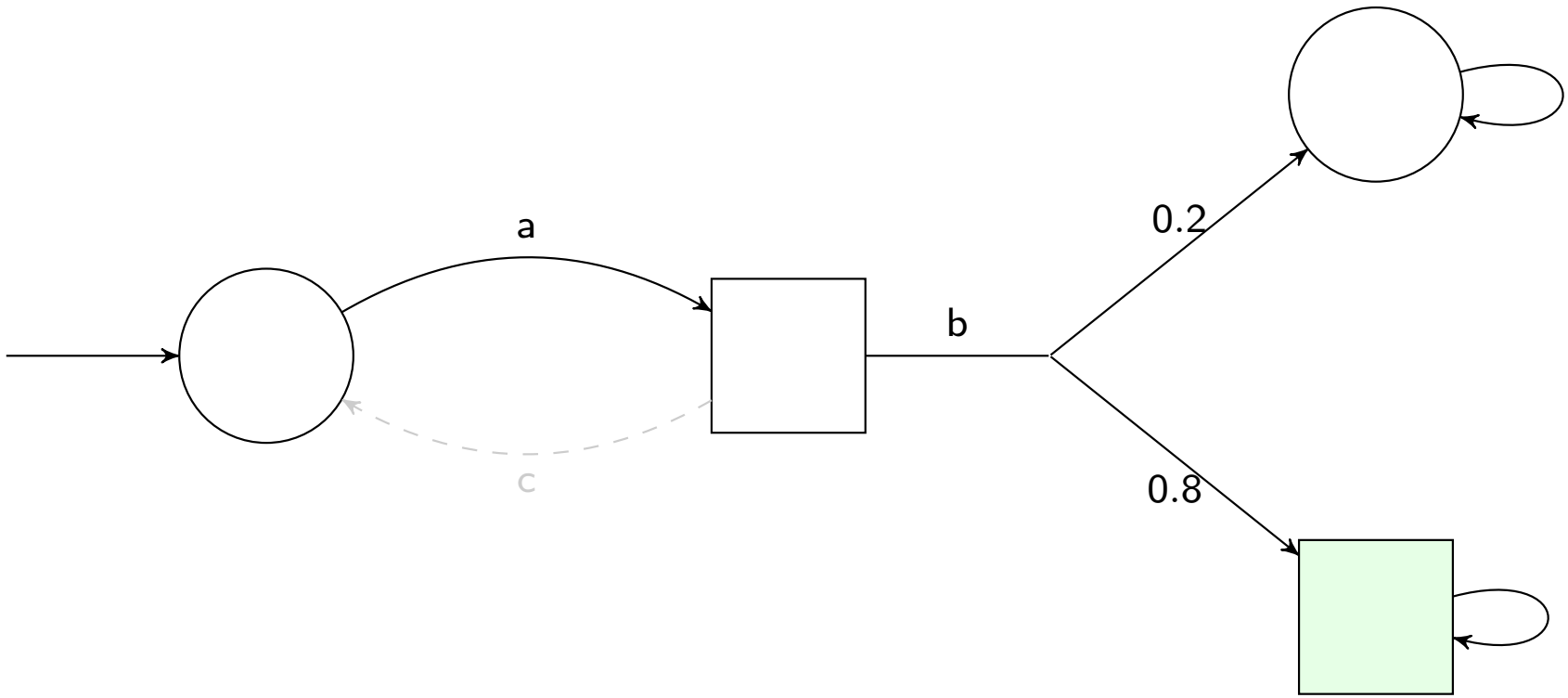
Reachability



Objective
□ player: maximize P(F □)
○ player: minimize P(F □)

Stochastic Game

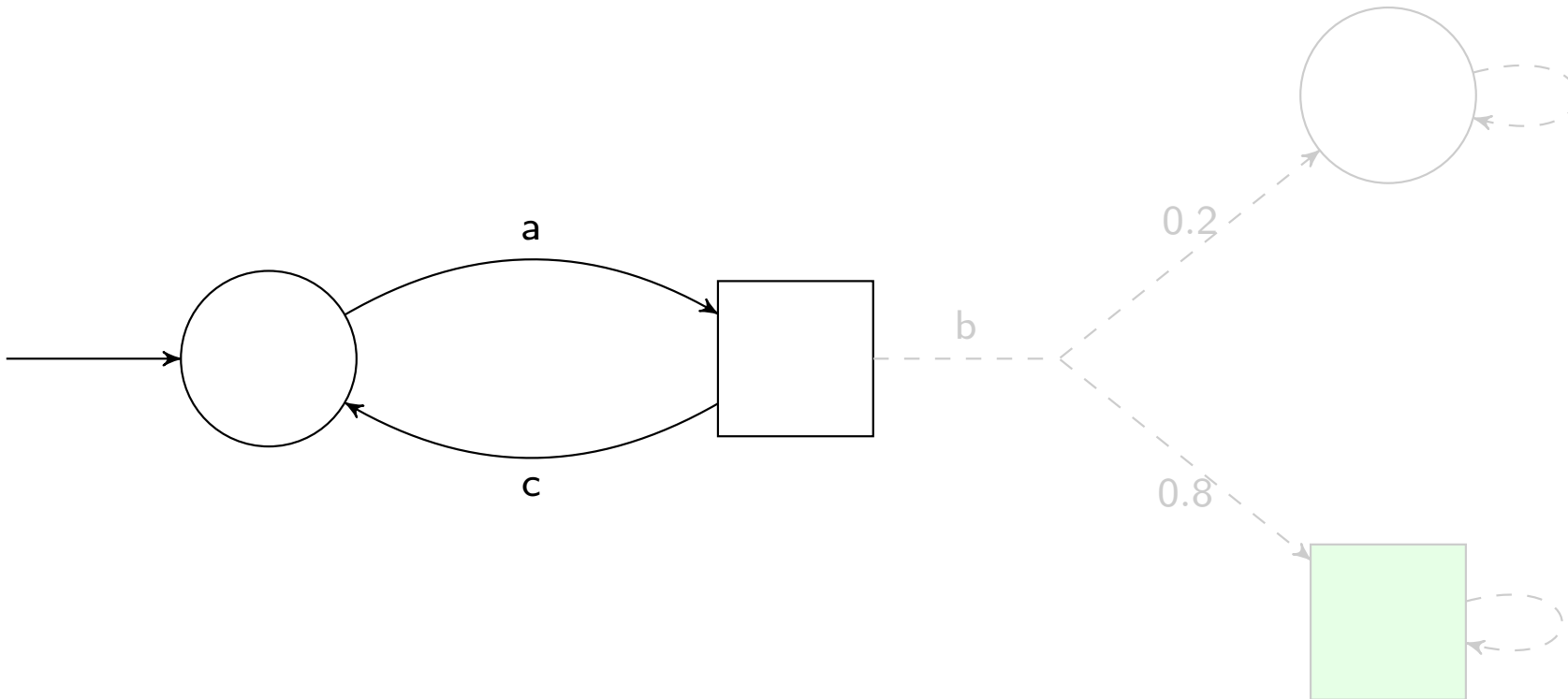
Reachability



Objective
□ player: maximize P(F □)
○ player: minimize P(F □)

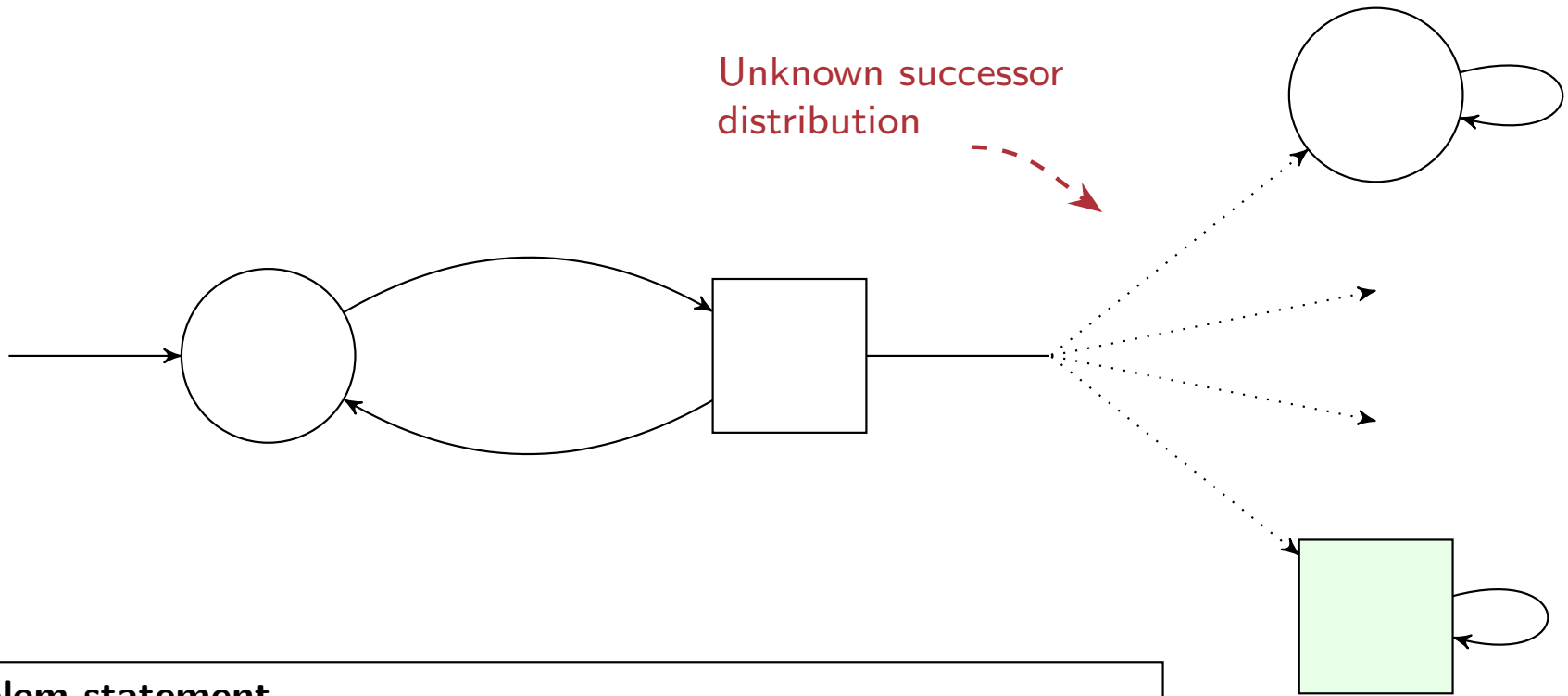
Stochastic Game

Reachability



Objective
□ player: maximize P(F □)
○ player: minimize P(F □)

This work: Black-box (limited information setting)



Problem statement
Compute $V(s) = \max_\sigma \min_\tau P_s^{\sigma,\tau}(F\ \square) = \min_\tau \max_\sigma P_s^{\sigma,\tau}(F\ \square)$
with guarantees

Background

- Seminal paper on Stochastic Games [*Condon 90*]
  quadratic programming, strategy iteration, value iteration

Background

- Seminal paper on Stochastic Games [*Condon 90*]
  quadratic programming, strategy iteration, value iteration
- Algos not directly applicable on general SG
- First practical algorithm for general SG giving guarantees
  [*Kelmendi et. al. 2018*]

Background

- Seminal paper on Stochastic Games [*Condon 90*]
  quadratic programming, strategy iteration, value iteration
- Algos not directly applicable on general SG
- First practical algorithm for general SG giving guarantees [*Kelmendi et. al. 2018*]
- This work: first algorithm for limited information SG

The Algorithm

Similar to Kelmendi et. al. 2018

while $U - L$ is large

1. Simulate and estimate
2. Back-propagate

The Algorithm

Similar to Kelmendi et. al. 2018

while $U - L$ is large

1. Simulate and estimate
2. Back-propagate

The how

- Simulation finds important parts of state space

The Algorithm

Similar to Kelmendi et. al. 2018

while $U - L$ is large

1. Simulate and estimate
2. Back-propagate

The how

- Simulation finds important parts of state space
- Simulation computes Hoeffding confidence intervals

  ball around estimate such that real prob. falls in the ball with high confidence

The Algorithm

Similar to Kelmendi et. al. 2018

while $U - L$ is large

1. Simulate and estimate
2. Back-propagate

The how

- Simulation finds important parts of state space
- Simulation computes Hoeffding confidence intervals
  ball around estimate such that real prob. falls in the ball with high confidence
- Information conservatively back-propagated

The Algorithm

Similar to Kelmendi et. al. 2018

while $U - L$ is large

1. Simulate and estimate
2. Back-propagate

The how

- Simulation finds important parts of state space
- Simulation computes Hoeffding confidence intervals
  ball around estimate such that real prob. falls in the ball with high confidence
- Information conservatively back-propagated
- Other tricks to ensure fixpoint convergence

Conclusion

- Algorithm for reachability in limited information MDP/SG result $\in [0.6 - \epsilon, 0.6 + \epsilon]$ with prob of going wrong 10^{-8}
- Implemented and benchmarked in PRISM Model Checker
- First algorithm to do so for SG
- First practical algorithm for MDPs