Policy Gradient Methods in the Evolutionary Pricing Game

Sahar Jahani, Galit Ashkenazi-Golan,

Edward Plumb, Katerina Papadaki, Bernhard von Stengel

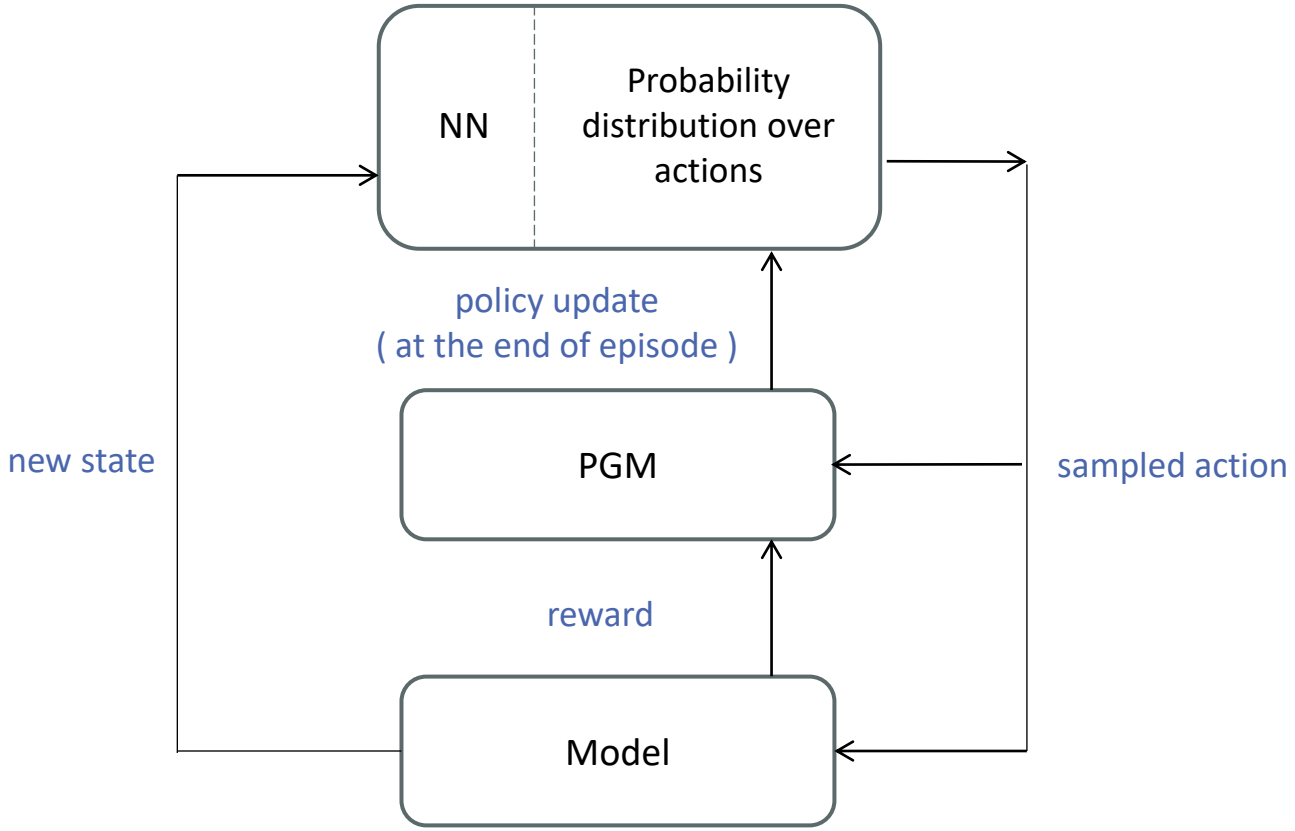
ALGA 2023

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

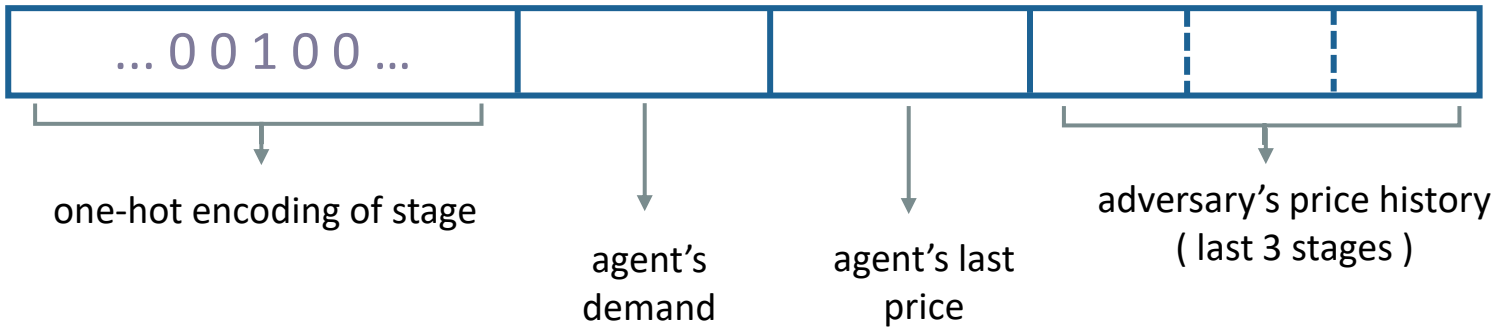
RL trained agents become players in the game

- Model: a classic duopoly pricing game by Selten 1965
- Agents set prices over 25 stages
- Firm with higher price loses costumers to the other firm at next stage
- Trade-off between short term vs long term payoff

Learning using Policy Gradient methods



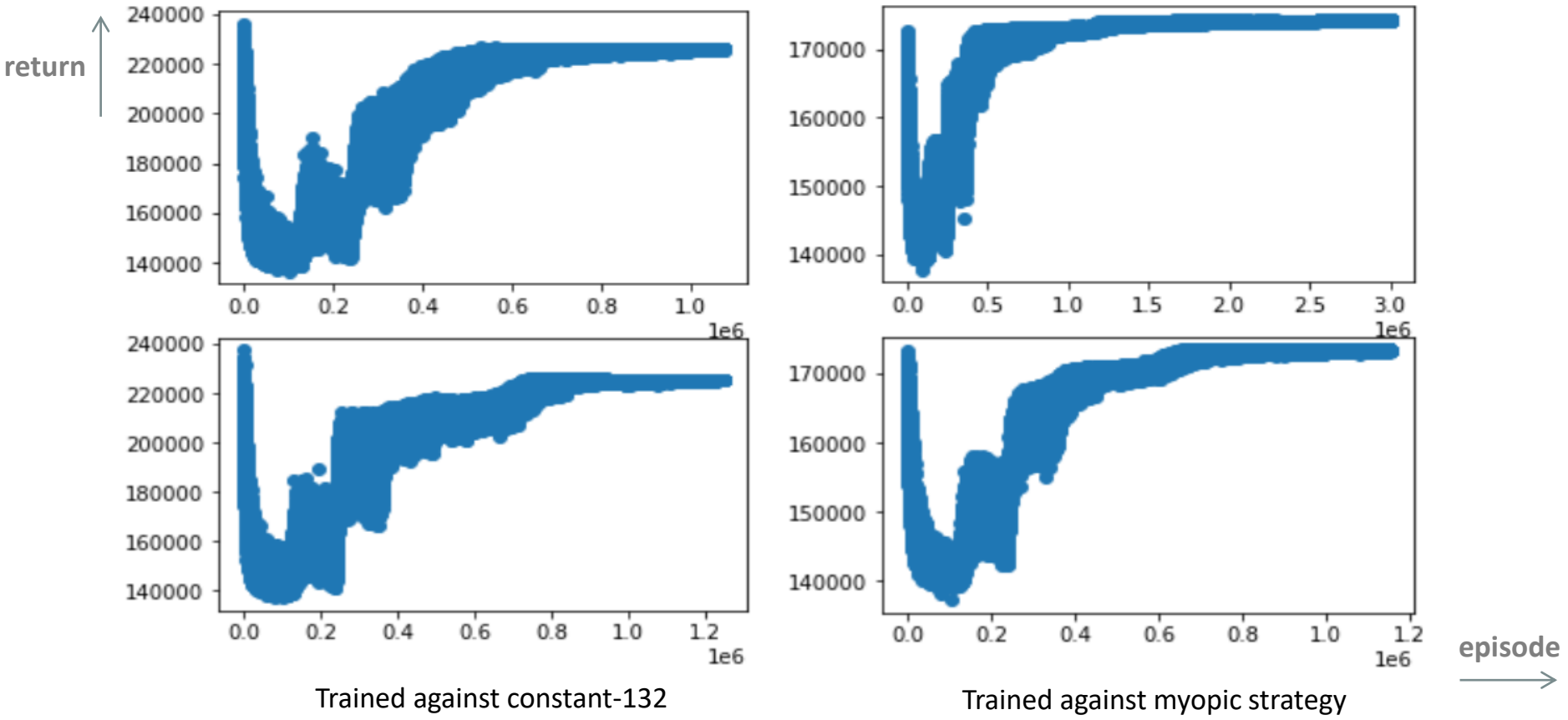
- State representation:



- $price = monopoly\ price - action, \quad action \in \{0, 3, 6, ..., 60\}$
- Reward = (game's payoff – BASE)
- BASE = game's payoff if agent would play monopoly price at every stage

Not easily trained!

- Learning backward from the last stage



Trained against constant-132

Trained against myopic strategy

Agent's return will be the player's payoff in the later phase.

Mixed equilibrium as learning framework

- new low-cost and high-cost agents are trained against equilibrium strategies