Discrete Payments Optimization Using Reinforcement Learning

Pablo Samuel Castro[1], Ajit Desai[2], Han Du[2], Rodney Garratt[3], Francisco Rivadeneyra[2]

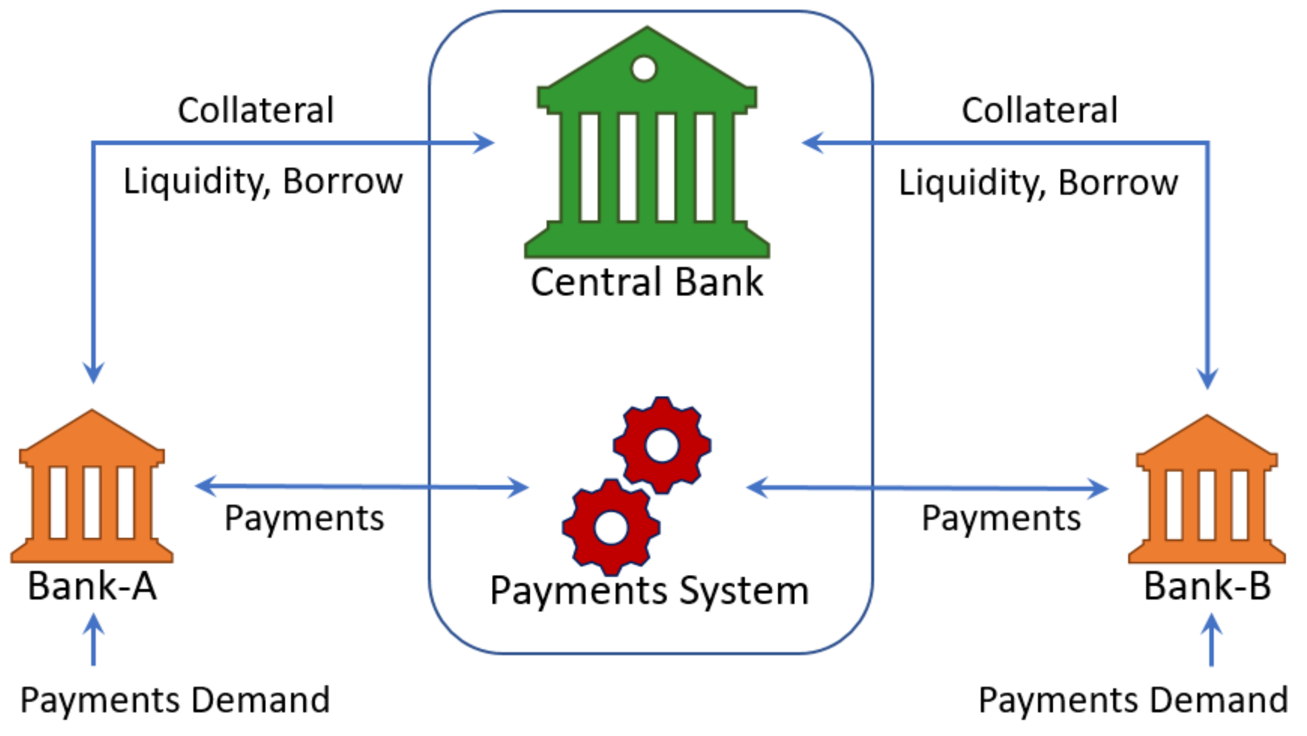
[1]Google Research, Brain Team, [2]Bank of Canada, [3]University of California Santa Barbara

Objective

Introduction: we use reinforcement learning (RL) to estimate the policy function of a bank participating in a wholesale payments system.

Bank's objective: Learn the policy function to *jointly* choose the amount of initial liquidity and the rate at which to pay intraday payments to minimize total cost.

Payment system environment: a simple real-time gross settlement system with two banks (one non-strategic). A portion of payments are indivisible leading to an integer optimization problem.



Model and Methodology

Timing:

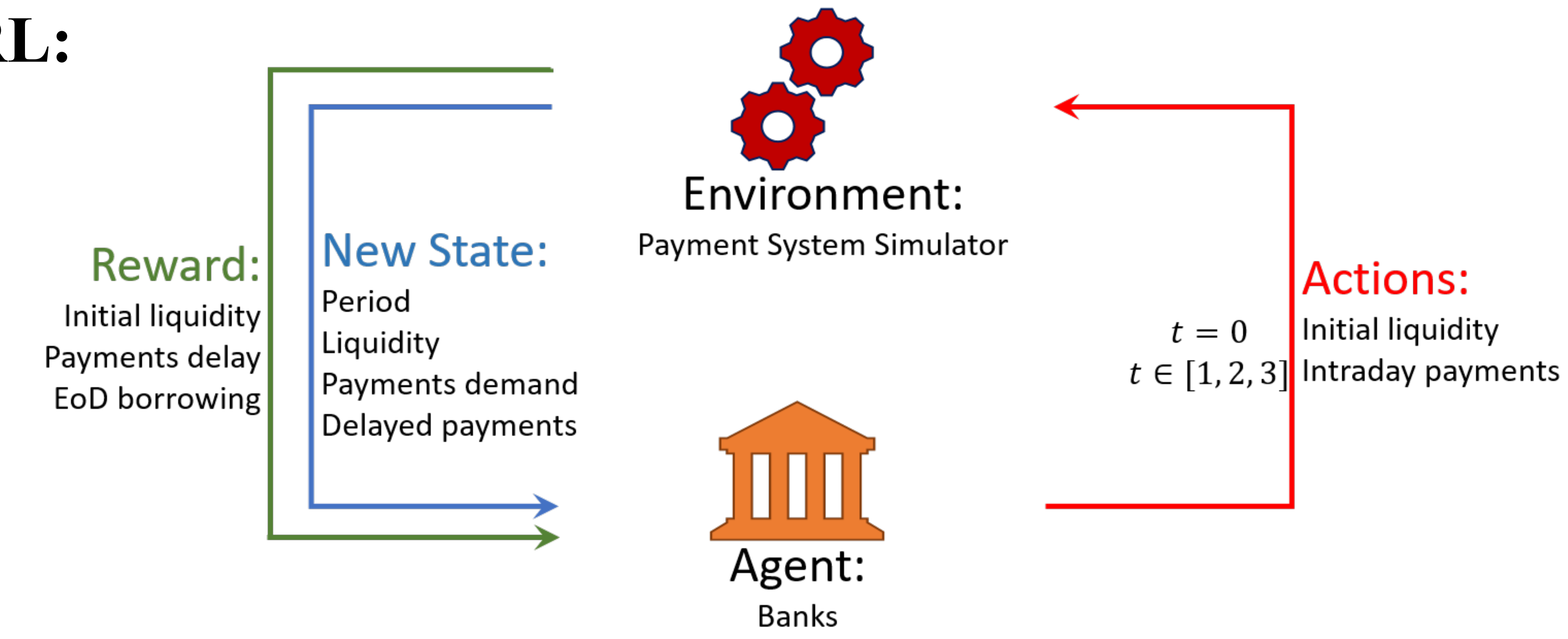
- $t = 0$: choice of costly initial liquidity;
- $t \in [1,2,3]$: payment requests arrival; delaying them is costly. A portion of request in $t = 2$ is indivisible;
- Payments not processed incur a high borrowing cost.

Total cost: $\mathscr{R}$ = **initial liquidity + payment delay + borrowing**

Key trade-off:

- Optimal policy should balance the cost of initial liquidity with the potential delay and borrowing costs.
- Indivisibility of a portion of the $t = 2$ payment creates a wedge in the delay costs in $t = 1$ and 2.

RL:

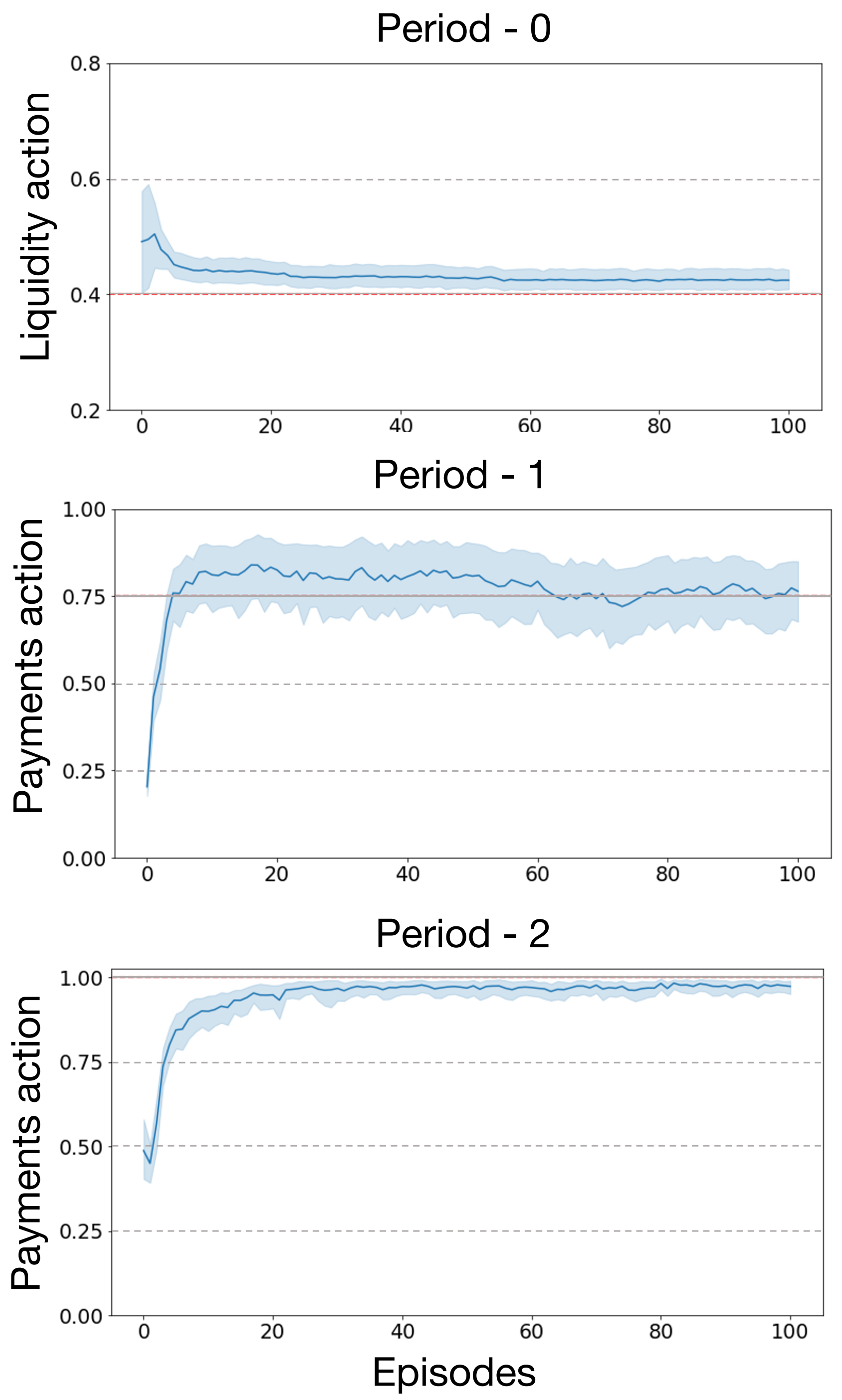


Deep Q-learning: Reinforcement learning algorithm to learn the value of an action (a) in a given state (s).

$$Q(s,a) = r(s,a) + \gamma \max_a Q(s',a)$$

Results

Intuition of the precautionary motive:
By sending less than the entirety of the payment in $t = 1$, the bank conserves the appropriate amount of liquidity to be able to send the indivisible payment in $t = 2$.



In $t = 0$: agent learns to allocate enough liquidity

In *t=1*: agent learns the precautionary policy to conserve liquidity

In $t = 2$: agent is able and choses to send all payments (divisible and non-divisible)

Results show agents trained with RL can solve complex integer optimization problems in a real-world environment.