

Towards Blackwell Optimality: Bellman Optimality is All You Can Get

Victor Boone, Adrienne Tuynman — AISTATS 2026

May 3, 2026

ANITI, IRIT, CNRS, LAAS, Université de Toulouse, Université de Lille, Inria

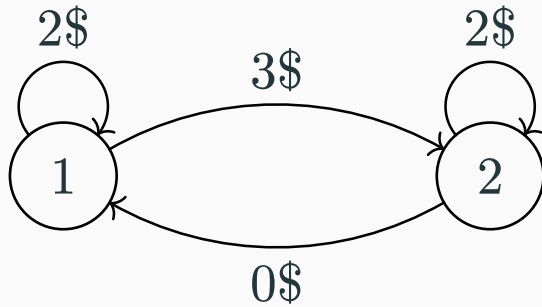
Learning optimal policies in Markov decision processes

- There is a Markov decision process M — *but it's hidden to you.*
- You can **query** state-action pairs **to learn** the MDP:

At time t , query $(S_t, A_t) \rightsquigarrow$ observe a transition to S_{t+1} and a reward R_t .

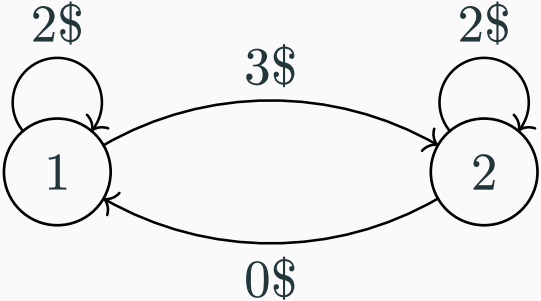
⇒ How quickly can you learn an **optimal policy**?

Higher order optimalities: gain optimality, bias optimality and beyond

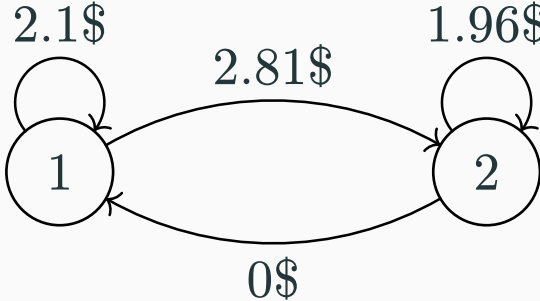


True environment

Higher order optimalities: gain optimality, bias optimality and beyond

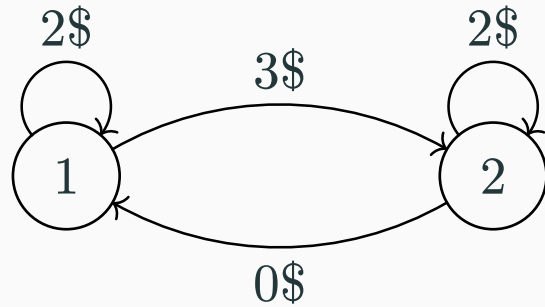


True environment

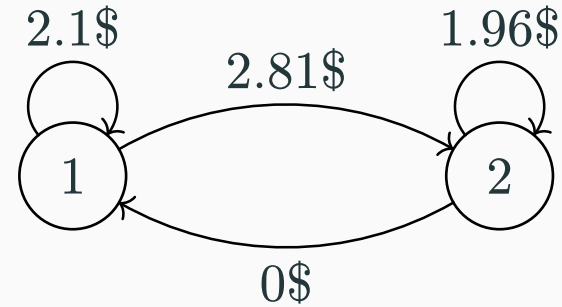


What you see while learning

Higher order optimalities: gain optimality, bias optimality and beyond



True environment



What you see while learning

A whole **hierarchy** of optimalities:

$$\Pi_{-1}^* \supseteq \Pi_0^* \supseteq \Pi_1^* \supseteq \dots \supseteq \Pi_\infty^*$$

gain optimality bias optimality higher bias optimality Blackwell optimality

Learning higher order optimality in MDPs

Theorem (Boone, Tuynman, 2025). Learning stops at **Bellman optimality**

$$\begin{array}{ccccc} \Pi_{-1}^* & \supseteq & \Pi_B^* & \supseteq & \Pi_0^* \\ \text{gain optimality} & & \text{Bellman optimality} & & \text{bias optimality} \end{array}$$

and further refined optimality cannot be learned.¹

¹Check our paper for more details!

Victor Boone & Adrienne Tuynman, (2025). Towards Blackwell Optimality: Bellman Optimality Is All You Can Get. *arXiv preprint arXiv:2510.13476*.