Sound Algorithms in Imperfect Information Games

Michal Šustr, Martin Schmid, Matej Moravčík, Neil Burch, Marc Lanctot, Michael Bowling
michal.sustr@aic.fel.cvut.cz, {mschmid, moravcik, burchn, lanctot, bowlingm} @google.com

CTU CZECH TECHNICAL UNIVERSITY IN PRAGUE — DeepMind

Exploitability and ε-Nash equilibria are ill-suited to measure an online algorithm's worst-case performance. We examine how to do so properly, using the setting of repeated sequential games. https://arxiv.org/abs/2006.08740

Abstract

Search has played a fundamental role in computer game research since the very beginning. And while **online search** has been commonly used in perfect information games such as Chess and Go, online search methods for imperfect information games have only been introduced relatively recently. This paper addresses the question of *what is* a *sound online algorithm* in an *imperfect information setting* of *two-player zero-sum games.*
We argue that the fixed-strategy definitions of **exploitability** and **ε-Nash** equilibria are **ill-suited** to measure an online algorithm's **worst-case performance**.
We thus formalize **ε-soundness**, a concept that **connects** the worst-case performance of an **online algorithm** to the performance of an **ε-Nash** equilibrium. As ε-soundness can be difficult to compute in general, we introduce a **consistency framework** -- a **hierarchy** that connects an online algorithm's behavior to a Nash equilibrium. These multiple levels of consistency describe in what sense an online algorithm plays "just like a fixed Nash equilibrium". These notions further illustrate the **difference** between perfect and imperfect information settings, as the same consistency guarantees have different worst-case online performance in perfect and imperfect information games.
The definitions of soundness and the consistency hierarchy finally provide **appropriate tools** to analyze online algorithms in repeated imperfect information games.
We thus inspect some of the previous online algorithms in a new light, bringing new insights into their worst-case performance guarantees.

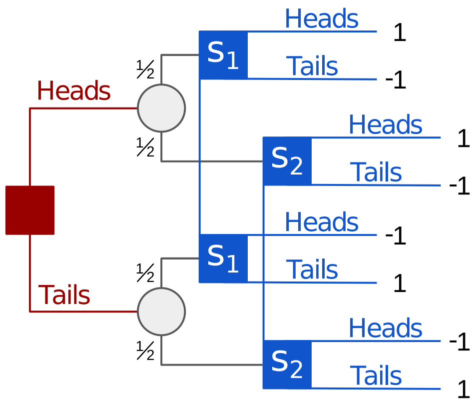
"Black-box" online algorithm

Soundness

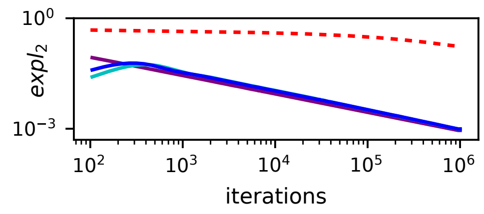
For an (k,ε)-sound online algorithm Ω, the expected average reward against any opponent is at least as good as if it followed an ε-Nash equilibrium fixed strategy σ for any number of matches k':

$$\forall k' \geq k \;\; \forall \Omega_2 : \mathbb{E}_{m \sim P^{k'}_{\Omega,\Omega_2}} [\mathcal{R}(m)] \geq \mathbb{E}_{m \sim P^{k'}_{\sigma,\Omega_2}} [\mathcal{R}(m)]$$

Coordinated Matching Pennies



A locally consistent algorithm

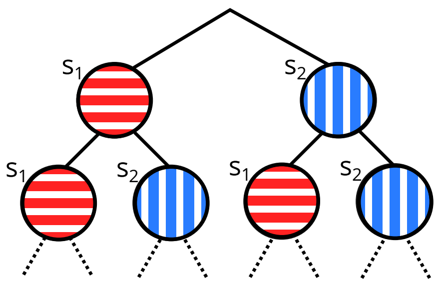


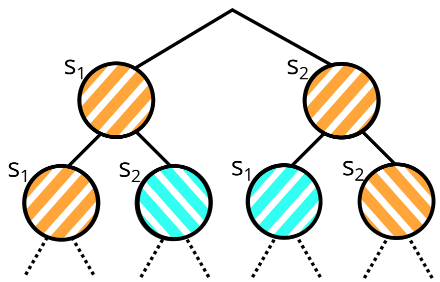
Consistency framework

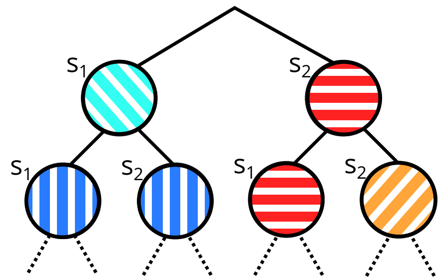
Infostate strategies

- $p_{Heads} = 1$, $p_{Tails} = 0$
- $p_{Heads} = 0$, $p_{Tails} = 1$
- $p_{Heads} = 0.6$, $p_{Tails} = 0.4$
- $p_{Heads} = 0.4$, $p_{Tails} = 0.6$

Strongly globally consistent ⊆ Globally consistent ⊆ Locally consistent







Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum provided under the program "Projects of Large Research, Development, and Innovations Infrastructures" (CESNET LM2015042), is greatly appreciated. This work was supported by Czech science foundation grant no. 18-27483Y.