Pure Exploration with Infinite Answers

Riccardo Poiani, Martino Bernasconi, Andrea Celli

Università Bocconi
DEPARTMENT OF COMPUTING SCIENCES

AISTATS 2026

Supported by Capiplo CRYPTONOMEX and ERC grant 101165466–PLA-STEER

Pure Exploration Problems

Setting

- $K \in \mathbb{N}$ probability distributions $\{\nu_k\}_{k=1}^K$ with means $\boldsymbol{\mu} = \{\mu_k\}_{k=1}^K$
- An answer space $\mathcal{X}$ and correct answer set-valued function $\mathcal{X}^\star(\boldsymbol{\mu}) \subseteq \mathcal{X}$
- At each round $t \in \mathbb{N}$, we select $A_t \in [K]$ and observe $R_t \sim \nu_{A_t}$
- Collect data and stop at some step τ to return an answer $\hat{x}_\tau \in \mathcal{X}$

Pure Exploration Problems

Setting

- $K \in \mathbb{N}$ probability distributions $\{\nu_k\}_{k=1}^K$ with means $\boldsymbol{\mu} = \{\mu_k\}_{k=1}^K$
- An answer space $\mathcal{X}$ and correct answer set-valued function $\mathcal{X}^\star(\boldsymbol{\mu}) \subseteq \mathcal{X}$
- At each round $t \in \mathbb{N}$, we select $A_t \in [K]$ and observe $R_t \sim \nu_{A_t}$
- Collect data and stop at some step τ to return an answer $\hat{x}_\tau \in \mathcal{X}$

Goal: given $\delta \in (0, 1)$, identify a correct answer in high probability

$$\mathbb{P}\left(\hat{x}_\tau \notin \mathcal{X}^\star(\boldsymbol{\mu})\right) \leq \delta$$

...while minimizing the expected stopping time $\mathbb{E}[\tau]$

Pure Exploration Problems

Setting

- $K \in \mathbb{N}$ probability distributions $\{\nu_k\}_{k=1}^K$ with means $\boldsymbol{\mu} = \{\mu_k\}_{k=1}^K$
- An answer space $\mathcal{X}$ and correct answer set-valued function $\mathcal{X}^\star(\boldsymbol{\mu}) \subseteq \mathcal{X}$
- At each round $t \in \mathbb{N}$, we select $A_t \in [K]$ and observe $R_t \sim \nu_{A_t}$
- Collect data and stop at some step τ to return an answer $\hat{x}_\tau \in \mathcal{X}$

Goal: given $\delta \in (0, 1)$, identify a correct answer in high probability

$$\mathbb{P}\left(\hat{x}_\tau \notin \mathcal{X}^\star(\boldsymbol{\mu})\right) \leq \delta$$

...while minimizing the expected stopping time $\mathbb{E}[\tau]$

Examples

- Best arm identification: $\mathcal{X} = [K]$ and $\mathcal{X}^\star(\boldsymbol{\mu}) = \arg\max_k \mu_k$
- ϵ-best arm identification: $\mathcal{X} = [K]$ and $\mathcal{X}^\star(\boldsymbol{\mu}) = \{k : \mu_k \geq \max_i \mu_i\}$
- Threshold problems: $\mathcal{X} = \{\text{yes}, \text{no}\}$; Decide if $\max_k \mu_k$ is above a threshold γ

Focus of this work

For problems with a single correct answer and $|\mathcal{X}| < +\infty$, there exists several algorithms which are asymptotically optimal for $\delta \to 0$ [e.g., Garivier and Kaufmann, 2016, Degenne et al., 2019, Wang et al., 2021]

Focus of this work

For problems with a single correct answer and $|\mathcal{X}| < +\infty$, there exists several algorithms which are asymptotically optimal for $\delta \to 0$ [e.g., Garivier and Kaufmann, 2016, Degenne et al., 2019, Wang et al., 2021]

Degenne and Koolen [2019] extend these result to problems with multiple correct answers and $|\mathcal{X}| < +\infty$

Focus of this work

For problems with a single correct answer and $|\mathcal{X}| < +\infty$, there exists several algorithms which are asymptotically optimal for $\delta \to 0$ [e.g., Garivier and Kaufmann, 2016, Degenne et al., 2019, Wang et al., 2021]

Degenne and Koolen [2019] extend these result to problems with multiple correct answers and $|\mathcal{X}| < +\infty$

What if $\mathcal{X}$ and/or $\mathcal{X}^\star(\boldsymbol{\mu})$ are **infinite**?

Examples

- ϵ-regression of some (possibly multi-dimensional) function $f(\boldsymbol{\mu})$ (e.g., the value of an optimal arm)
- learning ϵ-Nash equilibria by querying noisy values of a payoff matrix in a bimatrix game

Results

This work

Tight instance-dependent asymptotic-optimality theory for infinite answer problems

Results

This work

Tight instance-dependent asymptotic-optimality theory for infinite answer problems

In more detail:

- Tight instance-dependent lower bounds for this general setting

Results

This work

Tight instance-dependent asymptotic-optimality theory for infinite answer problems

In more detail:

- Tight instance-dependent lower bounds for this general setting
- Explanation of why existing methods for multiple but finite correct answers cannot be applied as-is

Results

This work

Tight instance-dependent asymptotic-optimality theory for infinite answer problems

In more detail:

- Tight instance-dependent lower bounds for this general setting
- Explanation of why existing methods for multiple but finite correct answers cannot be applied as-is
- Present a framework that enjoys asymptotic optimality for $\delta \to 0$

Results

This work

Tight instance-dependent asymptotic-optimality theory for infinite answer problems

In more detail:

- Tight instance-dependent lower bounds for this general setting
- Explanation of why existing methods for multiple but finite correct answers cannot be applied as-is
- Present a framework that enjoys asymptotic optimality for $\delta \to 0$
- The framework generalizes existing works on finite answer problems [Garivier and Kaufmann, 2016, Degenne and Koolen, 2019] and it allows to understand in which problems with infinite answers the existing methods remain optimal

Thank you!

Thank you for your attention

References

Aurélien Garivier and Emilie Kaufmann. Optimal best arm identification with fixed confidence. In *Conference on Learning Theory*, pages 998–1027. PMLR, 2016.

Rémy Degenne, Wouter M Koolen, and Pierre Ménard. Non-asymptotic pure exploration by solving games. *Advances in Neural Information Processing Systems*, 32, 2019.

Po-An Wang, Ruo-Chun Tzeng, and Alexandre Proutiere. Fast pure exploration via frank-wolfe. *Advances in Neural Information Processing Systems*, 34:5810–5821, 2021.

Rémy Degenne and Wouter M Koolen. Pure exploration with multiple correct answers. *Advances in Neural Information Processing Systems*, 32, 2019.