Sequential Test for the Lowest Mean

From Thompson to Murphy Sampling

Emilie Kaufmann

Wouter M. Koolen

Aurélien Garivier

CWI/INRIA workshop

Paris, Wednesday 26th September, 2018

Why are we here?

Team
[6]PAC

Guedj

Kaufmann

Grünwald

Koolen

[6]Making **Probably Approximately Correct** Learning

Safe, Active, Sequential, Structure-aware, Ideal, Efficient

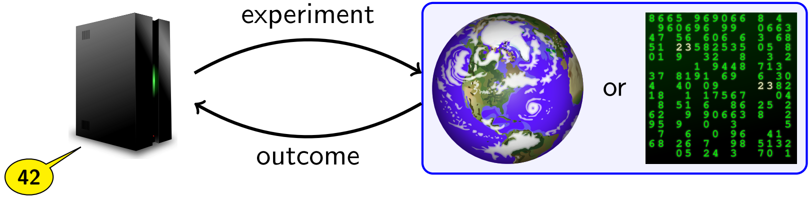
Grand Goal: Interactive Machine Learning

Query:

most effective drug dose?

most appealing website layout?

safest next robot action?



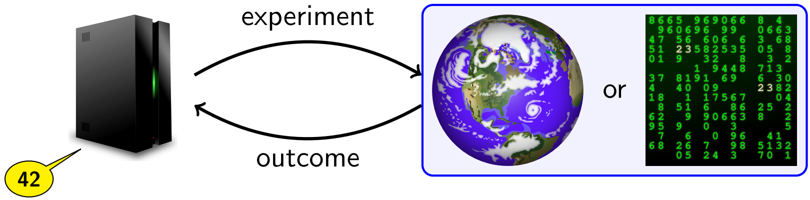
Grand Goal: Interactive Machine Learning

Query:

most effective drug dose?

most appealing website layout?

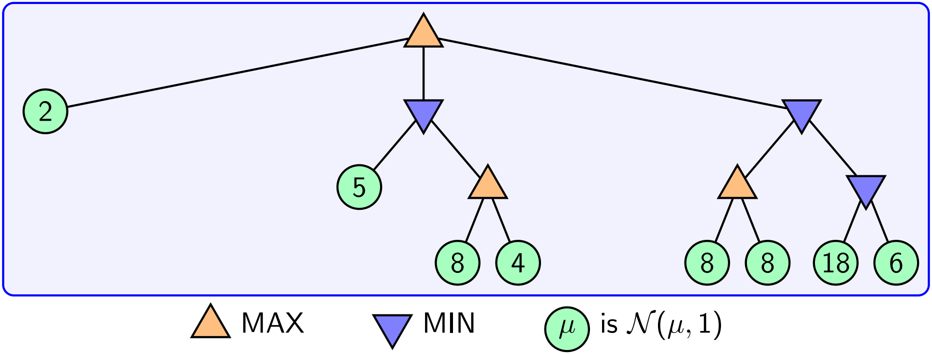
safest next robot action?



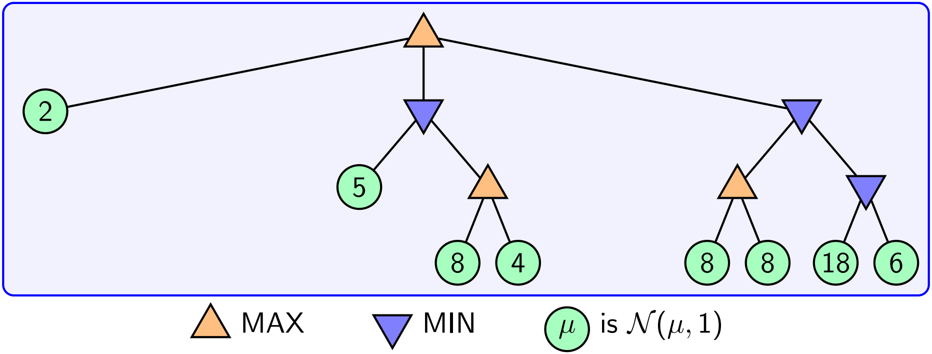
Main scientific questions

- **Efficient** systems
- **Sample complexity** as function of query and environment

Challenge Environment: Stochastic Game Tree Search



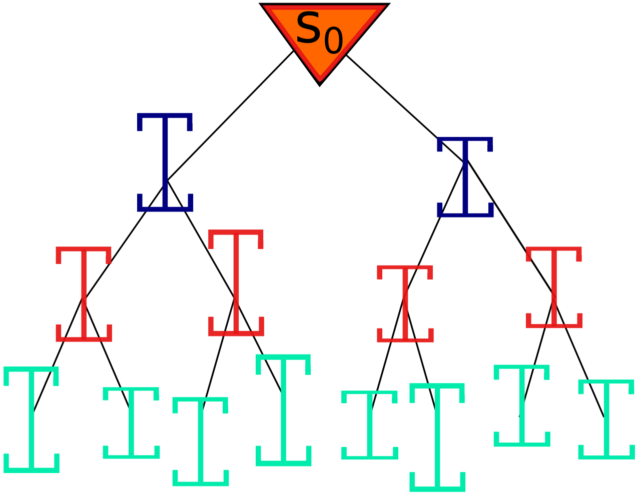
Challenge Environment: Stochastic Game Tree Search



Problem

Determine the optimal move at the root
From **sample access** to the leaf payoffs

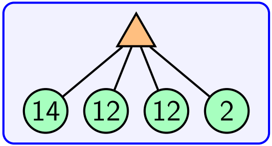
Flashback to Amsterdam Workshop [KK, NIPS'17]



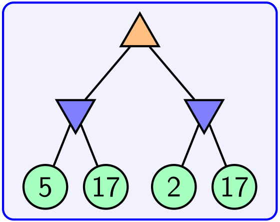
Revisit our Motivating Questions

- Design of pure exploration **algorithms** for **complex queries**?
  - Monte Carlo Tree Search
- Valid anytime **confidence intervals** for **derived quantities**?
  - maximum/minimum

Simplify

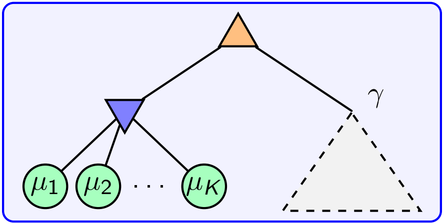


Best Arm Identification
[Garivier and Kaufmann, 2016]
Solved

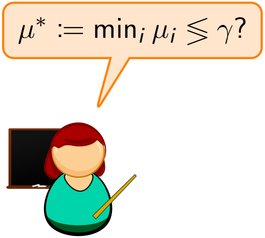


Depth 2
[Garivier, Kaufmann, and Koolen, 2016]
Open

Simple Instance: Minimum Threshold Identification



Fix threshold γ.



For $t = 1, \ldots, \tau$

- Pick leaf A_t
- See $X_t \sim \mu_{A_t}$

Recommend $\hat{m} \in \{<, >\}$

Goal: **fixed confidence** $\mathbb{P}_{\boldsymbol{\mu}} \{\text{error}\} < \delta$
and small **sample complexity** $\mathbb{E}_{\boldsymbol{\mu}}[\tau]$

Lower Bound

Generic lower bound [Castro, 2014, Garivier and Kaufmann, 2016] shows **sample complexity** for **any** δ-correct algorithm is at least

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau] \;\geq\; T^*(\boldsymbol{\mu}) \ln \frac{1}{\delta}.$$

Lower Bound

Generic lower bound [Castro, 2014, Garivier and Kaufmann, 2016] shows **sample complexity** for **any** δ-correct algorithm is at least

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau] \;\geq\; T^*(\boldsymbol{\mu}) \ln \tfrac{1}{\delta}.$$

For our problem the **characteristic time** and **oracle weights** are

$$T^*(\boldsymbol{\mu}) \;=\; \begin{cases} \dfrac{1}{d(\mu^*, \gamma)} & \mu^* < \gamma, \\ \displaystyle\sum_a \dfrac{1}{d(\mu_a, \gamma)} & \mu^* > \gamma, \end{cases} \qquad w_a^*(\boldsymbol{\mu}) \;=\; \begin{cases} \mathbf{1}_{a=a^*} & \mu^* < \gamma, \\ \dfrac{\frac{1}{d(\mu_a,\gamma)}}{\sum_j \frac{1}{d(\mu_j,\gamma)}} & \mu^* > \gamma. \end{cases}$$

Lower Bound

Generic lower bound [Castro, 2014, Garivier and Kaufmann, 2016] shows **sample complexity** for **any** δ-correct algorithm is at least

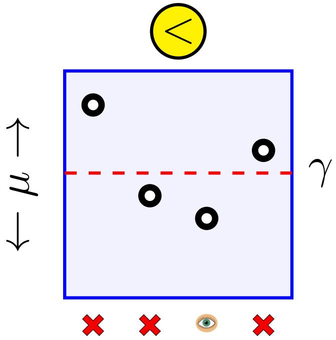
$$\mathbb{E}_{\boldsymbol{\mu}}[\tau] \;\geq\; T^*(\boldsymbol{\mu}) \ln \tfrac{1}{\delta}.$$

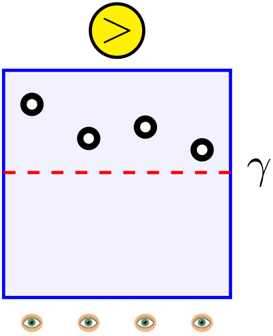
For our problem the **characteristic time** and **oracle weights** are

$$T^*(\boldsymbol{\mu}) \;=\; \begin{cases} \dfrac{1}{d(\mu^*, \gamma)} & \mu^* < \gamma, \\ \displaystyle\sum_a \dfrac{1}{d(\mu_a, \gamma)} & \mu^* > \gamma, \end{cases} \qquad w_a^*(\boldsymbol{\mu}) \;=\; \begin{cases} \mathbf{1}_{a=a^*} & \mu^* < \gamma, \\ \dfrac{\frac{1}{d(\mu_a,\gamma)}}{\sum_j \frac{1}{d(\mu_j,\gamma)}} & \mu^* > \gamma. \end{cases}$$

Lower-order refinements with Rémy Degenne (postdoc @ CWI ML)

Dichotomous Oracle Behaviour! Sampling Rule?





Sampling Rules

- **Lower Confidence Bounds**
  Play $A_t = \arg\min_a \ \mathrm{LCB}_a(t)$
- **Thompson Sampling** (Π_{t-1} is posterior after $t-1$ rounds)
  Sample $\boldsymbol{\theta} \sim \Pi_{t-1}$, then play $A_t = \arg\min_a \ \theta_a$.
- **Murphy Sampling** **condition on low minimum mean**
  Sample $\boldsymbol{\theta} \sim \Pi_{t-1}\left(\cdot | \min_a \theta_a < \gamma\right)$, then play $A_t = \arg\min_a \ \theta_a$. new

Intuition for Murphy Sampling

- When $\mu^* < \gamma$ conditioning is immaterial: $\boldsymbol{\theta} \approx \boldsymbol{\mu}$ and MS $\equiv$ TS.
- When $\mu^* > \gamma$ conditioning results in $\boldsymbol{\theta} \approx (\mu_1, \ldots, \gamma, \ldots, \mu_K)$. Index a lowered to γ with probability $\propto \frac{1}{d(\mu_a, \gamma)}$ [Russo, 2016].

Main Result 1 : Murphy Sampling Rule [KKG, NIPS'18]

Theorem

Asymptotic optimality: $N_a(t)/t \to w_a^*(\boldsymbol{\mu})$ *for all* $\boldsymbol{\mu}$

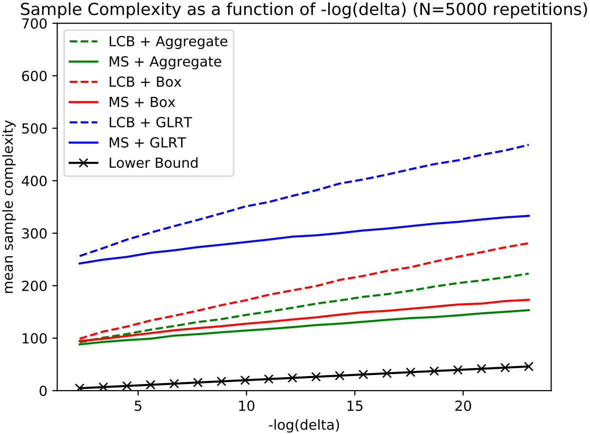
Sampling rule	$<$	$>$
Thompson Sampling	✓	✗
Lower Confidence Bounds	✗	✓
Murphy Sampling	✓	✓

Lemma

Any anytime sampling strategy $(A_t)_t$ *ensuring* $\frac{\boldsymbol{N}_t}{t} \to \boldsymbol{w}^*(\boldsymbol{\mu})$ *and* **good** *stopping rule* τ_δ *guarantee* $\limsup_{\delta \to 0} \frac{\tau_\delta}{\ln \frac{1}{\delta}} \leq T^*(\boldsymbol{\mu})$.

Numerical Results: sample complexity on $<$

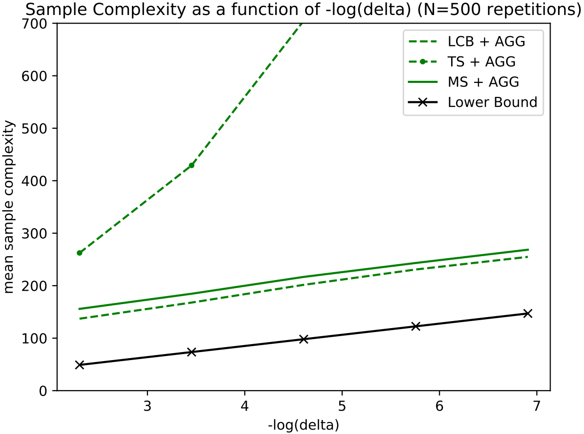
$$\boldsymbol{\mu} = \text{linspace}(-1, 1, 10) \in \mathcal{H}_<$$



Sample complexity $\mathbb{E}[\tau_\delta]$ as a function of $\ln(1/\delta)$. Throughout $\gamma = 0$.

Numerical Results: sample complexity on $>$

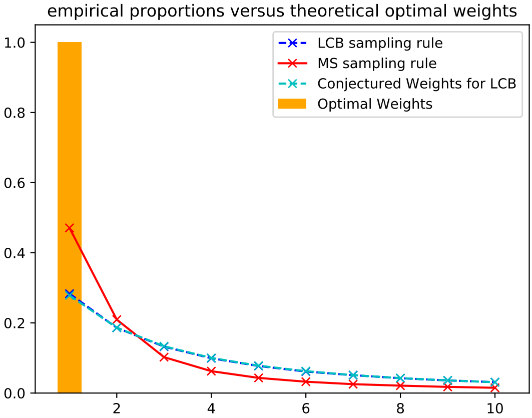
$$\boldsymbol{\mu} = \text{linspace}(1/2, 1, 5) \in \mathcal{H}_>$$



Sample complexity $\mathbb{E}[\tau_\delta]$ as a function of $\ln(1/\delta)$. Throughout $\gamma = 0$.

Numerical Results: proportions on $<$

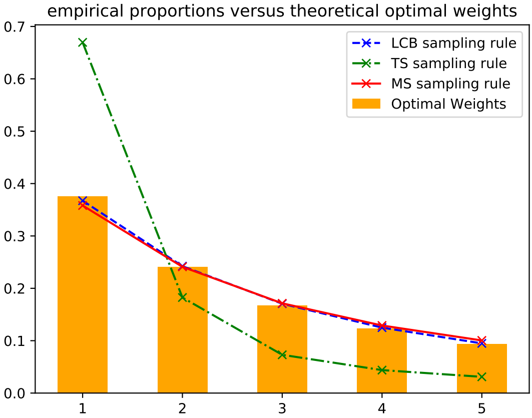
$$\boldsymbol{\mu} = \text{linspace}(-1, 1, 10) \in \mathcal{H}_<$$



Sampling proportions vs oracle, $\delta = e^{-23}$.

Numerical Results: proportions on $>$

$$\boldsymbol{\mu} = \text{linspace}(1/2, 1, 5) \in \mathcal{H}_>$$



Sampling proportions vs oracle, $\delta = e^{-7}$.

(Non-Asymptotic) Adaptivity

Multiple low arms
identical or similar

$\Rightarrow \begin{cases} \text{conclude } \mu^* < \gamma \textbf{ faster} \\ \textbf{tighter} \text{ confidence interval for } \mu^* \end{cases}$?

Confidence Interval for Minimum

For **LCB** we adopt the obvious $\mathrm{LCB}_{\mathsf{min}}(t) = \min_a \mathrm{LCB}_a(t)$.

For **UCB** we investigate three approaches:

- **Box**: Straightforward idea: $\mathrm{UCB}_{\mathsf{min}}(t) = \min_a \mathrm{UCB}_a(t)$.
- **GLRT**: New sum-of-deviations confidence bound.
- **Agg**: Pool samples from multiple arms. Upper bound on **any average** is upper bound on **minimum**. **Biased** but **narrower**.

Main Result 2: Deviation Inequalities [KKG, NIPS'18]

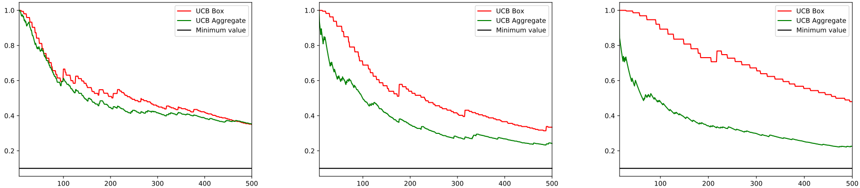
We identify a **threshold function** $T(x) = x + o(x)$ such that for every fixed subset $\mathcal{S} \subseteq [K]$, w.h.p. $\geq 1 - \delta$,

$$\forall t : \left[N_{\mathcal{S}}(t) d^+\big(\hat{\mu}_{\mathcal{S}}(t), \min_{a \in \mathcal{S}} \mu_a\big) - \ln \ln N_{\mathcal{S}}(t)\right]^+ \leq T\left(\ln \tfrac{1}{\delta}\right),$$

$$\forall t : \sum_{a \in \mathcal{S}} \left[N_a(t) d^+\big(\hat{\mu}_a(t), \min_{a \in \mathcal{S}} \mu_a\big) - \ln \ln N_a(t)\right]^+ \leq |\mathcal{S}| T\left(\frac{\ln \frac{1}{\delta}}{|\mathcal{S}|}\right).$$

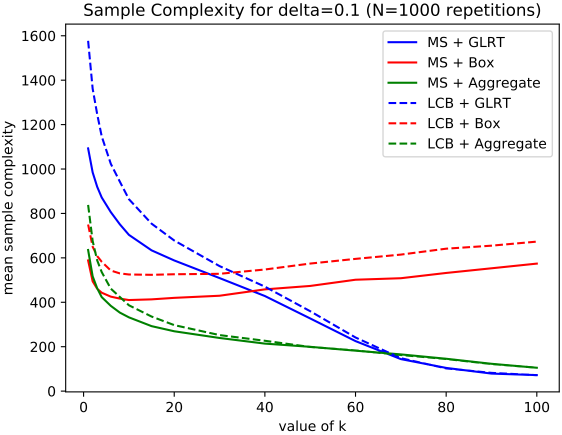
Weighted union bound over subsets **learns** useful low-mean arms.

Numerical Results



UCB for minimum: Agg **dominates** Box with 1, 3 and 10 low arms.

Numerical Results



Agg **beats** Box and GLRT in adapting to the number k of low arms. Here $\mu_a \in \{-1, 0\}$ and $\gamma = 0$.

What's Next

- Deep trees
  Extension to **regular** *depth 2 by Federico Girotti (MSc @ U. Milan)*
- Adaptive tree expansion
- Foundation for MCTS and RL

What's Next

- Deep trees
  Extension to **regular** *depth 2 by Federico Girotti (MSc @ U. Milan)*
- Adaptive tree expansion
- Foundation for MCTS and RL

Thank you!