PANDORA'S BOX WITH CORRELATIONS: LEARNING AND APPROXIMATION

Shuchi Chawla

University of Wisconsin-Madison

Based on joint work with Eva Gergatsouli, Yifeng Teng, Christos Tzamos, and Ruimin Zhang

PANDORA'S BOX PROBLEM [WEITZMAN'79]

- Given: boxes with (random) rewards drawn from known distributions; Can open each box at some fixed penalty
- Goal: select a single box to maximize the reward obtained minus total probing penalty

$R_1 \sim \mathcal{D}_1$ $\quad$ $\mathcal{D}_2$ $\quad$ $R_3 \sim \mathcal{D}_3$ $\quad$ $\mathcal{D}_4$ $\quad$ $\mathcal{D}_5$

t_1 $\quad$ t_2 $\quad$ t_3 $\quad$ t_4 $\quad$ t_5

Algorithm's net reward = $R_3 - (t_1 + t_3)$

PANDORA'S BOX PROBLEM: MINIMIZATION VERSION

- Given: boxes with (random) costs drawn from known distributions; Can open each box at some fixed penalty
- Goal: select a single box to minimize the cost incurred plus total probing penalty

$c_1 \sim \mathcal{D}_1$	$\mathcal{D}_2$	$c_3 \sim \mathcal{D}_3$	$\mathcal{D}_4$	$\mathcal{D}_5$
t_1	t_2	t_3	t_4	t_5

Algorithm's net cost = $c_3 + (t_1 + t_3)$

Question: what order to probe boxes and when to stop and select one?

WEITZMAN'S SOLUTION

[…, Dumitriu Tetali Winkler'03,
Kleinberg Waggoner Weyl'16,
Singla'18,
Bradac Singla Zuzic'19,
Beyhaghi Kleinberg'19,
Gupta Jiang Scully Singla'19, …]

Weitzman's algorithm:

- Compute an amortized cost (a.k.a. Gittins index).
- Probe boxes in greedy order of increasing amortized cost.
- Stop when an observed cost < all remaining indices. Select box with min observed cost.

Theorem: Weitzman's algorithm is optimal if the cost distributions $\mathcal{D}_1, \dots, \mathcal{D}_n$ are independent.

Our setting: correlated costs

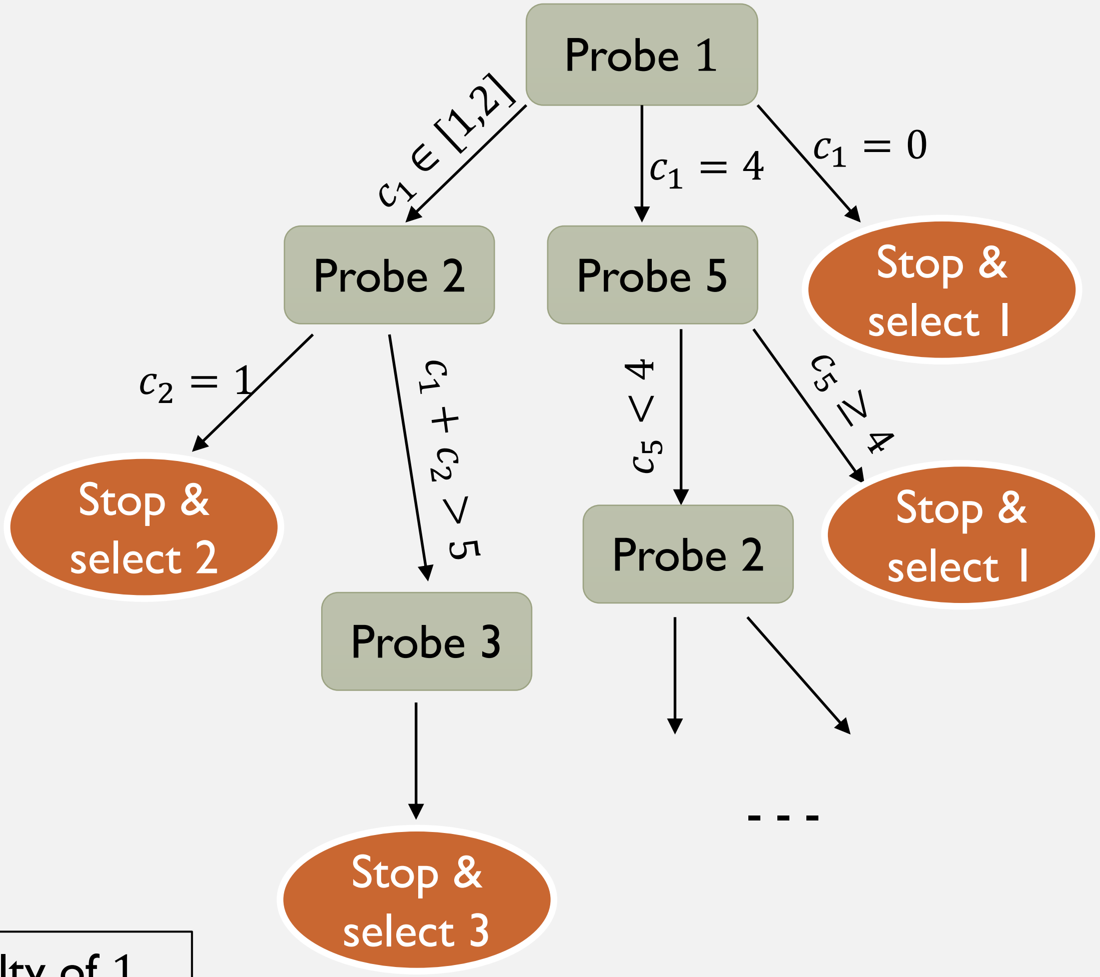
- $(c_1, c_2, \dots, c_n) \sim \mathcal{D}$ where $\mathcal{D}$ is a (large support) joint distribution
- Algorithm is provided sample access to $\mathcal{D}$

FULLY ADAPTIVE SOLUTIONS

- An algorithm is defined by a pair (Π, τ).
- Π: Probing Order over boxes
  - Π_i is a function of $c_{\Pi_1}, c_{\Pi_2}, \ldots, c_{\Pi_{i-1}}$.
- τ: Stopping time
  - At step τ, we stop and select box $\text{argmin}_{i \in [\tau]} \{c_{\Pi_i}\}$.
  - $\mathbb{I}(\tau = i)$ is a function of $c_{\Pi_1}, c_{\Pi_2}, \ldots, c_{\Pi_i}$.
- Objective: minimize $\text{E}[\tau + \min_{i \in [\tau]} \{c_{\Pi_i}\}]$

Probing penalty

Solution cost

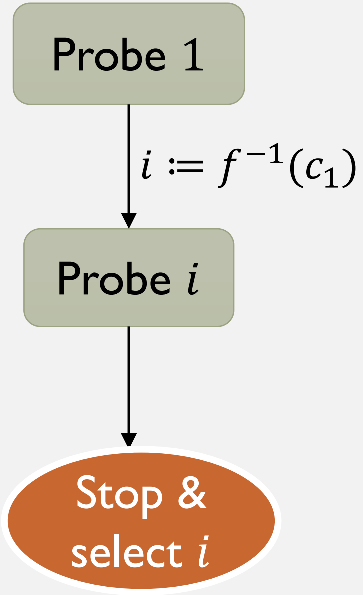


Simplifying assumption for this talk: each box has a probing penalty of 1.

BUT CORRELATED COSTS ARE HOPELESS!

- Let f be some hard to invert function.
- $C^{(i)} = \begin{cases} c_1 = f(i) \\ c_i = 0 \\ c_{i'} = \infty \quad \text{for } i' \neq 1, i \end{cases}$

- OPT = 2
- Alg cannot hope to invert f and find a zero-cost box quickly.



ALTERNATE PLAN: COMPETE AGAINST A SIMPLER BENCHMARK

Related but different: optimal decision tree problem; Assumes small support distribution

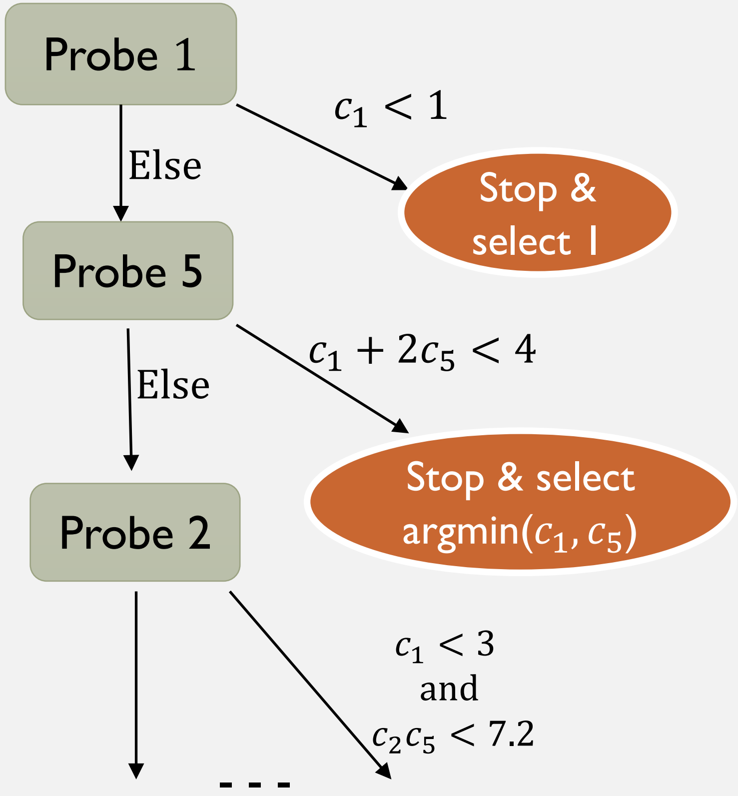
PARTIALLY ADAPTIVE PROBING STRATEGIES

Defined by a pair (Π, τ).

- Π: **Ordering** over boxes
  - Π is independent of instantiated costs.
- τ: **Stopping time**
  - At step τ, we stop and select box $\mathrm{argmin}_{i\in[\tau]}\{c_{\Pi_i}\}$.
  - $\mathbb{I}(\tau = i)$ is a function of $c_{\Pi_1}, c_{\Pi_2}, \ldots, c_{\Pi_i}$.

- Objective: minimize $\mathrm{E}[\,\tau + \min_{i\in[\tau]}\{c_{\Pi_i}\}\,]$ over PA strategies

- Stopping rule can still be quite complicated. Unclear if we can learn it with low sample complexity, or even represent it succinctly.



MAIN RESULT

[Chawla, Gergatsouli, Teng, Tzamos, Zhang'20]

There exists a simple class of PA strategies $\mathcal{C}$ with the following properties:

Theorem 1: For every joint distribution over costs, $\mathcal{C}$ contains an $\frac{e}{e-1}$ approximate strategy.

Theorem 2: Learning the optimal strategy in $\mathcal{C}$ requires $\mathrm{poly}(n)$ samples.

Theorem 3: Given a small support distribution over costs, can efficiently approximate the optimal strategy in $\mathcal{C}$ to within a small constant $(3 + 2\sqrt{2})$.

CAN LEARN AN APPROXIMATELY OPTIMAL Partially Adaptive STRATEGY EFFICIENTLY FROM DATA

MAIN RESULT

[Chawla, Gergatsouli, Teng, Tzamos, Zhang'20]

There exists a simple class of PA strategies $\mathcal{C}$ with the following properties:

Theorem 1: For every joint distribution over costs, $\mathcal{C}$ contains an ~~$\frac{e}{e-1}$~~ $\mathbf{2}$ approximate strategy.

Theorem 2: Learning the optimal strategy in $\mathcal{C}$ requires $\mathrm{poly}(n)$ samples.

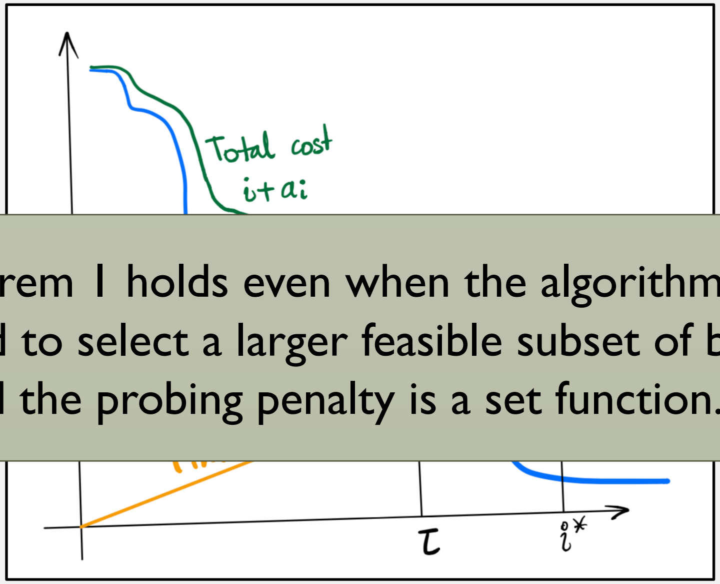
$|\boldsymbol{\mathcal{C}}| = \boldsymbol{n}!$

A strategy in $\mathcal{C}$ is parameterized by the ordering Π.

Stop when probing penalty $>$ solution cost:

$$\tau = \min\{i : i > \min_{j \le i} c_{\Pi_j}\}$$

"Myopic stopping"



Theorem 1 holds even when the algorithm is required to select a larger feasible subset of boxes and the probing penalty is a set function.

Let $a_i = \min_{j \le i} c_{\Pi_j}$ and $i^* = \mathrm{argmin}\{i + a_i\}$.

$\tau \le \max(i^*, a_{i^*})$. $\quad \Rightarrow \tau + a_\tau \le 2(i^* + a_{i^*})$.

EFFICIENT OPTIMIZATION OVER $\mathcal{C}$

Given: uniform distribution over m "scenarios" with cost vectors $c^{(s)} = (c_1^s, c_2^s, \ldots, c_n^s)$ for each scenario $s \in m$.

~~Goal: find a permutation Π such that $(\Pi, \text{myopic stopping})$ is approximately optimal.~~

Goal: find a permutation Π such that $(\Pi, \text{hindsight}-\text{optimal stopping})$ is approximately optimal.

Scenario-aware PA strategy

$\tau_s = \text{argmin}\{i + c_{\Pi_i}^s\}$

Special case: costs are 0 or ∞. "Min sum set cover"

- Minimize the expected time to find a 0, equivalently, "cover" the scenario.
- 4-approx. (tight!) via greedy and LP-rounding. [Feige Lovasz Tetali'02]
- Many variants studied. [Azar Gamzu Yin'09, Bansal Gupta Krishnaswamy'10, Azar Gamzu'11, …]

EFFICIENT OPTIMIZATION OVER $\mathcal{C}$

Given: uniform distribution over m "scenarios" with cost vectors $c^{(s)} = (c_1^s, c_2^s, \dots, c_n^s)$ for each scenario $s \in m$.

~~Goal: find a permutation Π such that $(\Pi, \text{myopic stopping})$ is approximately optimal.~~

Goal: find a permutation Π such that $(\Pi, \text{hindsight}-\text{optimal stopping})$ is approximately optimal.

An LP for scenario-aware PA strategies

$$
\begin{aligned}
\text{minimize} \quad & \sum_{i\in[n],s\in[m],t\in[n]} (t + c_i^s) z_{ist} \\
\text{subject to} \quad & \sum_{i\in[n]} x_{it} \le 1, \\
& \sum_{t\in[n]} x_{it} \le 1, \\
& z_{ist} \le x_{it}, && \forall s \in [m], i,t \in [n] \\
& \sum_{t'\le n, i\in[n]} z_{ist'} = 1, \\
& x_{it}, z_{ist} \in [0,1] && \forall s \in [m], i,t \in [n]
\end{aligned}
$$

Permutation constraints for probing order (first two constraints)

i is selected only if i is probed (third constraint)

At least one box is selected (fourth constraint)

x_{it}: i is probed at time t.

z_{ist}: In scenario s, i is selected at time t.

A RECAP OF OUR RESULTS

Feasibility constraint	**Approx. Ratio** (approx. Partially Adap using Partially Adap)	**Lower bound** (approx. Non Adap using Fully Adap)
Select 1 box	9.22	1.27
Select k boxes	$O(1)$	1.27
Select a matroid basis	$O(\log \text{rank})$	$\Omega(\log \text{rank})$

In each setting:

- Draw $\text{poly}(n)$ samples from distribution. Set up LP on samples and solve.
- Use LP-rounding in phases to find a good probing order.
- Use myopic stopping with the probing order to get final algorithm.

CONCLUDING THOUGHTS

A potential approach to data-driven algorithm design:

Identify a class of algorithms that

- Always contains a near optimal solution
- Has low "complexity" so as to be learnable

Some open directions

- Improved approximation? (through a different "simple" class of algorithms?)
- Are there other benchmarks between Partially Adaptive and Fully Adaptive that are approximable?
- Other combinatorial settings, e.g. metric probing penalties (parking problem)? shortest paths in a graph?

THANK YOU!

Questions?