PANDORA'S BOX:
LEARNING TO LEVERAGE COSTLY INFORMATION

Shuchi Chawla

The University of Texas at Austin

Based on joint work with Eva Gergatsouli, Jeremy McMahan, Yifeng Teng, Christos Tzamos, and Ruimin Zhang
UW-Madison UW-Madison Google UW-Madison U. Chicago

DATA DRIVEN DECISION PROCESSES

Examples of past instances
or
Model of input generating process
or
Predictions about future instance

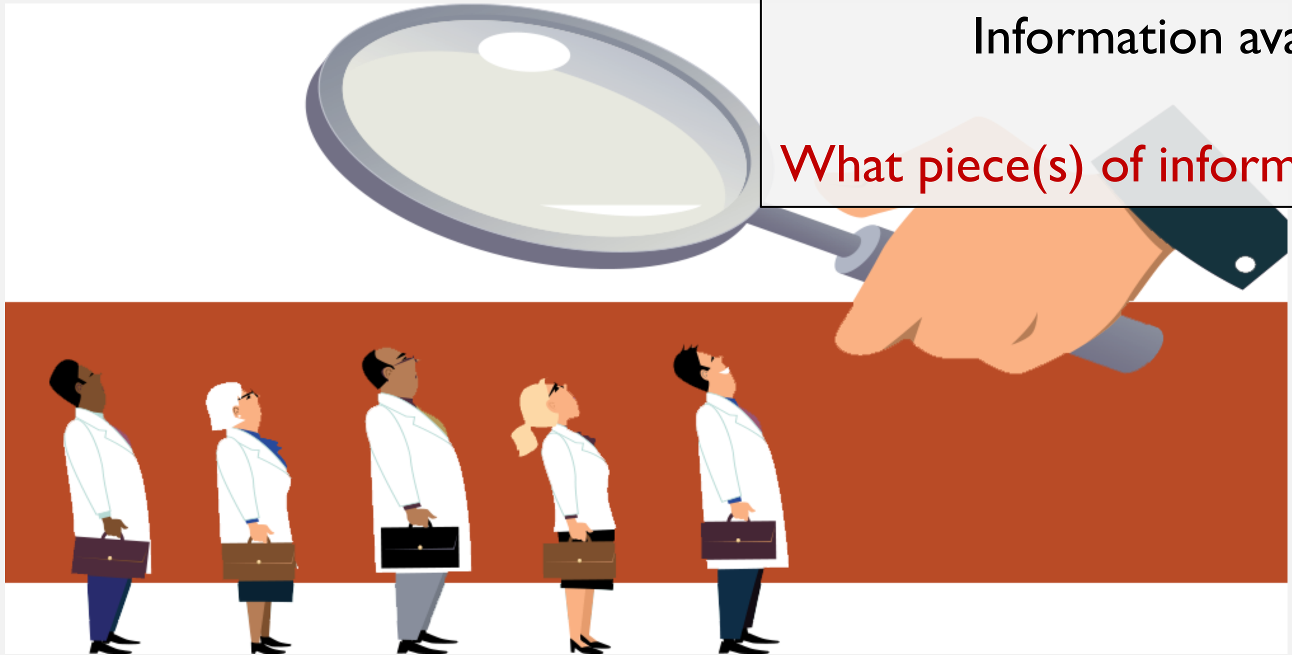
Sequential/Online decision making

Lack of complete knowledge of input

Focus of this talk:

Information available at a cost

What piece(s) of information should we buy?



WEITZMAN'S PANDORA'S BOX PROBLEM

- n boxes containing "costs" drawn from known distributions
- Goal: find the box with smallest cost
- Can open boxes in any order by paying a "probing penalty"

$c_1 \sim \mathcal{D}_1$ $\qquad$ $\mathcal{D}_2$ $\qquad$ $c_3 \sim \mathcal{D}_3$ $\qquad$ $\mathcal{D}_4$ $\qquad$ $\mathcal{D}_5$

t_1 $\qquad$ t_2 $\qquad$ t_3 $\qquad$ t_4 $\qquad$ t_5

Algorithm's net loss = cost of chosen box + total probing penalty $= c_3 + (t_1 + t_3)$

Want to minimize this in expectation

SOLUTIONS TO PANDORA'S BOX: DECISION TREES

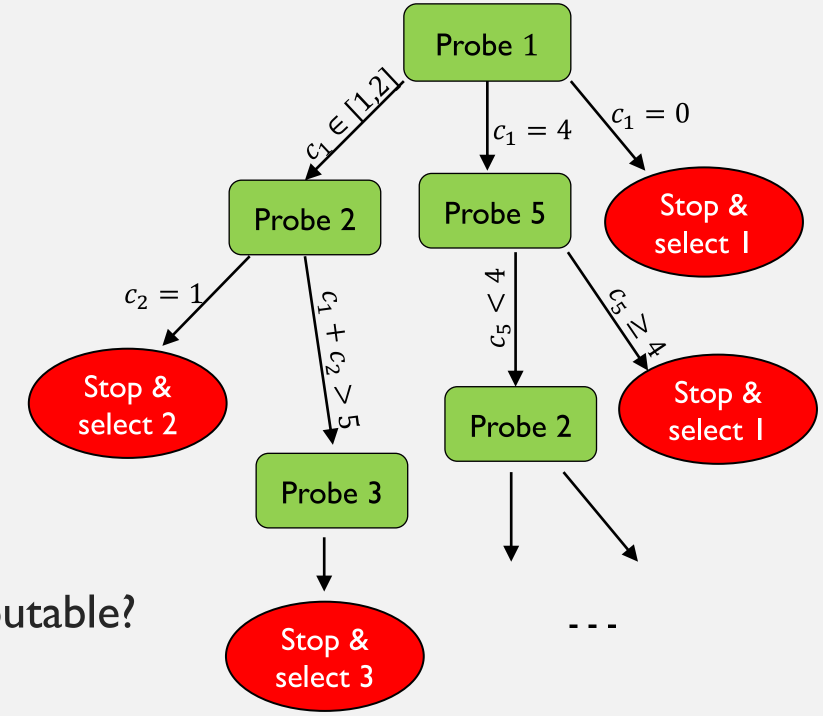
Two components of any algorithm:

- What to probe next?
- When to stop and choose?

A primary source of difficulty: both decisions can depend adaptively on information obtained previously.

Can the tree be large? Is it succinctly representable? Is it efficiently computable?

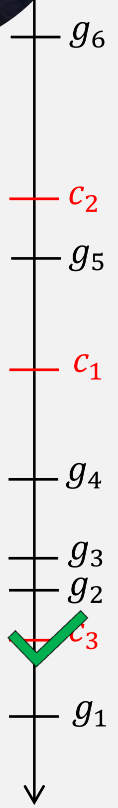
Upshot: in general, hard to compute or learn



WEITZMAN'S SOLUTION (1979)

- Compute an amortized cost (a.k.a. Gittins index). g_i is a function of only $\mathcal{D}_i$ and t_i.
- Probe boxes in greedy order of increasing amortized cost.
- Stop when an observed cost < all remaining indices. Select box with min observed cost.

Note: no adaptivity in probing order!



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

But it fails as soon as we modify the model, e.g.:

This talk

Non-obligatory inspection

i.e. can select boxes without opening them

[Doval'18, Beyhaghi Kleinberg'19]

Correlated costs

E.g. c_1 being high implies c_2 is also high

[C. Gergatsouli Teng Tzamos Zhang'20, C. Gergatsouli McMahan Tzamos'22]

Combinatorial/correlated selection

E.g. boxes are edges in a network and want to select a shortest path; boxes are attributes, and must select all or none

[Singla'18, Klabjan Olszewski Wolinsky'14]

A ROUGH OUTLINE

- Distributional models
- Some components of a Pandora's Box algorithm
  - Challenges
  - Related optimization problems
- Benchmarks
- Putting everything together and some results
- Open questions

Objective:

Develop an algorithm to efficiently compute and learn an approximately optimal probing strategy

MODELING THE UNCERTAINTY IN COSTS

$(c_1, c_2, \dots, c_n) \sim \mathcal{D}$

Explicitly-described distribution

- $\mathcal{D}$ is a "small support" distribution over m scenarios or states of the world.
- The size of the input, and therefore also the complexity of the problem, depend on m.

Arbitrary distribution with sample access

- The distribution $\mathcal{D}$ is arbitrary and we have no direct access to it
- But we are given m "samples" drawn from it.
- Each sample is a possible scenario or state of the world.

Example: $\mathcal{D} = \mathrm{Unif}\begin{pmatrix} 1, 4, 10 \\ 8, 2, 10 \\ 8, 4, 1 \\ 8, 4, 10 \\ 10, 2, 10 \end{pmatrix}$

Terminology cheat-sheet:

- Alg's loss = cost of chosen box + total probing penalty
- Scenario or state of the world ≡ particular realization of costs in boxes

COMPONENTS OF A PANDORA'S BOX ALGORITHM

- Exploration: gathering data about which "world" we're in
- Exploitation: opening a low cost box quickly
- Knowing when to stop
- Learning from limited data: sample access to the distribution

Terminology cheat-sheet:

- Alg's loss = cost of chosen box + total probing penalty
- Scenario or state of the world ≡ particular realization of costs in boxes

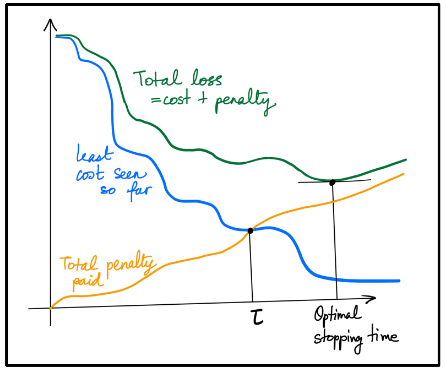
COMPONENTS OF A PANDORA'S BOX ALGORITHM

- Exploration: gathering data about which "world" we're in
- Exploitation: opening a low cost box quickly
- Knowing when to stop
- Learning from limited data: sample access to the distribution

- Optimal stopping rule can be quite complicated: unclear if we can learn it effectively from data, or even represent it succinctly

  But we can effectively approximate it!

- "Myopic" stopping: stop when probing cost exceeds the cost of the best solution found



- Let τ = stopping time. Then $\text{Cost}_\tau = \text{Penalty}_\tau$
- At any other time, either cost or penalty is higher.

$\Longrightarrow$ 2-approx. ($^e/_{e-1}$ with randomness)

[C. Gergatsouli Teng Tzamos Zhang'20]

Upshot: can approximate the **hindsight optimal** stopping rule easily with small loss in approximation factor.

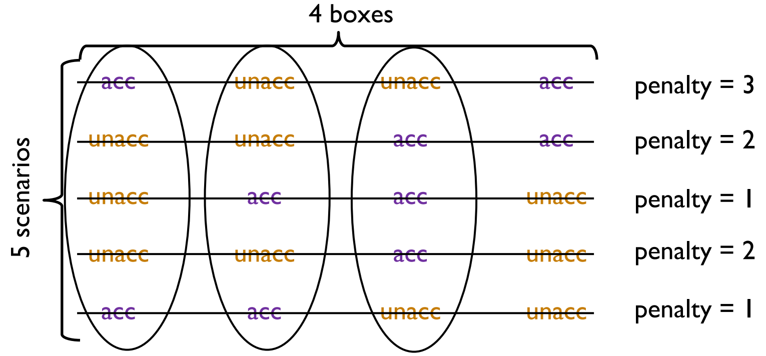
Terminology cheat-sheet:
- Alg's loss = cost of chosen box + total probing penalty
- Scenario or state of the world ≡ particular realization of costs in boxes

COMPONENTS OF A PANDORA'S BOX ALGORITHM

- Exploration: gathering data about which "world" we're in
- Exploitation: opening a low cost box quickly
- Knowing when to stop
- Learning from limited data: sample access to the distribution

A special case: the Min Sum Set Cover problem

- All costs are 0 or ∞. I.e. boxes are "acceptable" or "unacceptable".
- For simplicity: all probing penalties are 1; Uniform distribution over m scenarios.
- Think of the scenarios as elements and boxes as sets. If an "acceptable" box is opened in a scenario, the scenario gets covered.
- Objective: minimize total penalty a.k.a. covering time.



Terminology cheat-sheet:

- Alg's loss = cost of chosen box + total probing penalty
- Scenario or state of the world ≡ particular realization of costs in boxes

COMPONENTS OF A PANDORA'S BOX ALGORITHM

- Exploration: gathering data about which "world" we're in
- Exploitation: opening a low cost box quickly
- Knowing when to stop
- Learning from limited data: sample access to the distribution

A special case: the Min Sum Set Cover problem

- All costs are 0 or ∞. I.e. boxes are "acceptable" or "unacceptable".
- For simplicity: all probing penalties are 1; Uniform distribution over m scenarios.
- Think of the scenarios as elements and boxes as sets. If an "acceptable" box is opened in a scenario, the scenario gets covered.
- Objective: minimize total penalty a.k.a. covering time.

NP-Hard but approximable. [Feige Lovasz Tetali'02]

E.g. greedy 4-approx.: at every step, pick the box that covers the most remaining scenarios

Alternatively: LP-relaxation and rounding.

But doesn't capture the richness of informational structure in Pandora's Box

Terminology cheat-sheet:

- Alg's loss = cost of chosen box + total probing penalty
- Scenario or state of the world ≡ particular realization of costs in boxes

COMPONENTS OF A PANDORA'S BOX ALGORITHM

- Exploration: gathering data about which "world" we're in
- Exploitation: opening a low cost box quickly
- Knowing when to stop
- Learning from limited data: sample access to the distribution

Related problem: the Optimal Decision Tree problem

Use experiments/tests to discover which of m given hypotheses is true

Many applications: experiment design; disease diagnosis; fault-tolerant computing

NP-hard even when tests are binary and prior is uniform [Hyafil Rivest'76]

Some approaches:

- Greedy w.r.t. a "nice" proxy for information gain
- Successive elimination of scenarios

m states of the world $\Longrightarrow O(\log m)$-approximation. [Gupta Nagarajan Ravi'18, Li Liang Mussmann'20]

Tight in general! [Chakaravarthy et al.'07]

- Open: constant approx. when the prior over hypotheses is uniform?

Upshot: exploration/isolation can be solved approximately when there are "few" possible states of the world.

COMPONENTS OF A PANDORA'S BOX ALGORITHM

- Exploration: gathering data about which "world" we're in
- Exploitation: opening a low cost box quickly

Can solve approximately when there are few states of the world.

- Knowing when to stop

Can approximate hindsight optimal stopping rule without any stochastic info.

- Learning from limited data: sample access to the distribution

??

APPROXIMATION FOR THE EXPLICIT DISTRIBUTION SETTING

Explicitly-described distribution: $\mathcal{D}$ is a "small support" distribution over m scenarios or states of the world.

Pandora's Box $\Leftarrow\Rightarrow$ Min Sum Set Cover Min Sum Set Cover with "feedback" $\Rightarrow$ Optimal Decision Tree

(log log m loss in approximation)

"Guess" an acceptable-cost threshold that covers a constant fraction of scenarios; Iterate.

Theorem: $O(\log m \log\log m)$-approximation where m = number of scenarios.

[C. Gergatsouli McMahan Tzamos'22]

Can we say anything interesting in the large support, sample access setting?

Terminology cheat-sheet:

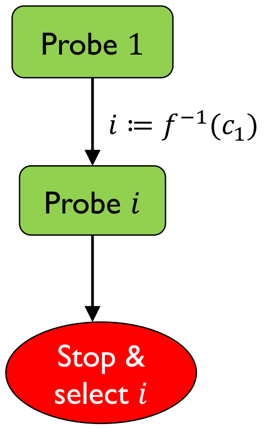
- Alg's loss = cost of chosen box + total probing penalty
- Scenario or state of the world ≡ particular realization of costs in boxes

COMPONENTS OF A PANDORA'S BOX ALGORITHM

- Exploration: gathering data about which "world" we're in
- Exploitation: opening a low cost box quickly
- Knowing when to stop
- Learning from limited data: sample access to the distribution

Upshot: competing against the optimum is hopeless!

- Let f be a random mapping from $[n]$ to a large domain.
- $$C^{(i)} = \begin{cases} c_1 = f(i) \\ c_i = 0 \\ c_{i'} = \infty & \text{for } i' \neq 1, i \end{cases}$$
- OPT = 2
- Alg cannot hope to invert f and find a zero-cost box with few samples.



COMPONENTS OF A PANDORA'S BOX ALGORITHM

- Exploration: gathering data about which "world" we're in
- Exploitation: opening a low cost box quickly

Can solve approximately when there are few states of the world.

- Knowing when to stop

Can approximate hindsight optimal stopping rule without any stochastic info.

- Learning from limited data: sample access to the distribution

Cannot learn any approx. optimal strategy from few samples.

BENCHMARK REVISITED

Fully Adaptive Opt (FA-OPT)

- Probing order and stopping rule depend adaptively on obtained info.

Hard to approximate, and hard to learn from data

Inspired by Weitzman's solution

Partially Adaptive Opt. (PA-OPT)

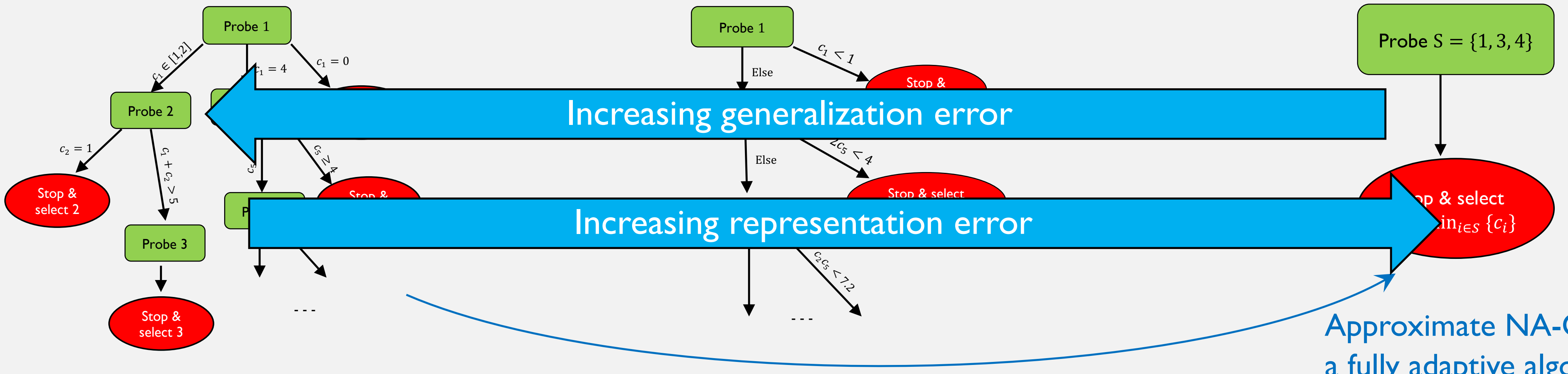
- Probing order is non-adaptive
- Stopping rule depends adaptively on obtained info.

Can approximate and learn from data!

Non Adaptive Optimum (NA-OPT)

- Selects a subset of boxes
- Opens all simultaneously
- Chooses the best box

Hard to approximate better than $\Omega(\log m)$ (captures set cover); Easy to learn from data.



Approximate NA-OPT using a fully adaptive algorithm?

LEARNING AND APPROXIMATING PA-OPT IN THE SAMPLES SETTING

Arbitrary distribution with sample access: given m "samples" drawn from $\mathcal{D}$. Each sample is a possible scenario or state of the world.

1. Simplify stopping rule at small expense: use myopic stopping
2. Only $n!$ different PA strategies with myopic stopping; algorithm's total loss is bounded

   $\Longrightarrow$ $poly(n)$ samples suffice to learn optimal PA strategy.
3. Draw $m = poly(n)$ samples from distribution.
4. Use LP relaxation and rounding to find optimal PA strategy over samples.

   Inspired by algorithms for MSSC.

   $\Longrightarrow$ $(3 + 2\sqrt{2})$-approximation to the optimal PA strategy over samples.

Theorem: Constant approximation to PA-OPT in polynomial time using polynomial # of samples.

[C. Gergatsouli Teng Tzamos Zhang'20]

MODELING THE UNCERTAINTY IN COSTS (REVISITED)

$(c_1, c_2, \ldots, c_n) \sim \mathcal{D}$

Explicitly-described distribution

- $\mathcal{D}$ is a "small support" distribution over m scenarios or states of the world.

$O(\log m \log\log m)$-approximation to FA-OPT

Arbitrary distribution with sample access

- $\mathcal{D}$ is arbitrary; We are given m "samples" drawn from it.
- Each sample is a possible scenario or state of the world.

$O(1)$-approximation to PA-OPT

Product distribution

- Weitzman's algorithm gives an exact solution

FA-OPT = PA-OPT

Mixture of m product distributions

- Again exhibits explore-exploit tradeoff. Related to "noisy" ODT.
- Challenge: aggregating weak signals about the instantiated scenario. Need a "gap" assumption!
- $O\left(\frac{\log m}{\text{gap}}\right)$-approximation in polynomial time. [Gan Jia Li'21]
- $(1+\epsilon)$-approximation in time $\exp(\text{poly}(m/\text{gap}) \log\left(\frac{1}{\epsilon}\right))$

[C. Gergatsouli McMahan Tzamos'22]

FURTHER DIRECTIONS AND OPEN QUESTIONS

- Other models of correlation or noise; better results for mixture of product distributions?

  [e.g. Liang Mu Syrgkanis'21, Bardhi'22]

- Other benchmarks?

- Combinatorial selection, e.g., selecting k boxes or matching in a network, etc.
  - Techniques generally extend to matroid-type settings [e.g. Singla'16, C. Gergatsouli Teng Tzamos Zhang'20]
  - Beyond matroids?

- Pandora's Box internalizes the cost-of-information versus optimization tradeoff

  Alternate approach to handle costly information: budgeted/constrained probing
  - Can we define Pandora-style models for other models of costly information, e.g. query complexity, sample complexity, …

THANK YOU!

Questions?