Memory Bounds for the Experts Problem

Vaidehi Srinivas, David P. Woodruff, Ziyu Xu, Samson Zhou

Abstract

Online learning with expert advice is a fundamental problem of sequential prediction with wide-ranging applications in machine learning, optimization, economics, and beyond. Classical algorithms for this problem have been well-studied in many fields since as early as the 1950s. However, every existing variation, to our knowledge, requires $\Omega(n)$ memory. We initiate the study of this problem in the streaming setting, and show upper and lower bounds.

Prediction with Expert Advice

Algorithm can access n experts, and must make $\{0,1\}$ predictions for T days. Each day,

1. algorithm sees the predictions of each expert,
2. makes a prediction,
3. then sees true outcome of that day, and incurs cost 1 if its prediction was incorrect.

We bound (average) **regret**:

$$\frac{\text{cost more than best expert}}{\#\text{ days}}.$$

In **online learning** version, algorithm commits to a particular expert, then true $[0,1]$ cost of each expert is revealed.

Classical Algorithms

Randomized Multiplicative Weights (MW): For any $\varepsilon > 0$, can construct an algorithm A, such that

$$\mathbf{E}\big[\#\text{ mistakes by } A\big] \le (1+\varepsilon)(\#\text{ mistakes by best expert}) + \frac{\ln n}{\varepsilon}.$$

This achieves $O\left(\sqrt{\frac{\ln n}{T}}\right)$ regret (best possible), using $\Omega(n)$ memory.

Similar guarantees can be achieved by **follow the perturbed leader.**

Streaming Models

Complete sequence of T days is the **data stream**:

$$(\text{prediction}_1, \text{outcome}_1), \ldots, (\text{prediction}_T, \text{outcome}_T).$$

Algorithm can maintain at most s bits of memory from one day to the next (size of streaming state).

Arbitrary-order streams: adversary chooses predictions and outcomes for each of the T days

Random-order streams: adversary chooses T days of predictions and outcomes, and order of days is randomly shuffled

Results

Theorem (Lower Bound)
Any algorithm that achieves average regret δ in expectation, must use

$$\Omega\left(\frac{n}{\delta^2 T}\right) \text{ space,}$$

even for random-order and i.i.d. streams.

Theorem (Upper Bound for Random-Order Streams)
For a target $\delta > \sqrt{\frac{16 \log^2 n}{T}}$, our algorithm achieves average regret δ in expectation, using

$$\widetilde{O}\left(\frac{n}{\delta^2 T}\right) \text{ space.}$$

Theorem (Upper Bound for Arbitrary-Order Streams)
For a target $\delta > \sqrt{\frac{128 \log^2 n}{T}}$, our algorithm achieves average regret δ in expectation using

$$\widetilde{O}\left(\frac{n}{\delta T}\right) \text{ space,}$$

when the best expert makes at most $O(\frac{\delta^2 T}{\ln n})$ mistakes.

Lower Bound

Our lower bound for i.i.d., random-order, and arbitrary-order streams uses a reduction to a custom-built problem using a novel masking technique, to show a smooth tradeoff between regret and memory. Matching the regret of MW requires $\widetilde{\Omega}(n)$ memory, but for constant regret we could do much better.

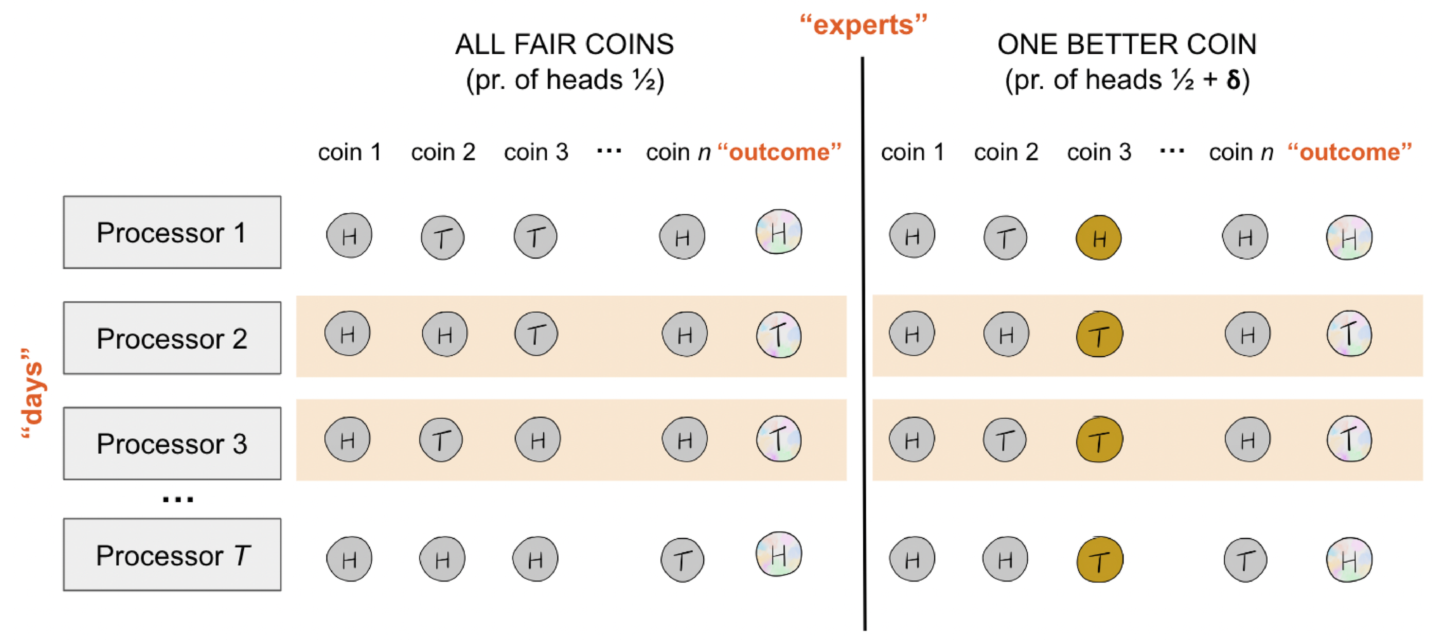

Figure 1: Reduction from communication problem to expert prediction. “Outcome” is chosen to be heads on every day. Then a **mask** is chosen uniformly at random for each day, and applied to predictions and outcome for that day (shown as orange bars).

δ-DiffDist problem: distinguish between inputs from ALL FAIR COINS case and ONE BETTER COIN case

- One index (one coin) problem requires $\Omega(1/\delta^2)$ communication, so n index (n coin) problem requires $\Omega(n/\delta^2)$ communication via *direct sum*
- δ-regret expert prediction algorithm with s bits of memory will predict well if and only if in ONE BETTER COIN case using sT bits of communication

Thus, any δ-regret expert prediction algorithm must use $s \in \Omega(n/\delta^2 T)$ bits of memory.

Upper Bounds

Our upper bounds show novel ways to run standard expert prediction algorithms on small “pools” of experts. For random-order streams, our upper bound is tight up to low order terms. For arbitrary-order streams, in the regime where the best expert makes a slightly subconstant fraction of mistakes, our upper bound beats our lower bound.

Let M be the number of mistakes the best expert makes. (For random-order, can find M using guess-and-double. For arbitrary-order, assume M is small.)

Goal: Given M, make at most $M + O(\delta T)$ mistakes

Strategy:

- Run MW on randomly sampled **“pool”** of $\frac{n \ln n}{\delta^2 T}$ experts
- If/when every expert in pool has average mistake rate $\ge \frac{M}{T} + \delta$, resample and start over

Total cost:

$$(1+\delta)[M + \delta T] + \frac{\ln n}{\delta}(\#\text{ rounds of sampling})$$

Random-order streams: with high probability, once we catch the best expert, we never resample (see Figure 2).

- Probability best expert caught in given pool: $\frac{n \ln n/(\delta^2 T)}{n}$
- Expected # rounds: $\frac{\delta^2 T}{\ln n}$
- Total cost: $M + O(\delta T)$

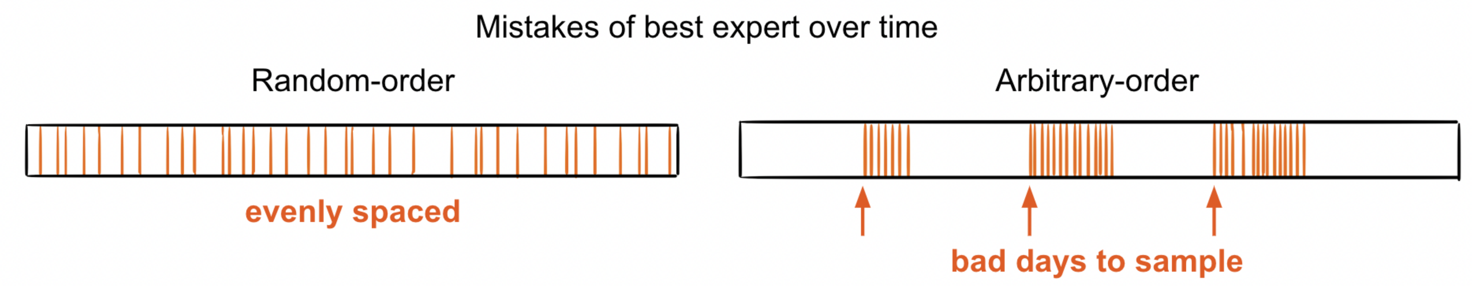

Figure 2: For random-order streams, whenever we catch the best expert, it will neve look like it is doing badly. For arbitrary-order streams, there may be some misleading “bad days,” but not too many.

Arbitrary-order streams: assume best expert makes $O(\frac{\delta^2 T}{\ln n})$ mistakes. There cannot be too many rounds until the best expert is caught forever.

- Total cost is dominated by regret, so pool size can be smaller
- Best expert makes few mistakes $\Longrightarrow$ not too many “bad days”
- Total cost does not increase too much

Open Questions

- **Tight bounds for arbitrary-order streams**

  For constant regret δ, can we tolerate the best expert making a constant fraction of mistakes?

- **Better bounds when expert costs have more structure**

  i.e. expert predictions are real numbers that are evaluated against the true outcome with some loss function