Sequential Information Maximization

When is Greedy Near-optimal?

Yuxin Chen†, S. Hamed Hassani†, Amin Karbasi‡, Andreas Krause†

| COLT, Paris | July 5th, 2015 |

† ETH Zürich
‡ Yale University
The Sequential Information Maximization Problem
The Sequential Information Maximization Problem
The Sequential Information Maximization Problem
The Sequential Information Maximization Problem

\[ \mathbb{H}(Y \mid \emptyset) = 2 \]
The Sequential Information Maximization Problem

\[ \mathbb{H}(Y \mid \emptyset) = 2 \]
The Sequential Information Maximization Problem

\[ H(Y \mid \emptyset) = 2 \]
The Sequential Information Maximization Problem

$$\mathbb{H}(Y \mid \text{Obs}_1) = 1.2$$
The Sequential Information Maximization Problem

\( \mathbb{H}(Y \mid \text{Obs}_1) = 1.2 \)
The Sequential Information Maximization Problem

\[ \mathbb{H}(Y \mid \text{Obs}_1) = 1.2 \]
The Sequential Information Maximization Problem

\[
\text{\texttt{\textbf{X}}}_1^0 \rightarrow 1 \rightarrow \text{\texttt{\textbf{X}}}_2^1 \rightarrow 0 \rightarrow 1 \rightarrow \text{\texttt{\textbf{X}}}_2^1
\]

\[
\mathbb{H}(Y \mid \text{Obs}_2) = 1.4
\]
The Sequential Information Maximization Problem

\[ \mathbb{H}(Y \mid \text{Obs}_2) = 1.4 \]
The Sequential Information Maximization Problem

\[ \mathbb{H}(Y \mid \text{Obs}_2) = 1.4 \]
The Sequential Information Maximization Problem

\[ \mathbb{H}(X_1 | \text{Obs}_3) = 0.3 \]
The Sequential Information Maximization Problem

\[ H(Y | \text{Obs}_3) = 0.6 \]

Diagram showing nodes labeled $X_1$, $X_2$, $X_3$, $X_5$ connected with edges 0 and 1. Nodes are associated with labels like Bacterial, Viral, Blood, Parasitic.
The Sequential Information Maximization Problem

What is the optimal policy $\pi^*$?
The Sequential Information Maximization Problem

What is the optimal policy $\pi^*$?

**Greedy:** Pick the test with the maximal reduction in entropy, given the past observations.
The most informative selection policy has been used since 1950’s [Lindley, 1956]
The Greedy Algorithm

The most informative selection policy has been used since 1950’s
[Lindley, 1956]

In the **non-adaptive** setting, Greedy is near-optimal

[Krause and Guestrin, 2005]
The Greedy Algorithm

The most informative selection policy has been used since 1950’s [Lindley, 1956]

In the non-adaptive setting, Greedy is near-optimal [Krause and Guestrin, 2005]

In the noiseless setting, Greedy is near-optimal [Dasgupta, 2005; Golovin and Krause 2011]
The Greedy Algorithm

The most informative selection policy has been used since 1950’s [Lindley, 1956]

In the **non-adaptive** setting, Greedy is near-optimal

[Krause and Guestrin, 2005]

In the **noiseless** setting, Greedy is near-optimal

[Dasgupta, 2005; Golovin and Krause 2011]

Tests are **noisy**, but can be **repeated** with i.i.d. outcome

— Reduction to Noiseless Case

[Nowak 2009]
The Greedy Algorithm

The most informative selection policy has been used since 1950’s
[Lindley, 1956]

In the **non-adaptive** setting, Greedy is near-optimal
[Krause and Guestrin, 2005]

In the **noiseless** setting, Greedy is near-optimal
[Dasgupta, 2005; Golovin and Krause 2011]

Tests are **noisy**, but can be **repeated** with i.i.d. outcome
— Reduction to Noiseless Case
[Nowak 2009]

We present **the first rigorous analysis** of the greedy policy
in the **persistent noise** setting.
Main Result

For any $\delta > 0$, it holds that

$$\mathbb{I}(\pi_{\text{Greedy}[k']} ; Y) \geq (\mathbb{I}(\pi_{\text{OPT}[k]} ; Y) - \delta) \left(1 - \exp\left(-\frac{k'}{k} \cdot \frac{S_{\text{min}}}{7 \max\{\log n, \log \frac{1}{\delta}\}}\right)\right)$$

- Gain of the greedy policy in $k'$ steps
- Gain of the optimal policy in $k$ steps
- Number of hypotheses
- Characterizes the severity of noise
To gain information close to OPT \([k]\), we need to run Greedy \(O\left(\frac{k \cdot \log n}{S_{\text{min}}}\right)\) times.
Main Result

To gain information close to OPT \([k]\), we need to run Greedy \(O\left(\frac{k \cdot \log n}{S_{\text{min}}}\right)\) times.

We show that the parameter \(S_{\text{min}}\) is necessary in the bound.
Main Result

To gain information close to OPT \([k]\), we need to run

\[
\text{Greedy } O\left( k \cdot \frac{\log n}{S_{\text{min}}} \right) \text{ times.}
\]

We show that the parameter \(S_{\text{min}}\) is necessary in the bound.

Come to our poster / ICML Greed is Great workshop for more details!