Smooth Interactive Submodular Set Cover

Neural Information Processing Systems Foundation

Bryan He, Yisong Yue

Submodularity

- Many problems formulated as optimizing discrete functions, but discrete optimization is difficult
- Difficulty is avoided by using submodular functions
- Submodular functions have diminishing returns
  - $A \subseteq B \Rightarrow F(A \oplus q) - F(A) \geq F(B \oplus q) - F(A)$
- Also restricted to be monotone non-decreasing
  - $F(A \oplus q) \geq F(A)$
- Approximately optimized with a greedy algorithm
- Common way to model diversity

Interactive Submodular Set Cover (ISSC)

- Class of hypotheses H, and each hypothesis has a submodular utility function F_h
- Interactive process where decision maker selects an action, and environment provides a response according to an unknown true hypothesis h^*
- Requires queries until the utility function of the true hypothesis F_{h^*} is above a pre-selected threshold α
- **Limitation:** requires h^* to be in H

Noisy ISSC

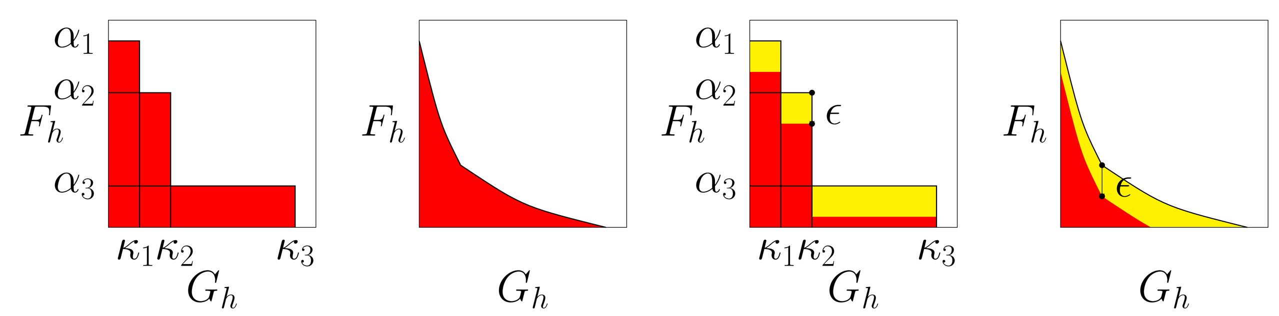
- Removes assumption that $h^* \in H$ by introducing a distance G_h for each hypothesis
- A hypothesis is considered satisfied if $F_h \geq \alpha$ and is considered irrelevant if $G_h \geq \kappa$
- Requires queries until $F_h \geq \alpha$ or $G_h \geq \kappa$ (considers tradeoff between the utility of a hypothesis and the likelihood that it is true)
- Characterizes explore/exploit tradeoff
- **Limitation:** not flexible because of binary decisions (satisfied/unsatisfied and relevant/irrelevant)

Problem Statement

Smooth ISSC is a generalization of noisy ISSC that considers a hypothesis satisfied if $F_h \geq \alpha(G_h)$ for an arbitrary function $\alpha(\cdot)$.

Versions of Smooth ISSC

- multiple thresholds **(a)**
- continuous convex threshold **(b)**
- approximate versions **(c)** and **(d)**



(a) (b) (c) (d)

Multiple Thresholds

- Thresholds are $\alpha_1 > \alpha_2 > \cdots > \alpha_N$ and $\kappa_1 < \kappa_2 < \cdots < \kappa_N$
- Make queries until $F_h \geq \alpha_n$ or $G_h \geq \kappa_n$ for all n and $h \in H$
- Equivalent to selecting

$$\alpha(\kappa) = \begin{cases} \infty & \text{if } \kappa < 0 \\ \alpha_1 & \text{if } 0 \leq \kappa < \kappa_1 \\ \vdots & \vdots \\ \alpha_N & \text{if } \kappa_{N-1} \leq \kappa < \kappa_N \\ 0 & \text{if } \kappa_N \leq \kappa \end{cases}$$

Doubly truncated version of each function:

$$F_{h,\alpha_n,\alpha_j} \triangleq \max(\min(F_h, \alpha_n), \alpha_j) - \alpha_j$$
$$G_{h,\kappa_n,\kappa_j} \triangleq \max(\min(G_h, \kappa_n), \kappa_j) - \kappa_j$$

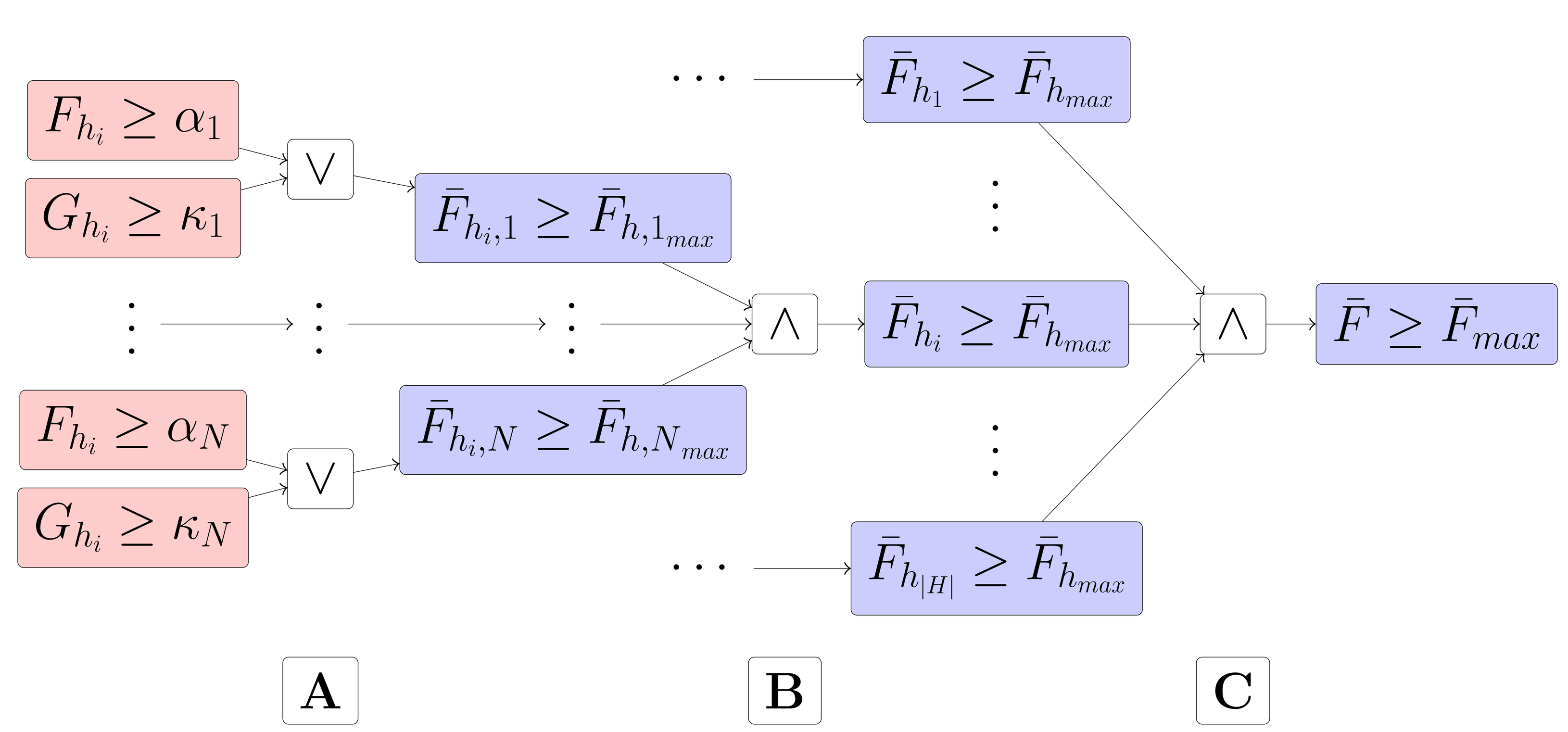
Using the truncated functions, we define $\bar{F}$:

$$\bar{F}_{h,n}(\hat{S}) \triangleq \left((\kappa_n - \kappa_{n-1}) - G_{h,\kappa_n,\kappa_{n-1}}(\hat{S})\right) F_{h,\alpha_n,\alpha_{n+1}}(\hat{S}) + G_{h,\kappa_n,\kappa_{n-1}}(\hat{S})(\alpha_n - \alpha_{n+1}),$$

$$\bar{F}_h(\hat{S}) \triangleq C_{\bar{F}} \sum_{n=1}^{N} \left[\left(\prod_{j \neq n} (\kappa_j - \kappa_{j-1}) \right) \bar{F}_{h,n}(\hat{S}) \right],$$

$$\bar{F}(\hat{S}) \triangleq \sum_{h \in H} \bar{F}_h(\hat{S}), \quad \bar{F}_{max} \triangleq |H| C_F C_G.$$

The following figure shows the relationship between the components of $\bar{F}$.



This indicates that $\bar{F} \geq \bar{F}_{max}$ is equivalent to satisfying $F_h \geq \alpha(G_h)$ for all hypotheses. In addition, $\bar{F}$ is submodular as long as the sequence $\langle \frac{\alpha_n - \alpha_{n+1}}{\kappa_n - \kappa_{n-1}} \rangle_{i=1}^{N}$ is non-increasing (the points (α_n, κ_n) are convex).

General Form of Solution

- Meta-function $\bar{F}$ of F_h, G_h, and $\alpha(\cdot)$ that
  - is submodular
  - reaches maximum value iff $F_h \geq \alpha(G_h)$ for all $h \in H$
- Equivalent to satisfying original requirements
- Reduces to conventional submodular set cover
- Can solve using simple greedy algorithm

Greedy Algorithm

```
input: F̄                                   // Submodular Meta-Objective
input: F̄_max                               // Termination Threshold for F̄
input: Q                                   // Query or Action Set
input: R                                   // Response Set
S ← ∅
while F̄(S) < F̄_max do
    q̂ ← argmax_{q∈Q} min_{r∈R} (F̄(S ⊕ (q,r)) − F̄(S)) / c(q)
    Play q̂, observe r̂
    S ← S ⊕ (q̂, r̂)
end while
```

Approximate Multiple Thresholds

- Terminate once $F_h \geq \alpha(G_h) - \epsilon$ for all $h \in H$
- Define surrogate functions and thresholds

$$F'_h(\hat{S}) \triangleq \frac{D}{\epsilon} \left\lfloor F_h(\hat{S}) + \frac{\epsilon}{D} \sum_{i=1}^{|\hat{S}|} (|\mathcal{Q}| + 1 - i) \right\rfloor_{\frac{\epsilon}{D}},$$

$$\alpha'_n \triangleq \frac{D}{\epsilon} \left\lceil \alpha_n - \frac{\epsilon}{D} \sum_{i=1}^{n} \left[(2N - 2i) D_G^{N-i+1} \prod_{j=i}^{N} (\kappa_j - \kappa_{j-1}) \right] \right\rceil_{\frac{\epsilon}{D}}$$

Continuous Convex Curve

- Reduce curve into equivalent multiple thresholds
- For the inexact case:
  - κ_n are selected as all distinct values of G_h
  - α_n are selected as $\alpha(\kappa_n)$
  - Solve using approximate multi-threshold version
- For the exact case
  - Set ϵ to smallest distinct difference between F_h values

Conclusion

- Smooth ISSC generalizes previous ISSC frameworks
- Allows the target threshold for a hypothesis to vary based on the plausibility of that hypothesis
- Introduces an approximate threshold concept that can be applied to real-valued functions