Satisficing Regret Minimization in Bandits

Qing Feng; Tianyi Ma; Ruihao Zhu

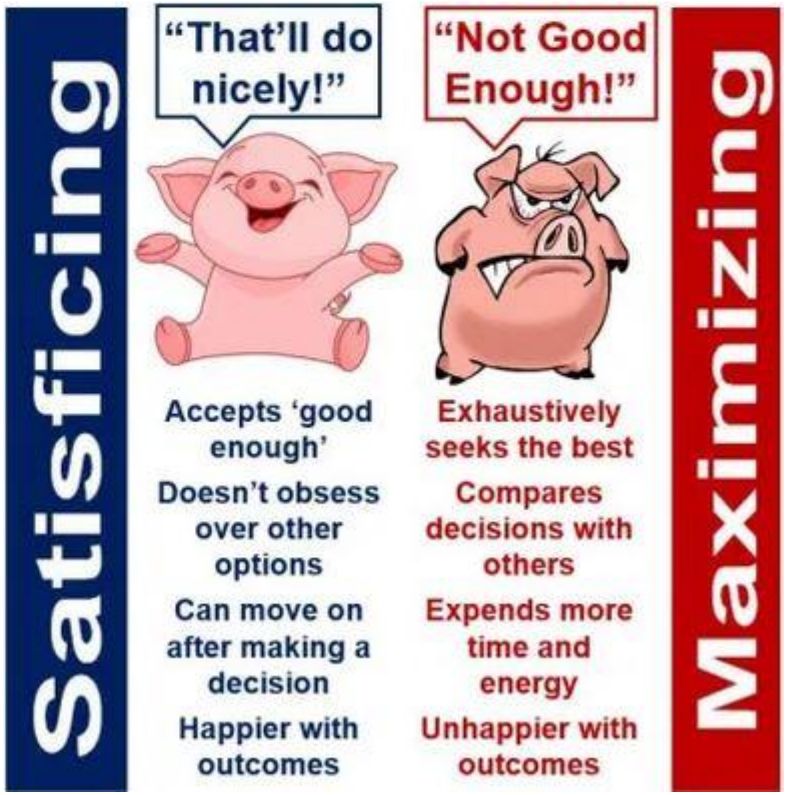
ICLR 2025

Satisficing Bandit - Motivation

Herbert Simon

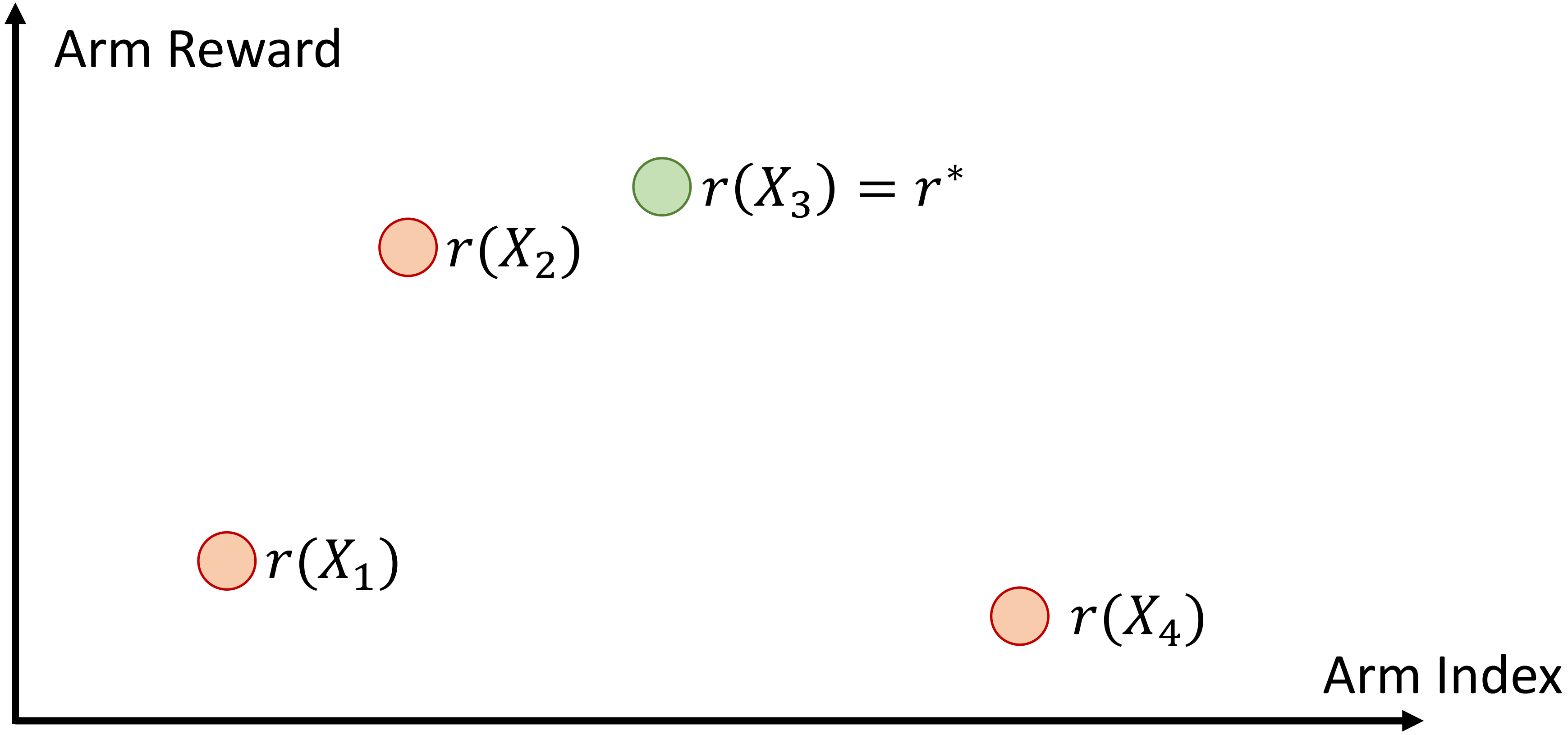
Nobel Prize (1978)

Turing Award (1975)



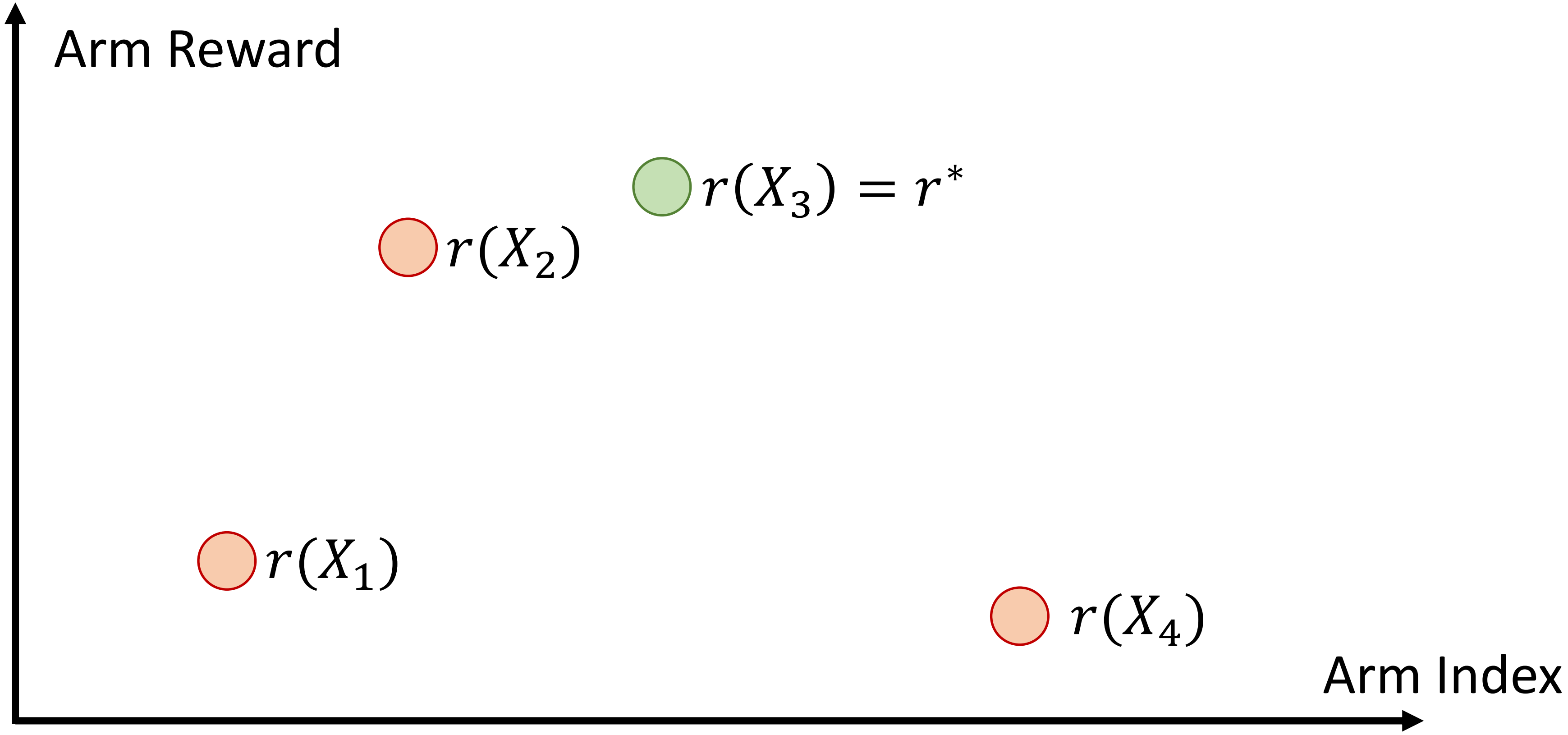
Satisficing is Everywhere!

Satisficing Bandit - Motivation



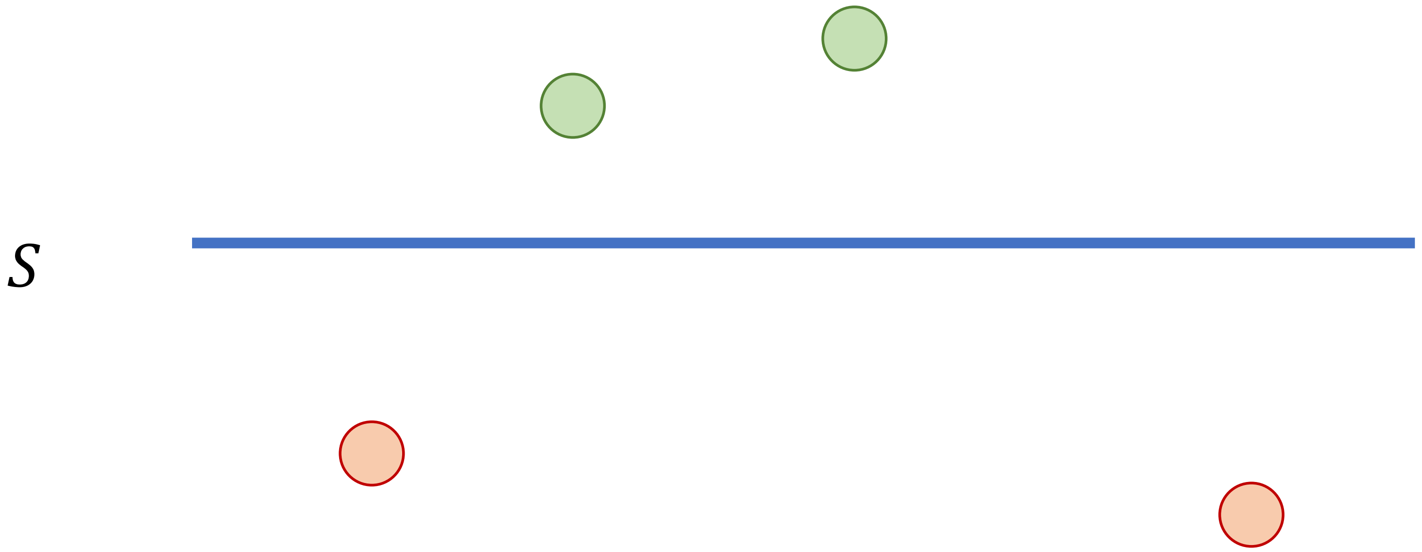
- A set of arms $\mathcal{X}$, T time steps. Arm X has unknown mean reward $r(X)$
- Select arm X_t in time step t brings a random reward has a mean $r(X_t) + \epsilon_t$

Satisficing Bandit - Motivation



$$\mathbf{Regret} = T \cdot r^* - \sum_{t=1}^{T} r(X_t)$$

Satisficing Bandit - Motivation

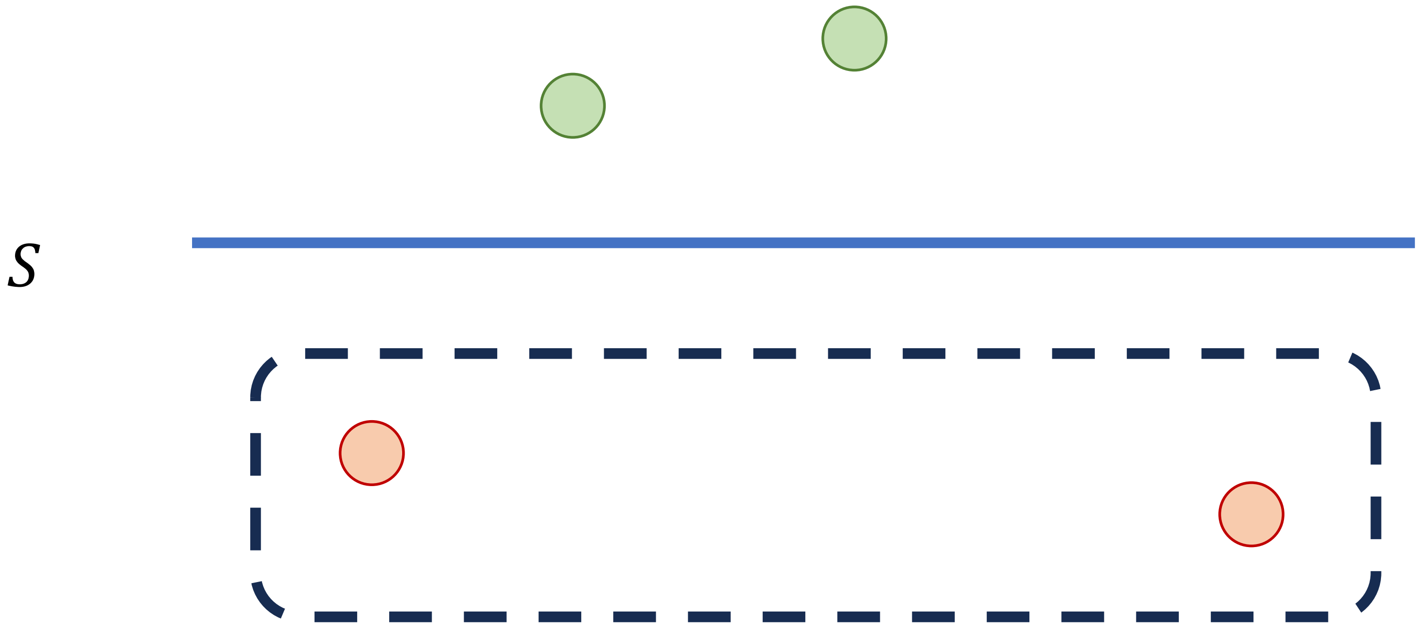


$$\mathbf{Regret_S} = \sum_{t=1}^{T} \max\{S - r(X_t), 0\}$$

$$\mathbf{Regret} = T \cdot r^* - \sum_{t=1}^{T} r(X_t)$$

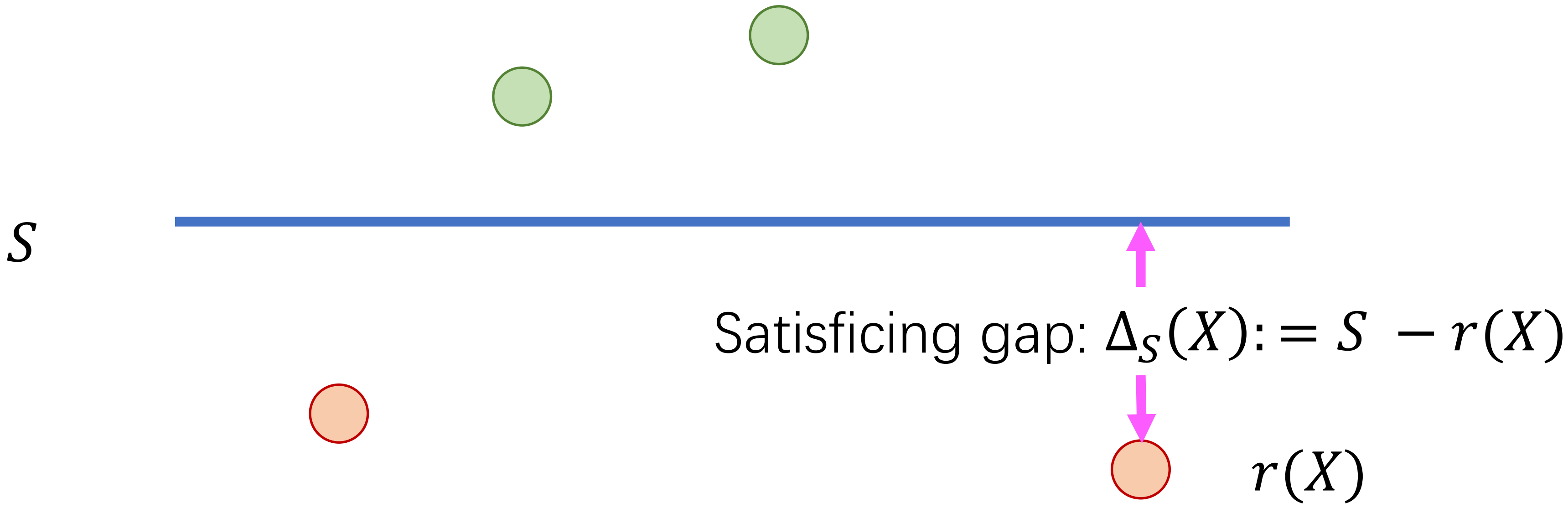
Satisficing Bandit - Motivation

- Remove all non-satisficing arms *(Garivier et al. 19, Michel et al. 23)* in finite-armed bandits



Satisficing Bandit - Motivation

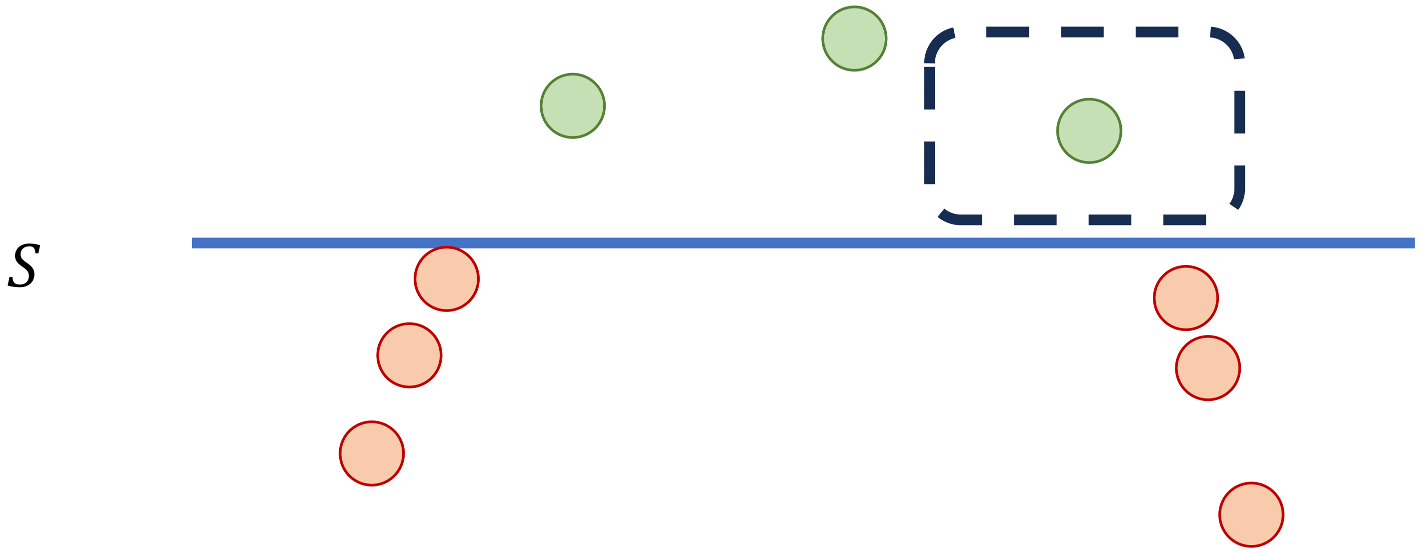
- Remove all non-satisficing arms *(Garivier et al. 19, Michel et al. 23)* in finite-armed bandits



Theorem. The satisficing regret of this algorithm is $O\left(\frac{1}{\min_X \Delta_S(X)}\right)$

Uniformly Bounded (Constant)! But generalizable?

Satisficing Bandit - Motivation



- Smallest satisficing gap approaches 0
- Almost impossible to remove all non-satisficing arms
- To achieve low satisficing regret, all we need is ONE satisficing arm!

Satisficing Bandit – Method & Results

- **Step 1:** Guess a candidate satisficing arm (by running oracle ALG)
- **Step 2:** Collect data sample from this arm
- **Step 3:** Test if the arm is truly satisficing (with LCB testing)

Theorem. If exits ALG with sub-linear standard regret, SELECT enjoys constant satisficing regret $\tilde{O}\left(\frac{1}{r^*-S}\right)$

Theorem. If no satisficing arm exists, SELECT shares the same standard regret as ALG

Satisficing Bandit - Insights

- **Settles for a good** choice is sometimes better than **hunting for the best**

- To reach satisficing objective, **find a good enough** choice instead of **delete all bad ones**

- Search for a good choice **as if searching for the best**