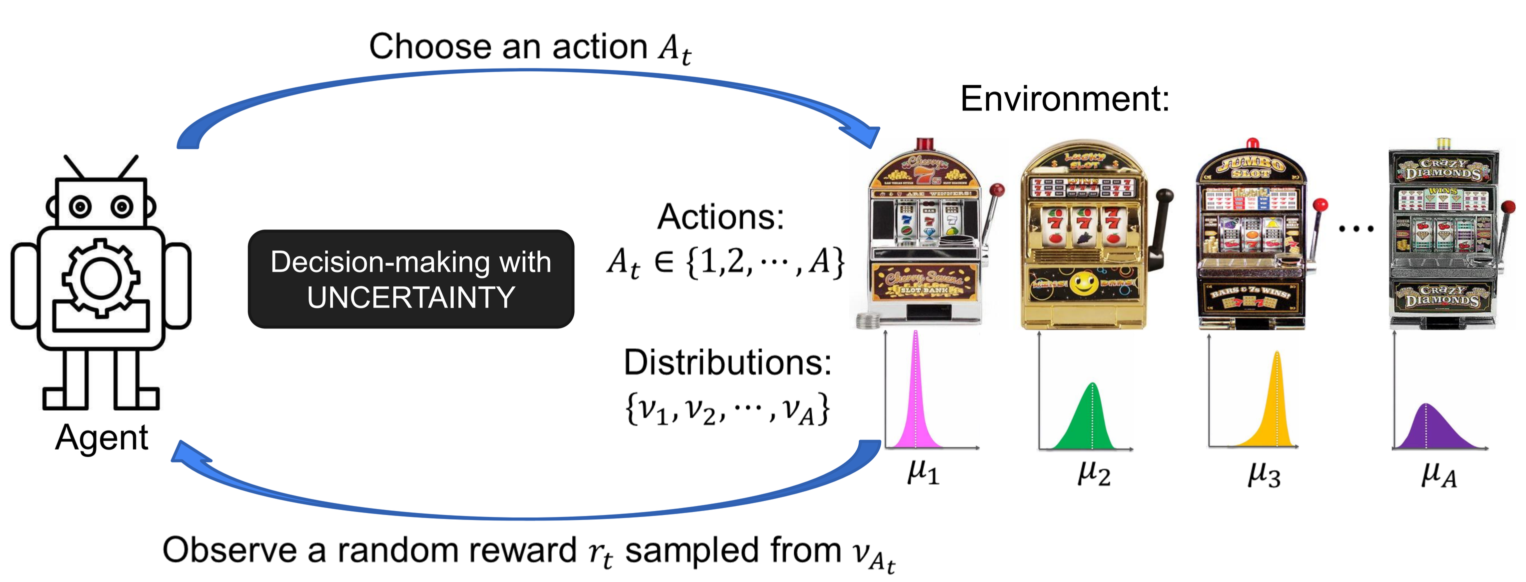
Cost Aware Best Arm Identification

Kellen Kanarios, Qining Zhang, Lei Ying
{kellenkk, qiningz, leiying}@umich.edu
MICHIGAN ENGINEERING
UNIVERSITY OF MICHIGAN
ECE ELECTRICAL & COMPUTER ENGINEERING UNIVERSITY OF MICHIGAN

https://arxiv.org/abs/2402.16710

Multi Armed Bandits

- Online sequential decision making.
- Many real world applications.



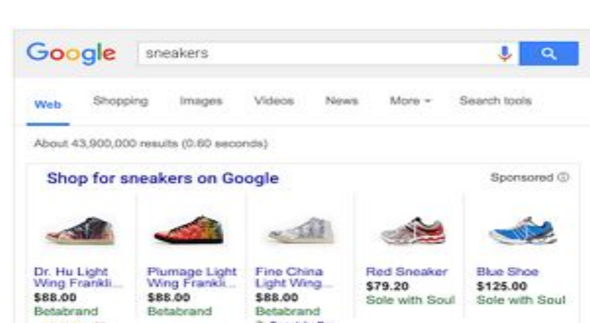
Learning / Implementation Phase

- Real world problems often have different objectives between learning and implementation

Clinic trials

- Learning: Minimize cost of trials
- Implementation: Cost is subsidized want most effective.

Recommendation



- Learning: Minimize cost to display ads
- Implementation: Click through rate.

Cost Aware BAI

- Best Arm Identification (Garivier, et al. 2016) ← Each arm has same cost

BAI

$$\min \tau \quad \text{s.t } \Pr(\hat{a}_\tau = a^*) \geq 1-\delta$$

CBAI

$$\min \sum_a \mathbb{E}\left[c_a N_a(\tau)\right] \quad \text{s.t } \Pr(\hat{a}_\tau = a^*) \geq 1-\delta$$

Cost Optimal Best Arm Identification

- Asymptotic optimality in incurred cost.
- Asymptotic lower bound on the cost for fixed confidence.

$$\mathbb{E}\left[\sum_a c_a N_a(\tau_\delta)\right] \geq T^*(\mu)\log(1/\delta) + o(\log(1/\delta))$$

$$T^*(\boldsymbol{\mu})^{-1} = \sup_{\boldsymbol{w}\in\Sigma_K} \inf_{\boldsymbol{\lambda}\in\{a^*(\boldsymbol{\lambda})\neq a^*(\boldsymbol{\mu})\}} \sum_a \frac{w_a}{c_a} d(\mu_a, \lambda_a).$$

Our Algorithm: CTAS

- Cost Aware Track and Stop
- Each arm makes up an optimal proportion of the final cost.
- Idea: track optimal proportion and stop as soon as possible.

Algorithm	Optimal?	$w_1(t)$ $(1.5, 1)$	$w_2(t)$ $(1, 0.1)$	$w_3(t)$ $(0.5, 0.01)$
TAS	×	0.46	0.46	0.08
CTAS	✓	0.23	0.72	0.05

- Optimal proportions for heterogeneous cost. (mean reward, mean cost) for each arm.

- Stopping is same as BAI: **ASAP**
- Use Generalized Likelihood Ratio.

Theoretical Results

Asymptotically optimal in expectation:

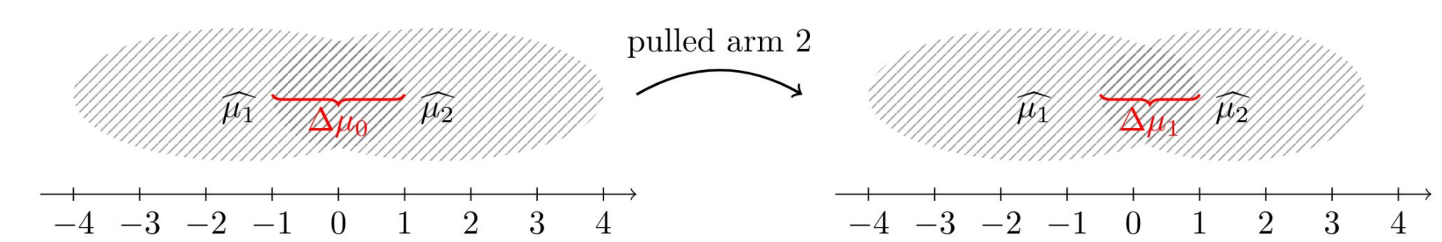
- Under CTAS (for suitable choices)

$$\limsup_{\delta\to 0} \frac{\mathbb{E}\left[\sum_a c_a N_a(\tau_\delta)\right]}{\log(1/\delta)} \leq T^*(\mu)$$

Low Complexity Algorithm

Idea: pull arm $a_t \in \arg\min_{a\in\mathcal{R}} \sqrt{c_a} N_a(t)$

- Asymptotically optimal in two armed Gaussian bandits.



	CO	CTAS	TAS	d-LUCB
Gaussian	85	1712	2410	82
Bernoulli	58	1995	2780	60
Poisson	96	3260	4633	101

Time (sec) of each algorithm over 1000 trajectories.

Numerical Results