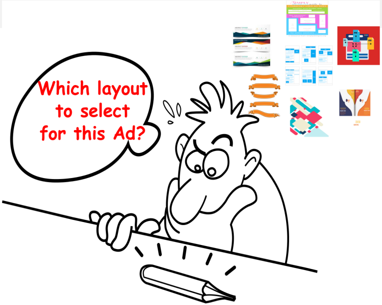
PAC Identification of Many Good Arms in Stochastic Multi-Armed Bandits

Arghya Roy Chaudhuri
under the guidance of
Prof. Shivaram Kalyanakrishnan

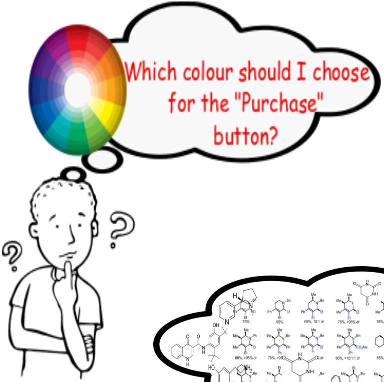
Indian Institute of Technology Bombay, India

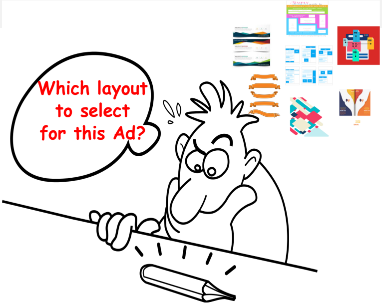
What Is It All About?

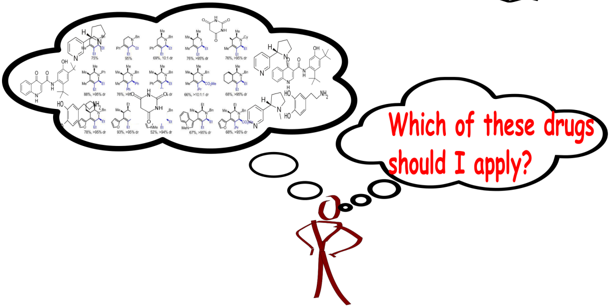
What Is It All About?



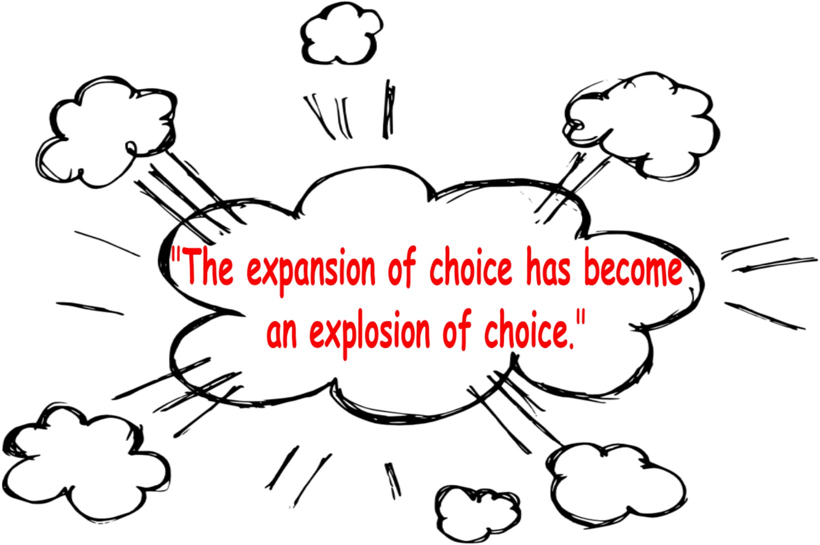
What Is It All About?



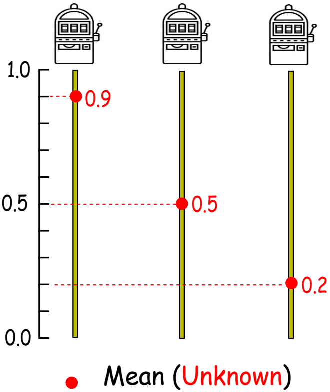




What Is It All About?



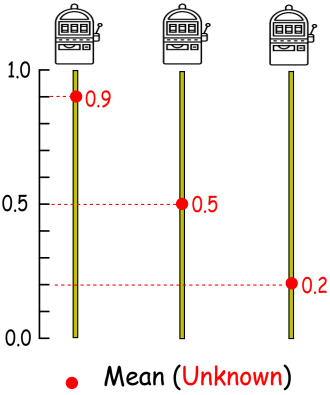
What Is a Multi-Armed Bandit?



Bandits: Slot machines

Mean: $\Pr[\text{Reward} = 1]$

What Is a Multi-Armed Bandit?



Bandits: Slot machines

Mean: $\Pr[\text{Reward} = 1]$

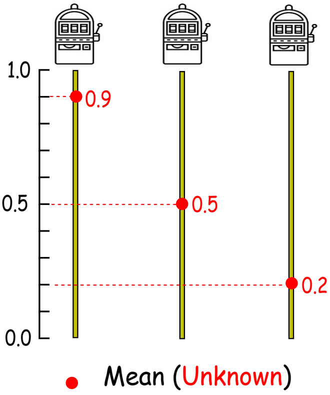
- To identify the best arm:

$$E[SC] = \Omega\left(\frac{\mathbf{n}}{\epsilon^2}\log\frac{1}{\delta}\right)$$

- To identify the best subset of size m:

$$E[SC] = \Omega\left(\frac{\mathbf{n}}{\epsilon^2}\log\frac{\mathbf{m}}{\delta}\right)$$

What Is a Multi-Armed Bandit?



Bandits: Slot machines

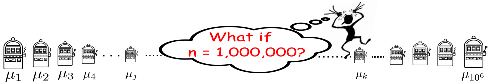
Mean: Pr[Reward = 1]

- To identify the best arm:

$$E[SC] = \Omega\left(\frac{\mathbf{n}}{\epsilon^2}\log\frac{1}{\delta}\right)$$

- To identify the best subset of size m:

$$E[SC] = \Omega\left(\frac{\mathbf{n}}{\epsilon^2}\log\frac{\mathbf{m}}{\delta}\right)$$



We need an alternative.

Large Bandit Instances

Difficulty for $n \gg T$:

$\lim_{n \to \infty} \frac{n}{\epsilon^2} \log \frac{1}{\delta} = \infty$.

Large Bandit Instances

Difficulty for $n \gg T$:

$\lim_{n\to\infty} \frac{n}{\epsilon^2} \log \frac{1}{\delta} = \infty$.

Get around:

- Identifying 1 from the best ρ-fraction is possible.

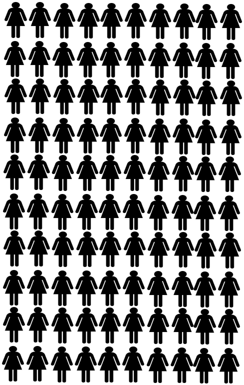
Large Bandit Instances

Difficulty for $n \gg T$:

$$\lim_{n\to\infty} \frac{n}{\epsilon^2} \log \frac{1}{\delta} = \infty.$$

Get around:

- Identifying 1 from the best ρ-fraction is possible.



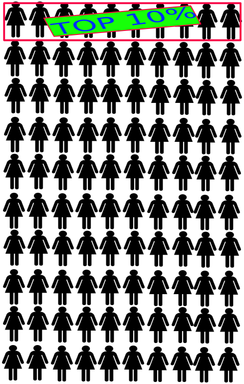
Large Bandit Instances

Difficulty for $n \gg T$:

$\lim_{n\to\infty} \frac{n}{\epsilon^2} \log \frac{1}{\delta} = \infty.$

Get around:

- Identifying 1 from the best ρ-fraction is possible.



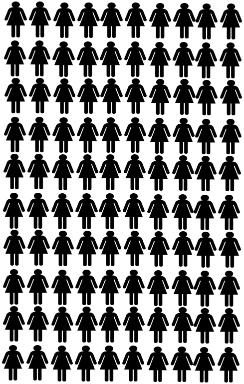
Large Bandit Instances

Difficulty for $n \gg T$:

$\lim_{n\to\infty} \frac{n}{\epsilon^2} \log \frac{1}{\delta} = \infty$.

Get around:

- Identifying 1 from the best ρ-fraction is possible.



Large Bandit Instances

Difficulty for $n \gg T$:

$\lim_{n \to \infty} \frac{n}{\epsilon^2} \log \frac{1}{\delta} = \infty.$

Get around:

- Identifying 1 from the best ρ-fraction is possible.

Large Bandit Instances

Difficulty for $n \gg T$:

$$\lim_{n \to \infty} \frac{n}{\epsilon^2} \log \frac{1}{\delta} = \infty.$$

Get around:

- Identifying 1 from the best ρ-fraction is possible.

Large Bandit Instances

Difficulty for $n \gg T$:

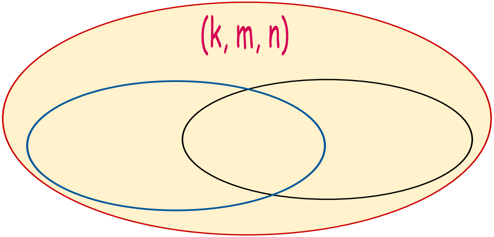
$$\lim_{n\to\infty} \frac{n}{\epsilon^2} \log \frac{1}{\delta} = \infty.$$

Get around:

- Identifying 1 from the best ρ-fraction is possible.
- Redefine the problem to identify 1 from the best m arms.
- Defining $\rho = \frac{m}{n}$, generalise the problem.
- What if we n is relatively small?

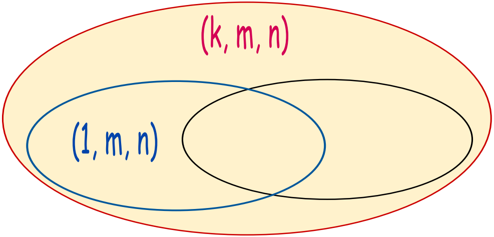
Finite-Armed Bandit Instances

$(\mathbf{k}, \mathbf{m}, \mathbf{n})$: To identify **any** distinct **k** arms from the **best m** arms in a set of **n** arms.



Finite-Armed Bandit Instances

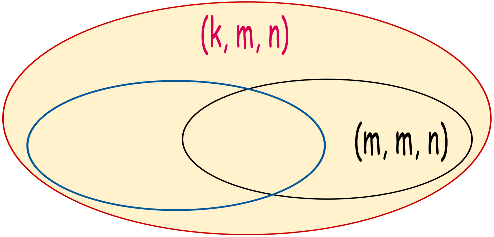
$(\mathbf{k}, \mathbf{m}, \mathbf{n})$: To identify **any** distinct **k** arms from the **best m** arms in a set of **n** arms.



- $\mathbf{k} = \mathbf{1}$: *Any* 1 arm out of the best *subset* of size m.

Finite-Armed Bandit Instances

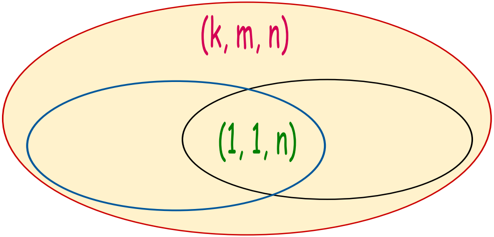
$(\mathbf{k}, \mathbf{m}, \mathbf{n})$: To identify **any** distinct **k** arms from the **best m** arms in a set of **n** arms.



- $\mathbf{k} = \mathbf{m}$: Best *subset* identification.

Finite-Armed Bandit Instances

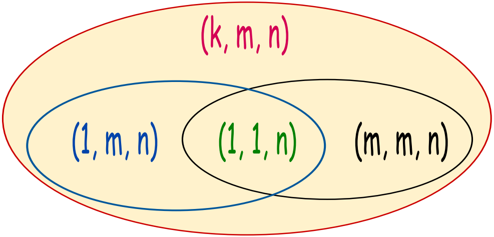
$(\mathbf{k}, \mathbf{m}, \mathbf{n})$: To identify **any** distinct **k** arms from the **best m** arms in a set of **n** arms.



- $\mathbf{k} = \mathbf{m} = \mathbf{1}$: Best arm identification.

Finite-Armed Bandit Instances

$(\mathbf{k}, \mathbf{m}, \mathbf{n})$: To identify **any** distinct **k** arms from the **best m** arms in a set of **n** arms.



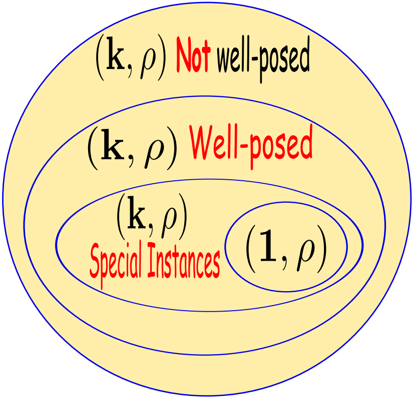
- $\mathbf{k} = \mathbf{1}$: *Any* 1 arm out of the best *subset* of size m.
- $\mathbf{k} = \mathbf{m}$: Best *subset* identification.
- $\mathbf{k} = \mathbf{m} = \mathbf{1}$: Best arm identification.

Contributions:

- LUCB-k-m (Fully sequential + Adaptive).
- Worst case upper and lower bound.

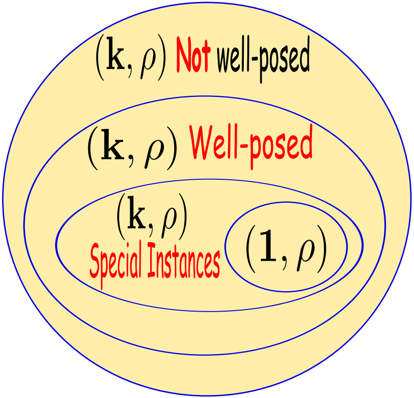
Infinite-Armed Bandit Instances

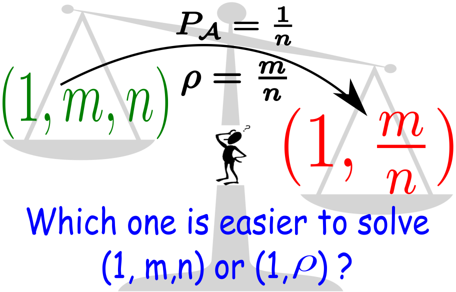
$(\mathbf{k}, \rho)$: To identify **any** distinct **k** arms from the **best** ρ fraction of arms.



Infinite-Armed Bandit Instances

$(\mathbf{k}, \rho)$: To identify **any** distinct **k** arms from the **best** ρ fraction of arms.





Thank You!

Poster: #54

Email: arghya@cse.iitb.ac.in