Input: n arms $\mathcal{X} = \{1, \ldots, n\}$, confidence level $\delta \in (0,1)$.
Let $\widehat{\mathcal{X}}_1 \leftarrow \mathcal{X}, \ell \leftarrow 1$
while $|\widehat{\mathcal{X}}_\ell| > 1$ **do**
  $\epsilon_\ell = 2^{-\ell}$
  Pull each arm in $\widehat{\mathcal{X}}_\ell$ exactly $\tau_\ell = \lceil 2\epsilon_\ell^{-2} \log(\frac{4\ell^2|\mathcal{X}|}{\delta})\rceil$ times
  Compute the empirical mean of these rewards $\widehat{\theta}_{i,\ell}$ for all $i \in \widehat{\mathcal{X}}_\ell$
  $\widehat{\mathcal{X}}_{\ell+1} \leftarrow \widehat{\mathcal{X}}_\ell \setminus \{i \in \widehat{\mathcal{X}}_\ell : \max_{j \in \widehat{\mathcal{X}}_\ell} \widehat{\theta}_{j,\ell} - \widehat{\theta}_{i,\ell} > 2\epsilon_\ell\}$
  $\ell \leftarrow \ell + 1$
Output: $\widehat{\mathcal{X}}_{\ell+1}$ (or play the last arm forever in the regret setting)

Elim $\sum_{i=2}^{n} \Delta_i^{-2} \log\left(\frac{n}{\delta}\right)$

lil'UCB $\sum_{i=2}^{n} \Delta_i^{-2} \log\left(\frac{1}{\delta}\right)$

$(a \vee b) := \max(a, b)$ WLOG: $\theta_1^* > \theta_2^* \geq \ldots \geq \theta_n^*$

Theorem Assume $\max_i \Delta_i \leq 4$. w.p. $\geq 1-\delta$

$$\sum_{i=2}^{n} \Delta_i T_i \leq \inf_{\nu \geq 0} \nu T + c \sum_{i=2}^{n} (\Delta_i \vee \nu)^{-1} \log\left(\frac{\log((\Delta \vee \nu)^{-1})|\mathcal{X}|}{\delta}\right).$$

Lemma Assume $\max_i \Delta_i \leq 4$. With prob $\geq 1-\delta$ we have $\{1\} \in \widehat{\mathcal{X}}_\ell \; \forall \ell$ and $\max_{i \in \widehat{\mathcal{X}}_\ell} \Delta_i \leq 8\varepsilon_\ell$.

Proof (of theorem): Fix any $\nu \geq 0$

$$\sum_{i=2}^{n} \Delta_i T_i \leq \sum_{i=2}^{n} T_i \max\{\nu, \Delta_i\}$$

$$\leq \sum_{i: \Delta_i \leq \nu} T_i \nu + \sum_{i: \Delta_i > \nu} T_i \Delta_i$$

$$\leq \nu T + \sum_{i: \Delta_i > \nu} T_i \Delta_i$$

$8\varepsilon_\ell \leq \nu$ when $\ell \geq \log_2\left(\frac{8}{\nu}\right)$

$$T_i = \sum_{\ell=1}^{\infty} \mathbb{1}\{i \in \widehat{\mathcal{X}}_\ell\} \tau_\ell \overset{\text{(by lemma)}}{=} \sum_{\ell=1}^{\log_2\left(\frac{8}{\Delta_i}\right)} \tau_\ell$$

$$\leq \nu T + \sum_{\ell=1}^{\log_2\left(\frac{8}{\nu}\right)} \sum_{i:\Delta_i>\nu}^{n} 8\varepsilon_\ell T_\ell \mathbb{1}\{\Delta_i \leq 8\varepsilon_\ell\}$$

$$\leq \nu T + \sum_{i:\Delta_i>\nu}^{n} \sum_{\ell=1}^{\log_2\left(\frac{8}{\Delta_i \vee \nu}\right)} 8\varepsilon_\ell T_\ell$$

$$\leq \nu T + \sum_{i=2}^{n} \sum_{\ell=1}^{\log_2\left(\frac{8}{\Delta_i \vee \nu}\right)} 8 \cdot \varepsilon_\ell \cdot 2\varepsilon_\ell^{-2} \log\left(\frac{4\ell^2 n}{\delta}\right)$$

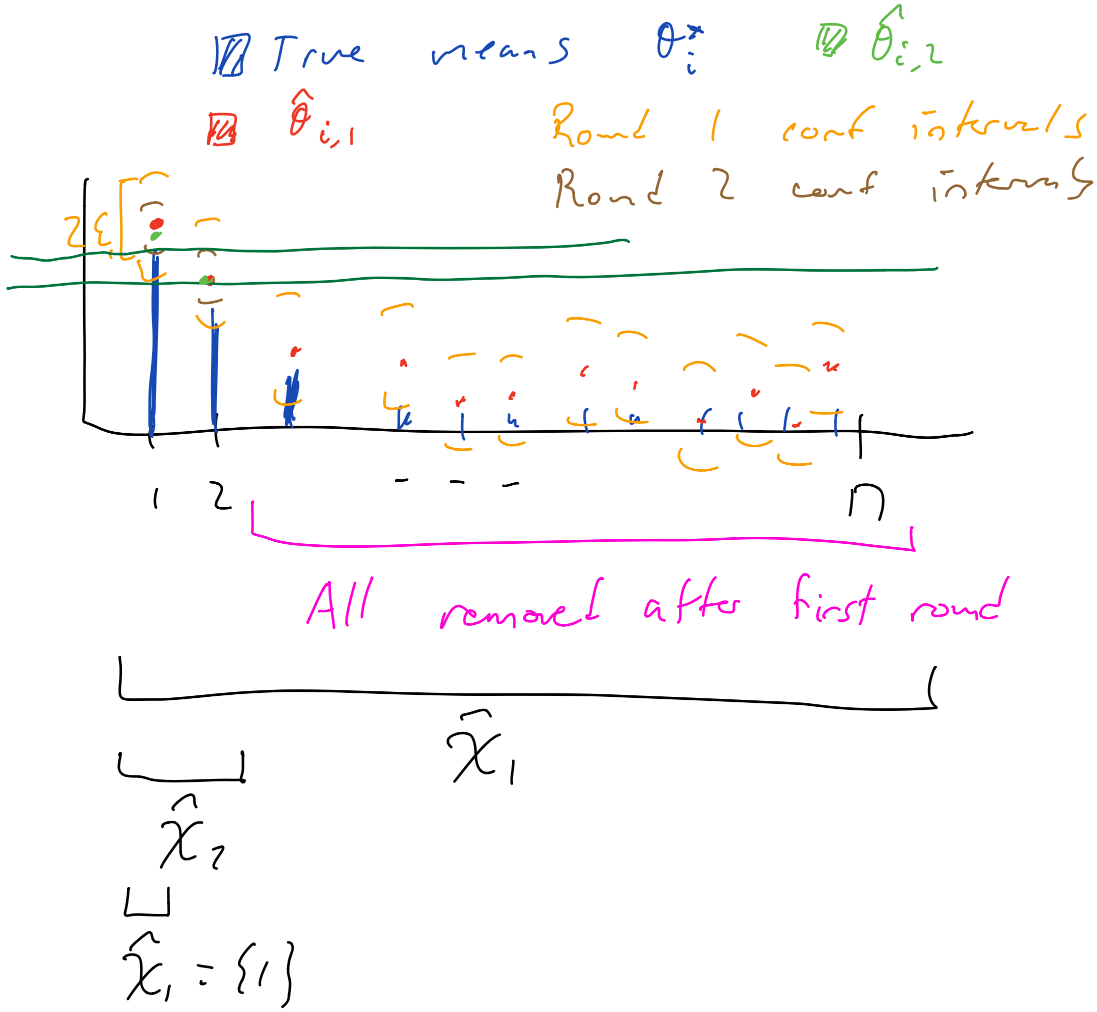
$$= \nu T + \sum_{i=2}^{n} \sum_{\ell=1}^{\log_2\left(\frac{8}{\Delta_i \vee \nu}\right)} 16 \cdot 2^{\ell} \cdot \log\left(\frac{4\ell^2 n}{\delta}\right)$$

$$\leq \nu T + \sum_{i=2}^{n} 16 \log\left(\frac{4\log^2\left(\frac{8}{\Delta_i \vee \nu}\right) n}{\delta}\right) \underline{\sum_{\ell=1}^{\log_2\left(\frac{8}{\Delta_i \vee \nu}\right)}} \underline{2^{\ell}}$$

$$\leq \nu T + \sum_{i=2}^{n} (16 \cdot 2 \cdot 8) \frac{1}{\Delta_i \vee \nu} \log\left(\frac{4\log^2\left(\frac{8}{\Delta_i \vee \nu}\right) n}{\delta}\right)$$

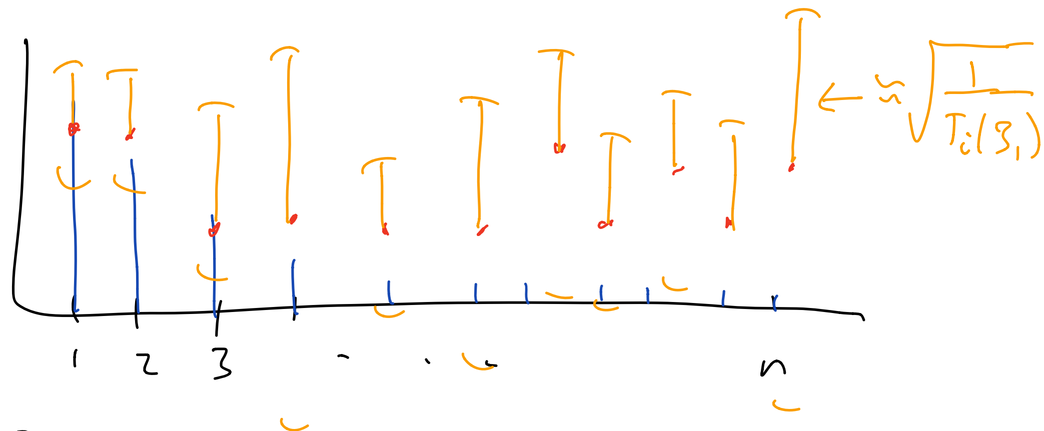
$$\leq \nu T + c \sum_{i=2}^{n} \frac{1}{\Delta_i \vee \gamma} \log\left(\frac{\log\left(\frac{1}{\Delta_i \vee \gamma}\right) n}{\delta} \right)$$

Elimination Algorithm picture.



Optimism in the Face of uncertainty

- True means
- Empirical means after 3, total pulls of all arms (not same # of times)
- conf intervals



UCB Algorithm

Pull each arm once

for $t = 1, 2, \ldots, T$

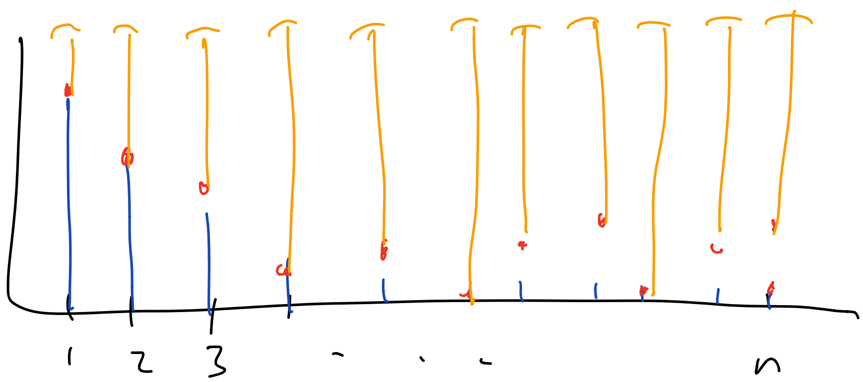
$$\text{Pull arm } \arg\max_i \overbrace{\underbrace{\hat{\theta}_{i,T_i(t-1)}}_{\substack{\text{empirical}\\ \text{mean after}\\ T_i(t-1) \text{ pulls}}} + \underbrace{c\sqrt{\frac{\log(\ )}{T_i(t-1)}}}_{\substack{\text{Confidenc}\\ \text{interval after}\\ T_i(t-1) \text{ pulls}}}}^{=: UCB_i}$$

Idea: With high prob, $\theta_i^* \le UCB_i \quad \forall i, t$

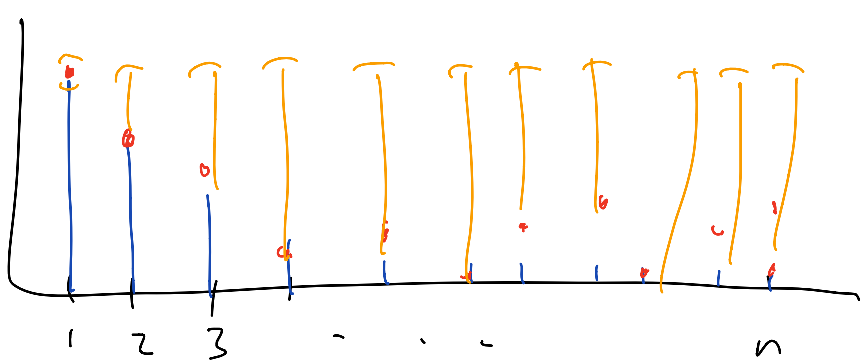
If pull $\arg\max_i UCB_i(t)$ then either

- θ_i^* is large (and get small regret)
- $UCB_i(t)$ is too large and shrinks after pulling it.

After many rounds...



Many more rounds $\theta_1^* \le UCB_1$



$\theta_i^* < \theta_1^*$

$i > 1$

$i \neq 1$ $\theta_i^* \le UCB_i \to \le \theta_1^*$

How many times is arm $i \neq 1$ played?

As many times so that $UCB_i \leq \theta_1^*$

$$UCB_i = \hat{\theta}_i + c\sqrt{\frac{1}{T_i}}$$

$$\leq \theta_i^* + 2c\sqrt{\frac{1}{T_i}}$$

$$\leq \theta_1^* \; ?$$

Happens when $T_i \geq \dfrac{4c^2}{(\theta_1^* - \theta_i^*)^2}$.