today improved learning (swap regret)

All strategies s, s'

$$\sum_{t=1}^{T} u^t(s^t) \geq \sum_{t: s^t \neq s} u^t(s^t) + \sum_{t: s^t = s} u^t(s') - Reg$$

All maps $\pi: S \to S$

$$\boxed{\sum_{t=1}^{T} u^t(s^t) \geq \sum_{t=1}^{T} u^t(\pi(s^t)) - Reg}$$

if $\pi(s) = s'$ all s, that regular Regret

Given any external no-regret alg
convert to swapregret

Each strategy s — no-regret alg A_s
"A_s in charge of maybe swapping outs"

Run each A_s
gives at time $q_s^t = (q_{s,1}^t \dots q_{sk}^t)$

Solve for $p^t = Q^t p^t$ | $p_s^t = \sum_i q_{si}^t p_i^t$

Q = rows

Note: solution always exists!

② selecting by p

or selecting by p alg A_s

& then using q_s to select

Feed alg A_s utilities

$p_s^t \, u^t(s')$ all s'

---

no-regret for A_s:

expected utility at time t

$$p_s^t \sum_i q_{s_i}^t \, u^t(s_i)$$

No-regret

$$\sum_t p_s^t \sum_i q_{s_i}^t \, u^t(s_i) \geq \sum_t p_s^t \, u^t(s') - \text{Reg}$$

use for $s' = \pi(s)$ & sum over all A_s

$$\sum_s \sum_t p_s^t \sum_i q_{s_i}^t \, u^t(s_i) \geq \sum_s \sum_t p_s^t \, u^t(\pi(s)) - k \cdot \overbrace{\text{Reg}}$$

$\parallel$

$$\sum_t \sum_i u^t(s_i) \sum_s p_s^t \cdot q_{s_i}^t = \sum_t \sum_i u^t(s_i)\, p_i^t$$

$\qquad$ ‖

utility of player

$$\sum_t \sum_s p_s^t u^t(\pi(s)) - k \text{ Regret}$$

‖

expected utility if you ~~you~~ swap $s \to \pi(s)$ each time

---

Regret we get

external ret was $O(\sqrt{T \ln k})$

this gives $O(k\sqrt{T \ln k})$

better version

$(1-\varepsilon)$ alternate $- \frac{\ln k}{\varepsilon}$

~~b~~ now we get $(1-\varepsilon) \ldots - k\frac{\ln k}{\varepsilon}$

choose ε now: gives $O(\sqrt{Tk \ln k})$

What if not full info?

take p as before $\longrightarrow$ s prob p_s

$$\bar{u}^t(s) = \begin{cases} 0 & \text{if } s \neq s^t \\ \dfrac{u^t(s)}{p^t_s} & \text{if } s = s^t \end{cases}$$

Run each A_s as before with utilities $\tilde{u}^t(s)$