SWAP and Internal regret

Modification rule: $F: X \to X$

Agent A selects p^t in $\forall$ step $t = 1 \dots T$

Agent A modified by F selects $f^t = (f_1^t, \dots, f_N^t)$

on X in $\forall$ step t where for $\forall j \in X$: $f_j^t = \sum_{\substack{i \in X \\ F(i)=j}} p_i^t$

Cummulative loss of this modified agent

is $L_{A,F}^T = \sum_{t=1}^{T} \sum_{i \in X} f_i^t \, l_i^t$

Let $\mathcal{F}$ be a class of modification rules

Regret of A with respect of $\mathcal{F}$ is $R_{A,\mathcal{F}}^T = \max_{F \in \mathcal{F}} \left(L_A^T - L_{A,F}^T \right)$

For class $\mathcal{F}^{EX} = \{ F_i : i \in X \}$, $F_i(j) = i$ for $\forall j \in X$

then $R_{A,\mathcal{F}^{EX}}^T$ is external regret.

For class $\mathcal{F}^{IN} = \{ F_{i,j} : i,j \in X, i \neq j \}$ where $F_{i,j}(k) = \begin{cases} j : k = i \\ k : k \neq i \end{cases}$ – almost identical except

then $R_{A,\mathcal{F}^{IN}}^T$ is internal regret. $\Rightarrow$ Comparing copies of agent, where they play almost the same, except the action j, which will be different.

For class $\mathcal{F}^{SW} = \{ F : X \to X \}$ we obtain

$R_{A,\mathcal{F}^{SW}}^T$ = SWAP Regrets

internal regret $\leq$ Swap regret

external regret $\leq$ Swap regret

$$R_{A,\mathcal{F}^{EX}}^T = \max_{j \in X} \sum_{t=1}^{T} \left(\sum_{i \in X} p_i^t l_i^t - l_j^t \right)$$

$\Rightarrow$ my cummulative loss $\Rightarrow$ including the other player's move.

$$R_{A,\mathcal{F}^{IN}}^T = \max_{\substack{i,j \in X \\ i \neq j}} \sum_{t=1}^{T} p_i^t \left(l_i^t - l_j^t \right)$$

$$R_{A,\mathcal{F}^{SW}}^T = \sum_{i=1}^{N} \max_{j \in X} \sum_{t=1}^{T} p_i^t \left(l_i^t - l_j \right)$$

Theorem 2.55.

For $\forall$ algorithm A with external regret $\leq R$

$\exists$ algorithm M with swap regret $\leq NR$

$(\forall A \exists M \, \forall F: X \to X: \, L_A^T \leq L_{A,F}^T + NR)$

Pf:

$A_1 \dots A_n$ be copies of A

$\forall t$: Let $q_i^t = (q_{i,1}^t, \dots, q_{i,N}^t)$ be prob. distr. created by A_i at step t

we choose $p^t = (p_1^t, \dots, p_N^t)$, the prob. distr. selected by M at step t,

by setting for $\forall j \in X$, $\underline{p_j^t = \sum_{i=1}^{N} p_i^t \cdot q_{i,j}^t}$ → solution exists, can be found effectively

$(p^t)^T = (p^t)^T Q^t$ where $(Q^t)_{ij} = q_{ij}^t$

After M selects p^t, it recieves $\ell^t = (\ell_1^t, \dots, \ell_N^t)$, then A_i recieves losses $p_i^t \cdot \ell_i^t$.

— A_i experiences expected loss $\underbrace{q_i^t}_{\text{scalar}} \cdot \underbrace{(p_i^t \, \ell^t)}_{\text{vector}} = p_i^t \cdot (q_i^t \cdot \ell^t)$

External regret of A_i is $\leq R \implies \forall j \in X: \sum_{t=1}^{T} p_i^t (q_i^t \, \ell^t) \leq \underbrace{\sum_{t=1}^{T} p_i^t \, \ell_j^t}_{\text{cummulative loss of } j} + R$ (*)

Losses of all A_is at step t $= \sum_{i=1}^{N} p_i^t (q_i^t \, \ell^t) = (p^t)^T \cdot Q^t \cdot \ell^t$

$= (p^t)^T \cdot \ell^t$ = loss of M at t. (**)

By summing (*) over $\forall i \in X$ and by using (**) for $F: X \to X$, $j = F(i)$

$$\underbrace{L_M^T}_{\text{extension of (*) } \forall i \in X} \leq \underbrace{L_{M,F}^T}_{\sum_{i=1}^{N} \sum_{t=1}^{T} p_i^t \, \ell_{F(i)}^t} + N \cdot R \qquad \boxed{\times}$$

If A = poly. w. alg. $\implies \exists$ algo. M with SWAP regret $\leq O(N\sqrt{T \cdot \log(N)})$

— M has average swap regret $\leq \varepsilon$.

Theorem:

$\forall G = (P, A, C)$ of n players, $\forall \varepsilon > 0$, $\forall T = T(\varepsilon)$, run no-swap-regret dynamics

$p^t = \prod_{i=1}^{n} p_i^t$, $p = \frac{1}{T} \sum_{t=1}^{T} p^t$, then p is ε-CE

p is CE $\left(\mathbb{E}_{a \sim p}[C_i(a) / a_i] \leq \mathbb{E}[C_i(a_i', a_{-i}) / a_i] \right) \iff$

$\iff \forall i \in P, \forall F: X \to X: \mathbb{E}_{a \sim p}[C_i(a)] \leq \mathbb{E}_{a \sim p}[C_i(F(a_i); a_{-i})]$

$p_i^t \, \ell_i^t$

Games in extensive form

Described by tree:

root : initial position

leaf : final state + payoff

internal nodes : decision nodes

move: outgoing edge from decision node.

Perfect information games: - we know exactly where in the tree we are.

Imperfect information games: decision nodes are partitioned into information sets

- $\forall$ nodes in it belong to the same player
- $\forall$ nodes in it have the same set of moves

H_i := set of information sets of player i

for $h \in H_i$, let C_h = { moves we can at node from h }

- we can simulate normal-form games with imperfect information game.

- in russian roulette, we introduce player 0, (instead fo just {1,2}), which represents the randomness of the gun being loaded.