Be The Leader

Alg 0: Be the Leader (BTL)

- let $V_j^i = \sum_{r=1}^{i} v_j^r$
- in round i choose:

$$j^i = \arg\max_{j'} V_{j'}^{i} \leftarrow$$

(Recall: FTL $j^i = \arg\max_{j'} V_{j'}^{i-1}$)

Example $k=2$ actions

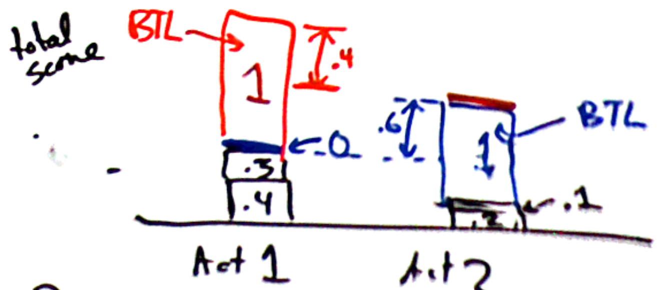
	1	2	3	4	
Action 1	$\boxed{.4}$	$\boxed{.3}$	0	$\boxed{1}$	⟸ OPT (best action in hindsight)
Action 2	.2	.1	$\boxed{1}$	0	
BTL Total	.4	.7	1.7	2.7	
OPT total	.4	.7	.6 1.3	.6 1.7	

Thm: $BTL \geq OPT$.

"change in BTL in each round"

$\geq$

"change in OPT in each round"



Proof:

- let OPT_i = best in hindsight after i rounds
- let $opt_i = OPT_i - OPT_{i-1}$
- let btl_i = payoff of BTL in round i
- claim: $btl_i \geq opt_i$
  - btl_i = payoff of best in hindsight action
    $\geq$ increase in best in hindsight ...

btl_i = payoff of best in hindsight action

$\geq$ increase in total payoff of best in hindsight.

$= opt_i$

- $BTL = \sum_i btl_i \geq \sum opt_i = OPT$.

Follow The Perturbed Leader

Alg 2: Follow The perturbed leader (FTPL)

- learning rate ϵ
- hallucinate: $\begin{cases} \text{heads w.p. } \epsilon \\ \text{tails w.p. } 1-\epsilon \end{cases}$

  $v_j^0 = h \times$ "number of tails of ϵ-bias coin that flip in a row"
- let $V_j^i = \sum_{r=1}^{i} v_j^r$
- in round i: choose action

  $j^i = \text{argmax}_j (v_j^0 + V_j^{i-1})$

Example:

	0	1	2	3	4	5
Action 1	2	$\frac{1}{2}$	0	1	0	1
Action 2	3	0	1	0	1	0

2.

OPT : 2.5 $\to n/2$

FTPL : 2 $\to n/2$

FTL : $1/2$ $\to 1/2$

Thm: payoffs in $[0,h]$

$$E[FTPL] \geq (1-\epsilon)OPT - O(\tfrac{h}{\epsilon}\log k)$$

Cor: in n rounds with payoffs in $[0,h]$

$$E(\text{Regret}_n(FTPL)) \leq 2h\sqrt{\frac{\ln k}{n}}.$$

Proof of Cor: same as for EW.

Q: why does FTPL work?

A:

1) stability: FTPL $\approx$ BTPL
2) small perturbation: BTPL $\gtrsim$ OPT

Def: Be the Perturbed Leader (BTPL)

"$j^i = \arg\max_j (v_j^0 + V_j^i)$"

Lemma 1 (stability): FTPL $\geq (1-\epsilon)$ BTPL

Lemma 2 (Small Perturb):

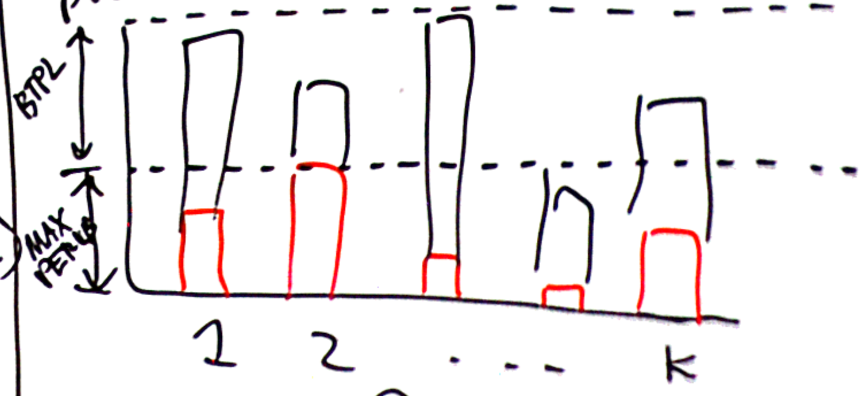
BTPL $\geq$ OPT $- O(\frac{h}{\epsilon} \log k)$

Proof of Thm:

combine lemmas.

Proof of Lemma 2:

picture



Thm $\Rightarrow \underbrace{\widetilde{BTPL}}_{=\text{BTPL} + \text{MAXPERTURB}} \geq \widetilde{OPT} \geq OPT$

$\Rightarrow$ BTPL $\geq$ OPT $-$ MAXPERTU

claim: E[MAX Perturb] $= O(\frac{h}{\epsilon} \log k)$

MAXPERTURB = # rounds until no actions left.

e.g. $\epsilon = 1/2 \Rightarrow \log_2 k$ ...

- $v_j^0 = h \times$ "number of ϵ bias coins that flip tails in a row"

$= h \times$ Geometric Random Variable with $p = 1-\epsilon$.

$$\text{MAXPERTURB} = \max_j v_j^0 .$$

perturbation round	Actions 1	2	...	k
1	T	$\cancel{H}$	T	T $(1-\epsilon)k$
2	$\cancel{H}$		T	$\cancel{H}$
⋮			T	
			$\cancel{H}$	

"$\log_{\frac{1}{1-\epsilon}} k$"

"$\frac{1}{\ln\frac{1}{1-\epsilon}} \ln k$"

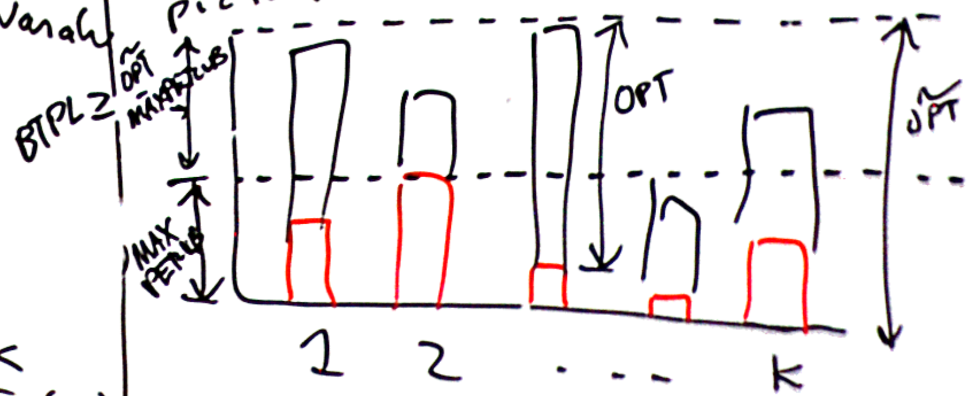
$\ln \frac{1}{1-\epsilon}$

$= -\ln(1-\epsilon)$

$\approx \epsilon$

Proof of Lemma 2:

picture



Thm $\Rightarrow \widetilde{BTPL} \geq \widetilde{OPT} \geq OPT$

$\widetilde{BTPL}$ = BTPL + MAXPERTURB

$\Rightarrow$ BTPL $\geq$ OPT − MAXPERTURB

claim: E[MAX PERTURB] $= O(\frac{1}{\epsilon} \log k)$

MAXPERTURB = # rounds until no actions left

e.g. $\epsilon = 1/2 \Rightarrow \log_2 k$ ...