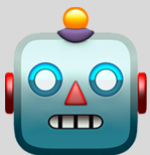


## Problem setting

For  $t = 1, \dots, T$ :



**Learner**

Make prediction  $\hat{c}_t \in \mathbb{B}^n$  of  $c^*$ .

Observe  $(X_t, x_t)$  and update from  $\hat{c}_t$  to  $\hat{c}_{t+1}$ .

Take action  $x_t \in \operatorname{argmax}_{x \in X_t} \langle c^*, x \rangle$ .



**Agent**

## Regret

**Opt. val.**

**Val. of learner's action**

$$R_T^{c^*} := \sum_{t=1}^T \langle c^*, x_t - \hat{x}_t \rangle = \sum_{t=1}^T \langle c^*, x_t \rangle - \sum_{t=1}^T \langle c^*, \hat{x}_t \rangle \text{ for } \hat{x}_t \in \arg \max \{ \langle \hat{c}_t, x \rangle \mid x \in X_t \}.$$

## Our results

- ONS-based method with  $R_T^{c^*} = O(n \log T)$ ; the per-bound complexity is independent of  $T$ .
  - The first method to have these two desirable properties!
- MetaGrad-based method to handle suboptimality of  $x_t$ .
- Lower bound of  $R_T^{c^*} = \Omega(n)$ .