Incentive Design for Adaptive Agents

Yiling Chen
Jerry Kung
David Parkes
Ariel Procaccia
Haoqi Zhang

AAMAS 2011

We are adaptive agents

- ☹ “This restaurant doesn’t live up to the hype.”
- ☹ “This napkins brand is good, but not *that* good.”
- ☺ “Boston weather isn’t so bad after all.”
- ☺ “This cereal is awesome!”

Influencing an **adaptive agent**

Influencing an adaptive agent with **rewards**

1-on-1
time

How can a principal use incentives to induce an **adaptive agent** to select a particular target action?

How can a principal use incentives to induce an **adaptive agent** to select a particular target action?

- Agent’s values for actions update with experience
- Principal observes actions, but does not know the agent’s values nor update process

Our work

Our work

- A simple model

Our work

- A simple model
- Optimal **incentive policies**
  - online vs. offline

Our work

- A simple model
- Optimal **incentive policies**
  - – online vs. offline
    - per-period budget (possible)
    - across-period budget (impossible)

Our work

- A simple model
- Optimal **incentive policies**
  - online vs. offline
    - per-period budget (possible)
    - across-period budget (impossible)
      - » know everything
      - » know something

related work on influencing agents

	adaptive or learning agents	intervention method
This paper	yes	incentives
Policy teaching [Z. et al.]	no	incentives
Ad-hoc teams [Stone & Kraus]	yes	actions
Partially-Controlled MAS [Brafman & Tennenholtz]	yes	actions

Model

Model

- Set of actions, one is the target.

Model

- Set of actions, one is the target.
- Discrete time $t = 1, 2, 3, \ldots$

Model

- Set of actions, one is the target.
- Discrete time $t = 1, 2, 3, \ldots$
- State of an action updates independently of other actions if selected, otherwise state is fixed.

Model

- Set of actions, one is the target.
- Discrete time $t = 1, 2, 3, \ldots$
- State of an action updates independently of other actions if selected, otherwise state is fixed.
- Agent's belief about an action's value is based only on its state.

Model

- Set of actions, one is the target.
- Discrete time $t = 1, 2, 3, \ldots$
- State of an action updates independently of other actions if selected, otherwise state is fixed.
- Agent's belief about an action's value is based only on its state.
- The principal can provide an external reward for choosing a particular action.

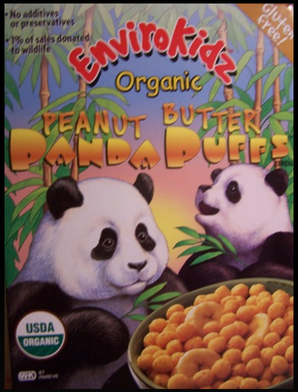
Model

- Set of actions, one is the target.
- Discrete time *t = 1, 2, 3, …*
- State of an action updates independently of other actions if selected, otherwise state is fixed.
- Agent's belief about an action's value is based only on its state.
- The principal can provide an external reward for choosing a particular action.
- Agent takes action with highest sum of value and external reward.

Example

4

5



2.5

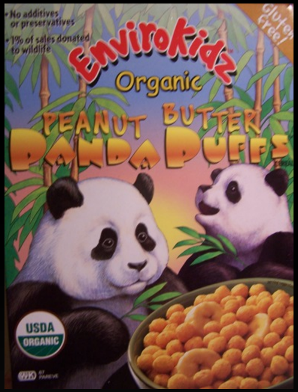
Example

4

5

2.5

Example



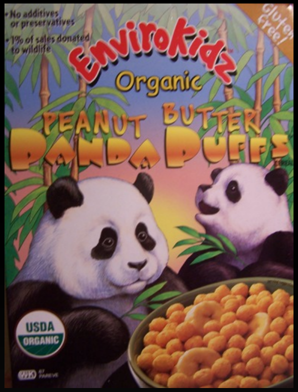
4

5

2.5

2

Example



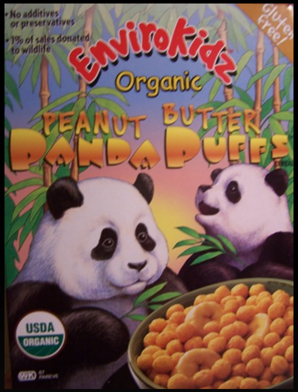
4

5

2.5

2

Example



4	5	2.5
6	2	

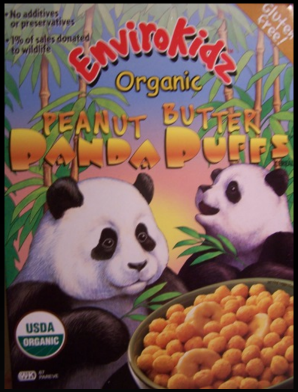
Example

4	5	2.5
6	2	

Example with incentives

4

5



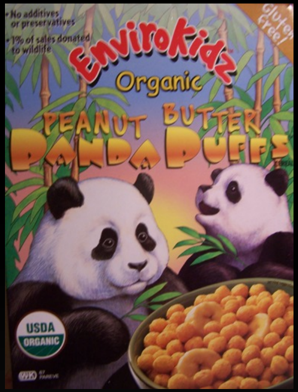
2.5

Example with incentives

Get a free plush toy!

4

5



2.5 + 2

Example with incentives

Get a free
plush toy!

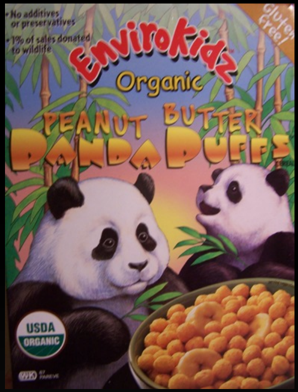
4

5

2.5 + 2

Example with incentives

Get a free plush toy!



4

5

2.5 + 2

2

Example with incentives

Get a free plush toy!

4

5

2.5 + 2

2

Example with incentives

Get a free
plush toy!

4

5

2.5 + 2

2

10

- Beliefs about value of an action can encode:
  - Empirical average of realized rewards
  - Explore/Exploit tradeoff (avg. + variance)
  - Bayesian learning

- Assumption: value updates based on experience ONLY, does not factor in incentives

- Assumption: decisions are myopic with respect to the principal's interventions

Per-period budget

- Can provide up to budget at each time step

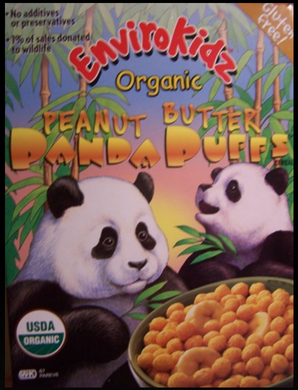
- Candidate policy: always provide to target

- Is this the best policy? May it benefit to intervene on other actions?

Theorem

Providing the budget to the target induces the target as soon as possible, and as many times as possible within a fixed number of time steps.

Threshold Lemma



4

...

5

...

2+1

Threshold

Implications

- Optimal incentive policy does not depend on the agent's values or update process
- The principal cannot otherwise speed up the agent's exploration of currently better actions

Across-Period budget

- Fixed budget to spend across time frame
- To induce target once, reduces to per-period budget case
- To induce target multiple times, we need to think about how to split the budget

Theorem

There is no (randomized) algorithm that provides a bounded competitive ratio for Induce-Multi, even if the algorithm can see the current values of the actions.

Implications

- Strong negative theoretical result, but in practice inputs may not be adversarial

- To make progress, important to consider **empirical or average case performance** for particular agent models and value distributions

Knowledge helps

- If know agent's values in any state, can compute optimal incentives in polynomial time
- In practice, more likely to only have distributional information on values

Case study: induce a new action

- Two actions
  - Incumbent action has fixed value of $1
  - New action's value is drawn Uniform[0,1]
- Belief updates by empirical average

Case study: induce a new action

- Two actions
  - Incumbent action has fixed value of $1
  - New action's value is drawn Uniform[0,1]
- Belief updates by empirical average

Question: What is the optimal incentive policy when there is $1 to spend across two rounds?

Case study: induce a new action

- Two actions
  - Incumbent action has fixed value of $1
  - New action's value is drawn Uniform[0,1]
- Belief updates by empirical average

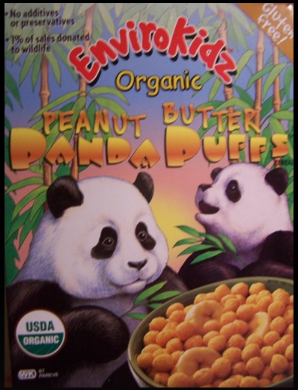
Question: What is the optimal incentive policy when there is $1 to spend across two rounds?

Answer: 4/9 in 1^{st} round, rest in 2^{nd}

Conclusion

- Incentive design for adaptive agents explores connections among incentives, actions, and learning
- Strong possibility and impossibility results
- Case study on using partial knowledge
- Rich space of computational and analytical problems

thank you



For more info: **poster R62**

For comments: **hq@eecs.harvard.edu**

For plush toys: **econcs.seas.harvard.edu**