Deep Bayesian Trust : A Dominant and Fair Incentive Mechanism for Crowd

Motivation

Quality Control Problem in Crowdsourcing

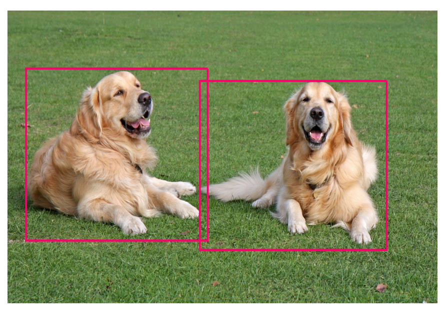
- Solving the tasks requires effort. Workers with different abilities and motivations submit answers.
- Performance based incentives are necessary to elicit effort and collect high quality answers.
- Answers can't be verified, which makes it very challenging to design incentive mechanisms.

- An important research problem because:
  - Crowdsourced data is used for training prediction algorithms and various other applications.
  - The prediction algorithms and the applications are only as good as the quality of the data.

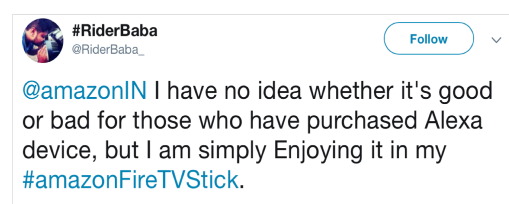
Examples of Crowdsourcing Tasks

Pollution Sensing

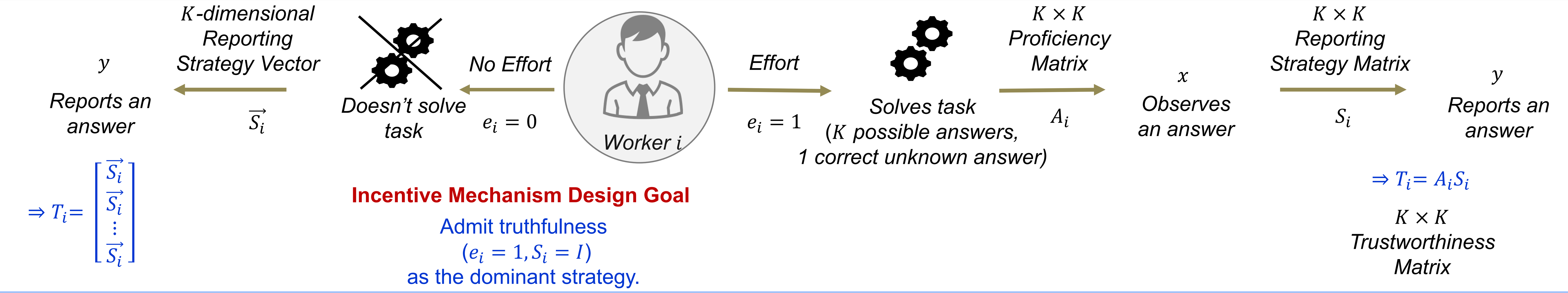
Image Labeling



Sentiment Identification



The Setting



Incentive Mechanism Design Goal

Admit truthfulness
$(e_i = 1, S_i = I)$
as the dominant strategy.

Related Work

Peer Prediction Mechanisms

Rewards determined by matching workers' answers with one another.

Dasgupta and Ghosh - WWW 2013
Radanovic, Faltings, Jurca – ACM TIST 2016
Shnayder, Agarwal, Frongillo, Parkes - EC 2016

and many others

Gold Standard Mechanisms

Rewards determined by matching with gold standard answers.

Gao, Wright, Leyton-Brown - arxiv 2016

Rewards for a constant number of workers are determined using gold standard answers and for the rest, using peer answers.

Deep Bayesian Trust Mechanism

1. Eliminates undesired equilibria of the peer-prediction mechanisms, while using the gold standard answers in a more scalable way.
2. Solves the unfairness problem of the peer-prediction mechanisms by ensuring that the reward of any worker is independent of the proficiency and strategy of her peer.

Transitive Trust Estimation

Given :

- Trustworthiness matrix of a worker j
- Answers of worker j and another worker i on a **(large)** set of common tasks.

Unknown :

- Trustworthiness matrix of the worker i.

Since the answers reported by the two workers are conditionally independent given the ground truth answer,

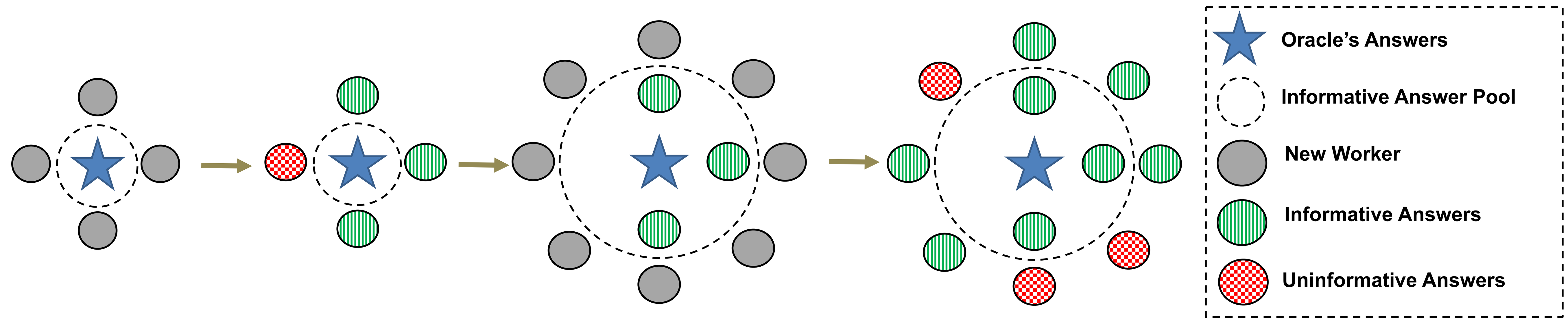
$$P(Y_i = 0 \mid Y_j = 0) = \sum_{k \in \{0,1\}} P(G = k \mid Y_j = 0) \cdot P(Y_i = 0 \mid G = k)$$
$$= \sum_{k \in \{0,1\}} \frac{T_j[k,0] \cdot P(k)}{P(Y_j = 0)} \cdot T_i[k,0]$$

Similarly for $P(Y_i = 1 \mid Y_j = 1)$

This system of linear equations can be solved for the unknown entries of the T_i matrix.

In non-binary answer space, we have a similar system of linear equations which can be solved to find all entries of the T_i matrix.

The Deep Bayesian Trust Mechanism



The DBT mechanism:

i. starts with an answer pool seeded with the oracle's answers,
ii. uses the answers in the pool to assess trust in answers submitted by other workers,
iii. rewards the workers based on the estimated trust, $r_i = \beta \cdot \{(\sum_{k \in [K]} T_i[k,k]) - 1\}$
iv. expands this pool based on the **informativeness** (determined by the estimated trust) of the workers' answers, and
v. repeats the process.

Properties*

*Please refer to the full paper for formal statements.

Theorem 1 : Given suitable scaling constant β, the Deep Bayesian Trust mechanism is Dominant Uniform Strategy Incentive Compatible with strictly positive expected reward in the truthful strategy.

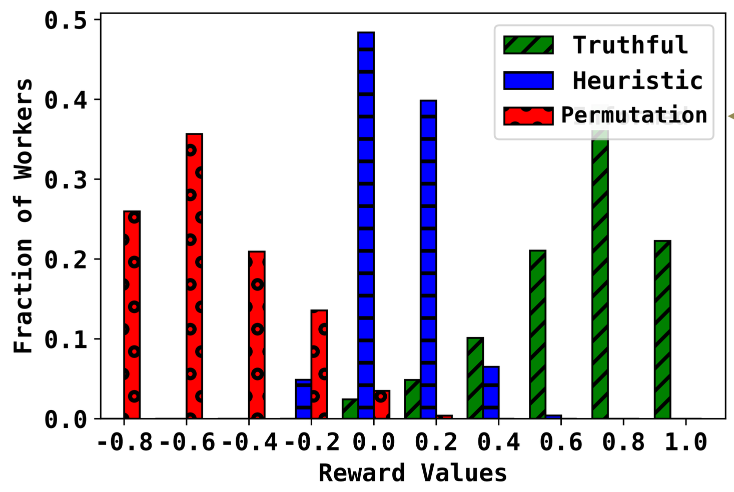
Corollary 1 : The scaling constant β of the Deep Bayesian Trust mechanism is independent of the probability of a worker getting oracle or another truthful worker as peer.

Theorem 2 : In the Deep Bayesian Trust mechanism, a heuristic strategy gives zero expected reward.

Theorem 3 : The Deep Bayesian Trust Mechanism is fair.

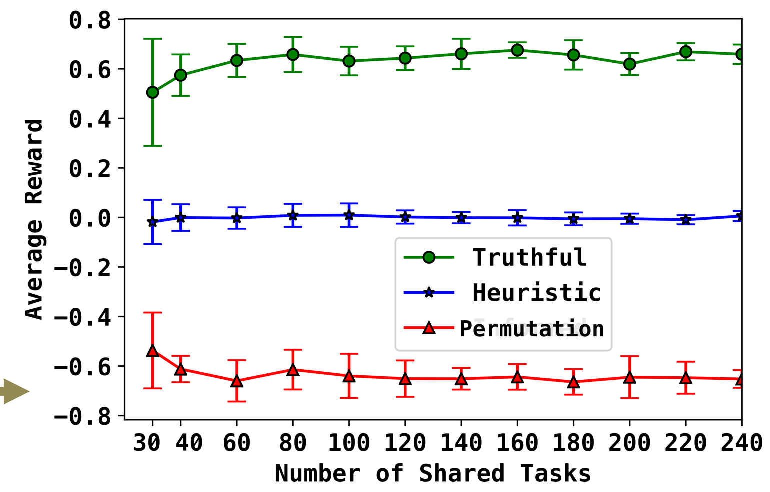
***(Fair Incentive Mechanism :** An incentive mechanism is called fair if the expected reward of any worker is directly proportional to the accuracy of the answers reported by her and independent of the strategy and proficiency of her random peer.)*

Simulations



The DBT mechanism correctly rewards the truthful workers and penalizes the non-truthful ones.

The DBT mechanism is robust *w.r.t.* the number of shared tasks between the peers.



ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Artificial Intelligence Laboratory
Naman Goel and Boi Faltings
{firstname.lastname}@epfl.ch