Eliciting High-quality Information

Boi Faltings

AI Laboratory, EPFL
boi.faltings@epfl.ch
http://liawww.epfl.ch/

2016

Eliciting Data

Organizations and individuals base decisions on data rather than principles:

- Financial markets
- Medicine
- Choosing a restaurant/hotel/spouse
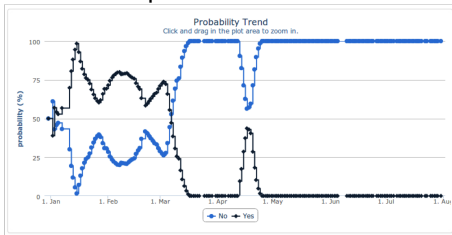- Law enforcement
- ...

Often, data must be obtained from others.

Product Reviews

- Reviews and ratings great to avoid poor products.
- Having reviews is essential for selling the product.
- But reviews are left for ulterior motives.

Forecasting Polls

Will Scotland become independent?

- Internet can be used to collect forecasts of important events.
- Important for many high-stakes decisions.
- How to encourage knowledgeable participants?

Example application: monitoring QoS

- Service provider serves a community of participants (internet, mobile phone, web service).
- Quality of Service, e.g. speed, response time, uptime, etc. are fixed by a contract.
- How to monitor quality of service?
  - route requests through a trusted intermediary: costly and inaccurate!
  - sample service: coarse measure of quality.
  - by clients themselves: incentives to be dishonest.
- Truthful reporting incentives can overcome this hurdle.
- $\Rightarrow$ much simpler, lower cost, very accurate.

Self-monitoring

- Groups of customers receive same service:
    - Mobile phone/Wifi in a certain area
    - Cloud computing
    - Students in a class
- Service Level Agreements can best be verified by customers.
- Self-monitoring could save a lot of cost.
- How can we ensure that customers report this honestly...
- ...given that they may get refunds for poor service.

Pollution Maps
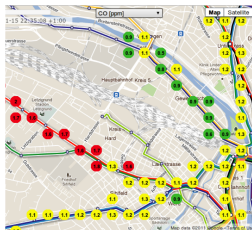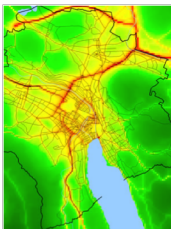
- Pollution, crop disease, etc. needs to be measured in many places.
- *Crowdsensing*: individuals deploy sensors and provide data to be aggregated into a map.
- But have to make sure to pay only accurate information.

Crowdwork

- Human computation: tasks solved by workers recruited through the internet (e.g. Amazon Mechanical Turk).
- Peer grading: students grade each others' homework.
- Huge benefits for knowledge acquisition, online courses, etc.

Why do we need Mechanisms?

Figure 3: Distribution of the Ratings on Amazon.com (fitted with a U-shaped curve) for a Music CD (Mr. A-Z)
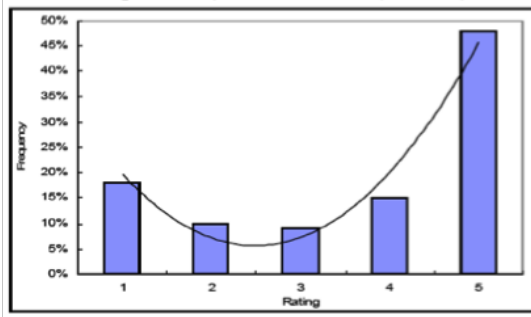
- laziness: most people do not write reviews, respond to polls, etc.
- self-selection: participation for ulterior motives, e.g. reviews paid for by hotel, push your own opinion, etc.
- malicious participants: paint a fake picture of reality.

Setting

1. Observe signals
2. Update beliefs
3. Report observations
4. Reward agents and publish the distribution of the reports
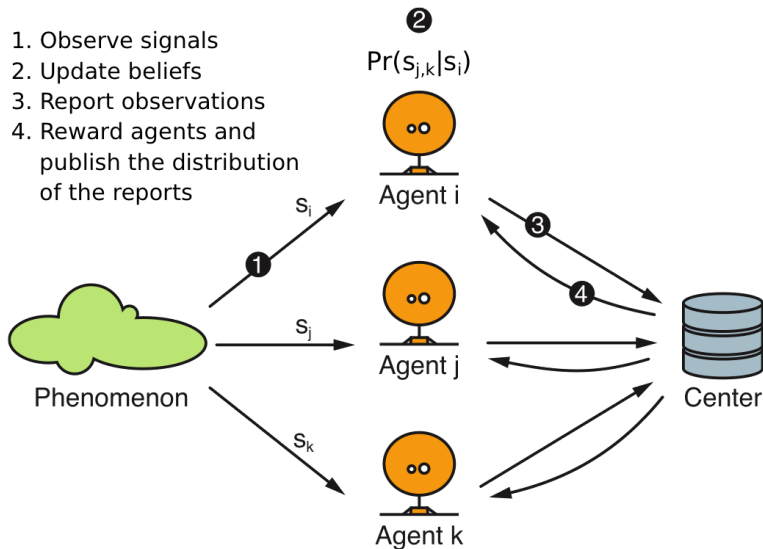
Principle underlying Truthful Mechanisms

Reward reports according to *consistency* with a *reference*:

- ground truth
- another *peer* report

Ground Truth as Reference

Forecasting: truth will eventually become known.

- value: reward accuracy.
- probability distribution: scoring rules.
- prediction markets: information aggregation as in a stock market.

Ground truth need not be known.
Reward data that is *surprisingly common*:

- common: agrees/is consistent with peer report.
- surprisingly: agreement is not expected.

Agents need to coordinate to obtain high rewards.
Coordination is only possible through the correct answer to the task.

Outline

- Scoring Rules
- Prediction Markets
- Peer Consistency
- Applications