Eliciting Truthful Information with the Peer Truth Serum

Motivation

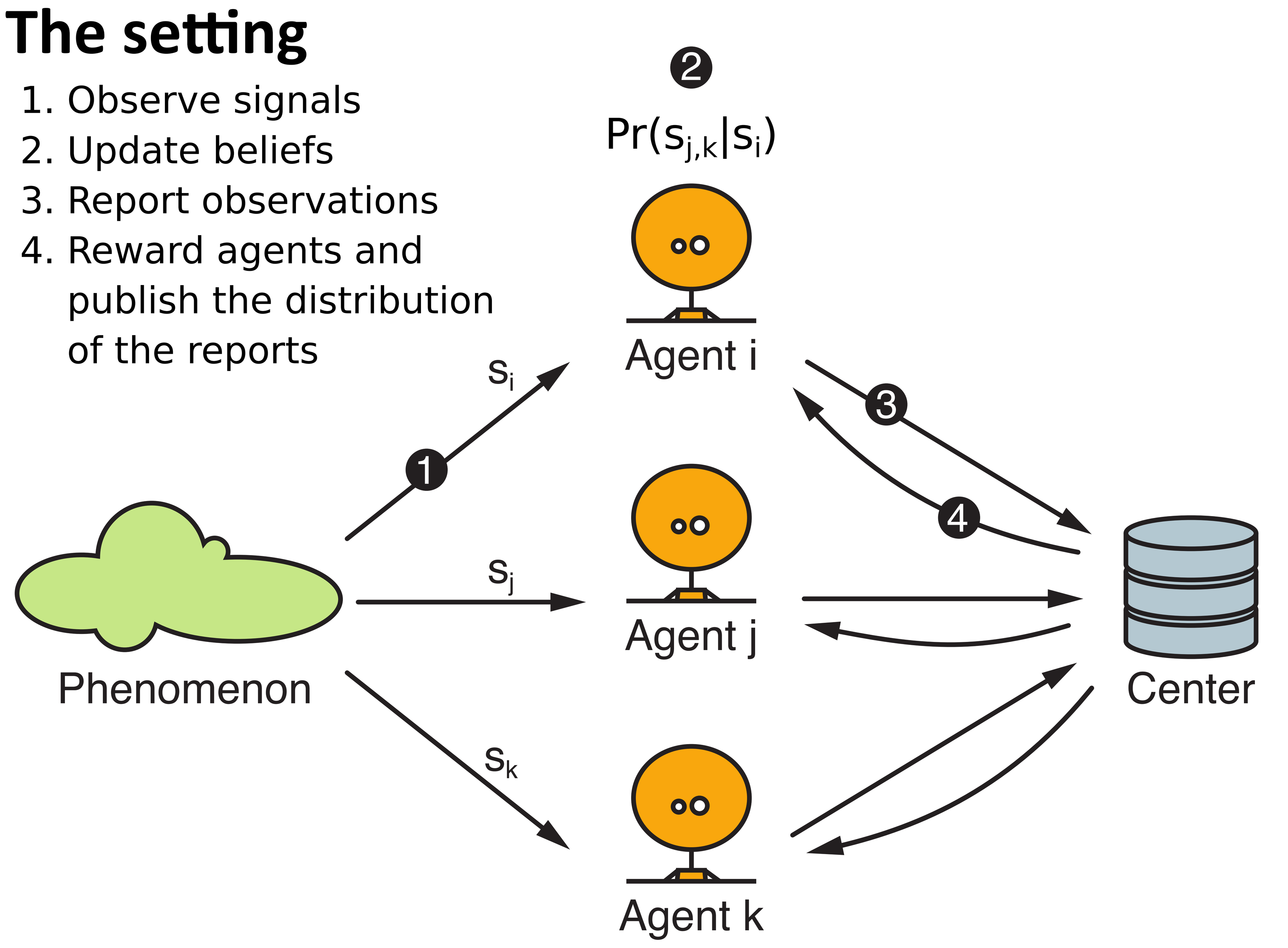
Information elicitation mechanisms represent an important component of many information aggregation techniques:

- Crowdsourcing
- Community sensing
- Product reviews
- Opinion polls

We investigate how to incentivize participants to reveal their private information when direct verification is not applicable.

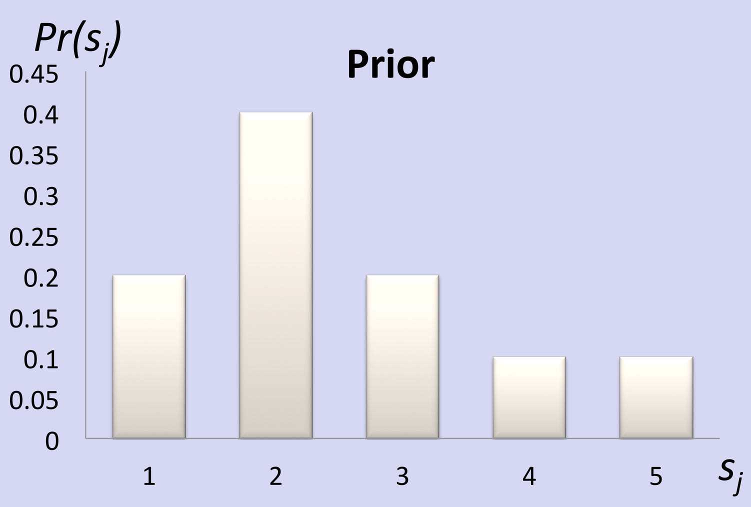
The setting

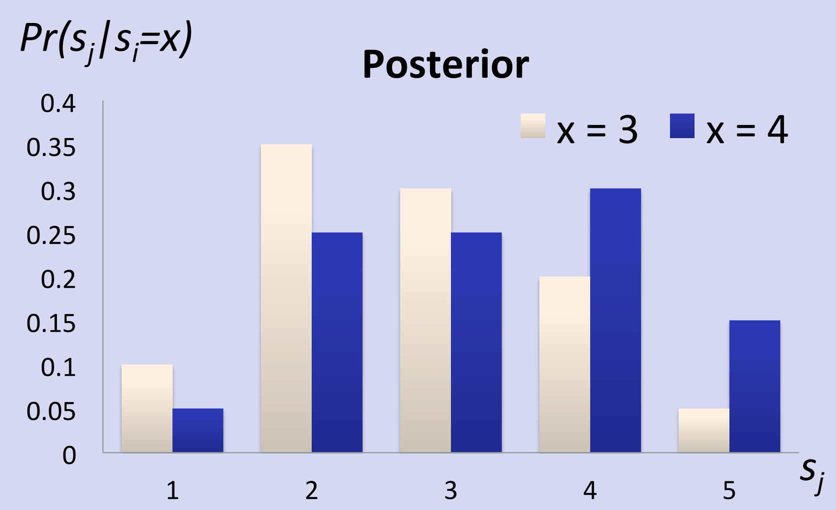
1. Observe signals
2. Update beliefs
3. Report observations
4. Reward agents and publish the distribution of the reports

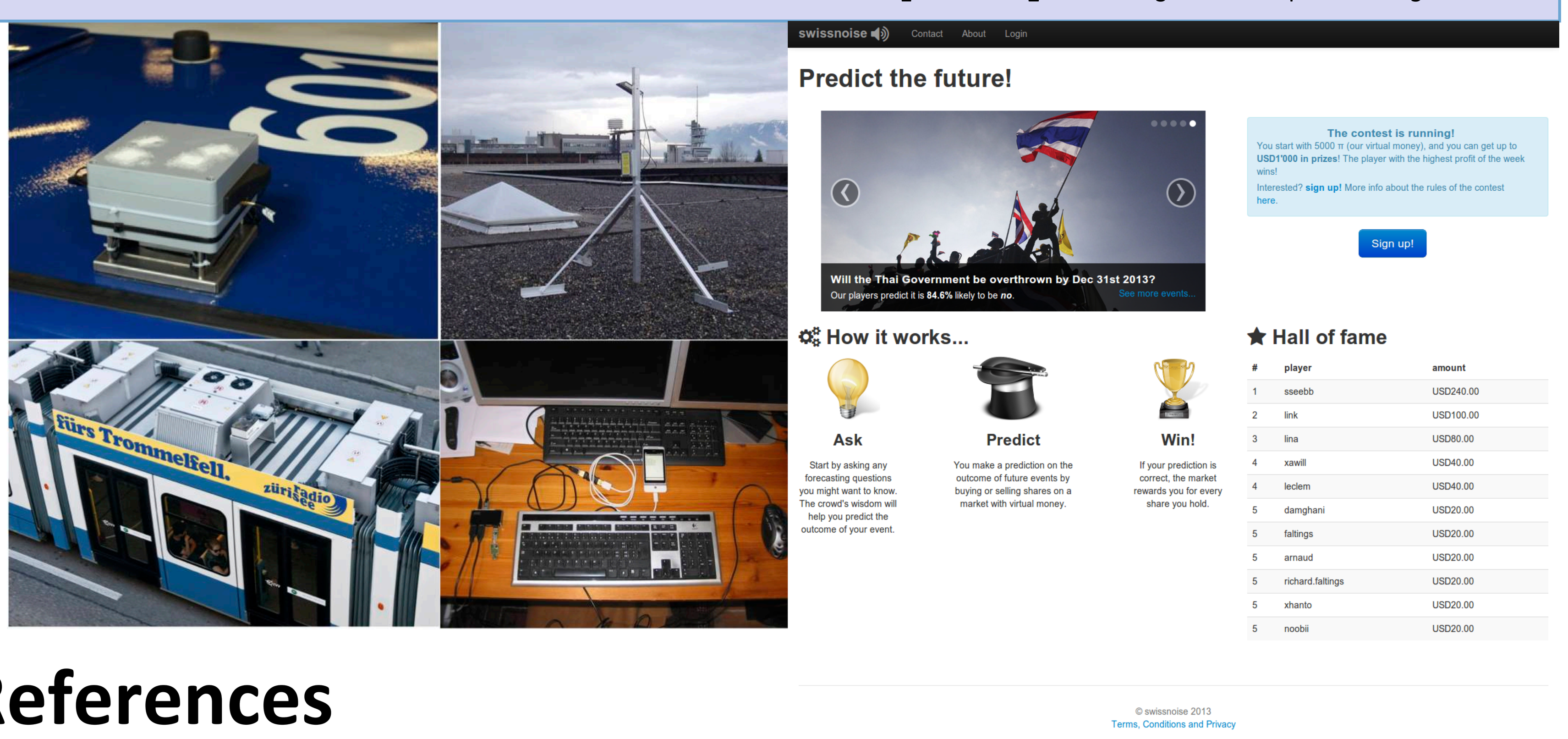


Self-predicting condition: Agents observation is a maximum-likelihood estimate of the true distribution, also seen by its peers.

$$\frac{\Pr(s_j = x_i \mid s_i)}{\Pr(s_j = x_i)} \geq \frac{\Pr(s_j = y \mid s_i)}{\Pr(s_j = y)} \qquad s_i = \text{argmax}_{\tilde{x}} \Pr(s_i \mid s_j = \tilde{x})$$







The Peer Truth Serum

Compute reward by comparing with peer report xj.
1 is an indicator variable, **R**(x_i) = histogram of x_i, a>0

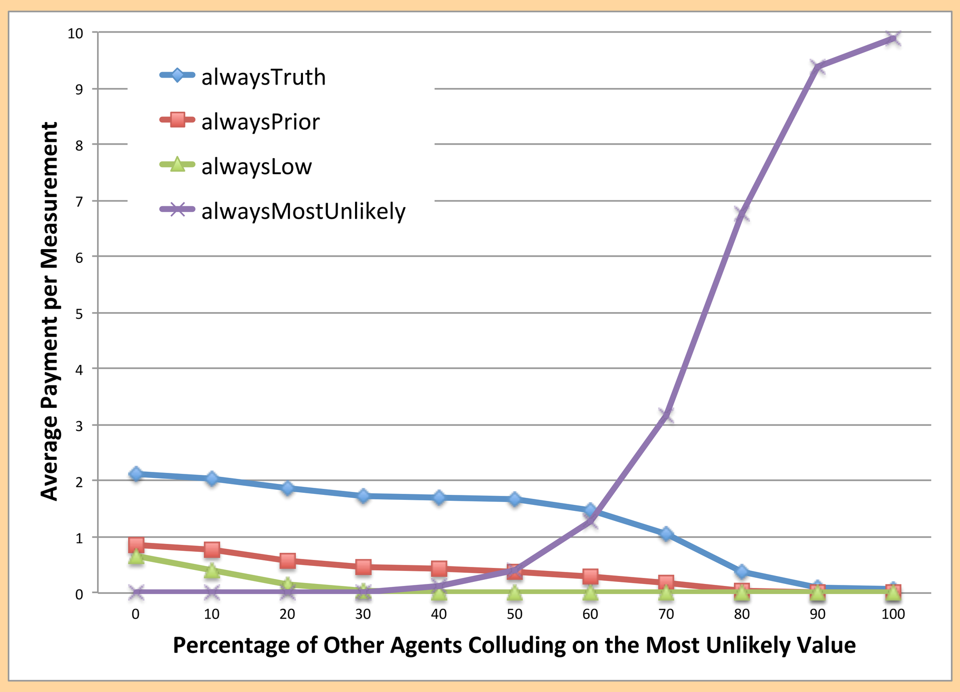
$$reward_i = a\frac{\mathbf{1}_{x_i = x_j}}{\mathbf{R}(x_i)} + b$$

Truthfulness: if agents' priors are close to **R**, and the self-predicting condition holds, truthful reporting ($x_i = s_i$) is a Bayes-Nash equilibrium.

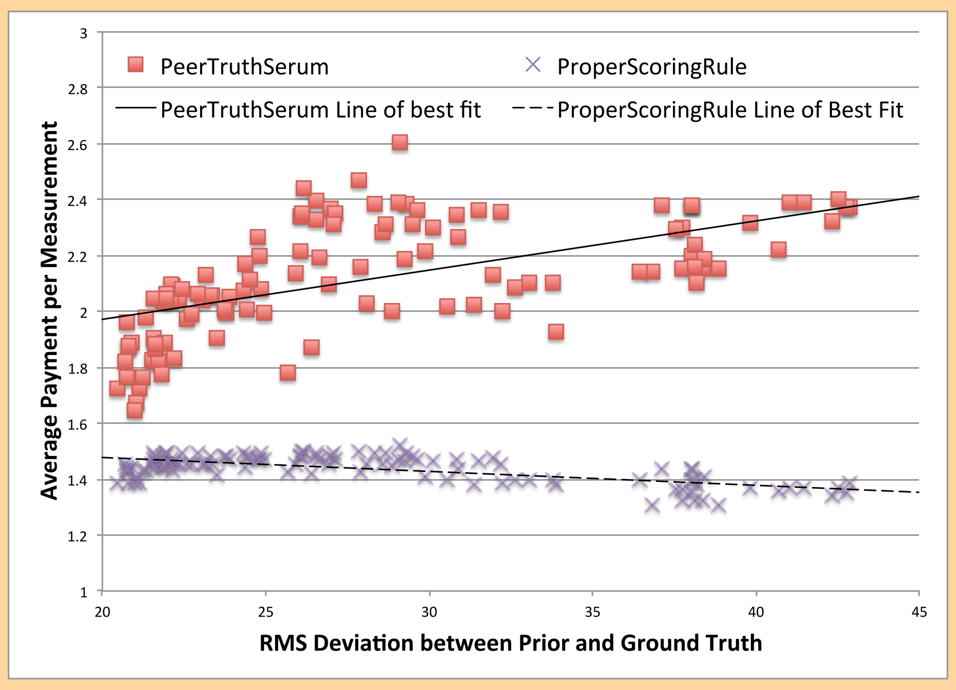
Uniqueness: if the self-predicting condition is the only assumption, any Bayes-Nash truthful scheme has the form of the peer truth serum.

Helpfulness: if agents' priors are far from **R**, but more informed (closer to the true distribution of the signal), and the self-predicting condition holds, PTS supports equilibria in helpful strategies that make **R** converge to the prior and are thus *asymptotically accurate*.

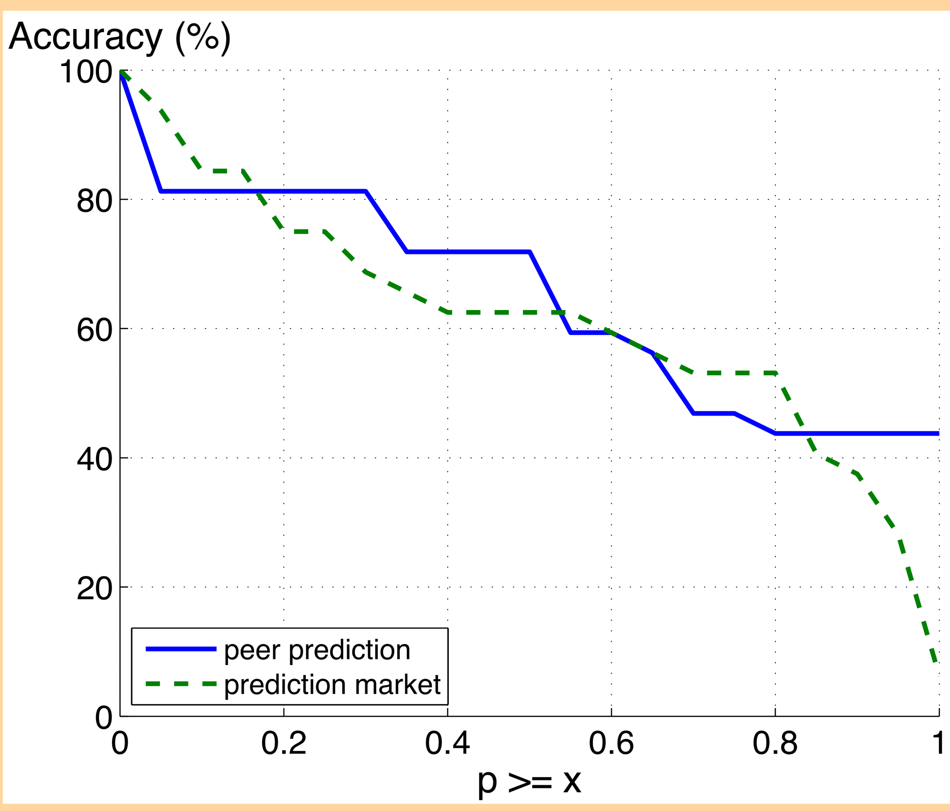
Empirical Performance



Sensing: truthfulness is the most profitable strategy even for a significant number of colluders (simulation using air quality model on real data from the city of Strassbourg).



Sensing: PTS encourages measurements that bring new information better than peer prediction with scoring rules, thus making self-selection work better (same simulation as above).



Prediction poll: peer prediction using PTS provides similar accuracy to classical prediction markets, but require no ground truth (data from the swissnoise.ch platform with about 200 users on 30 events).

References

- Jurca, R. and Faltings, B. *Incentives for answering hypothetical questions*. EC Workshop on Social Computing and User-Generated Content, 2011.
- Faltings, B., Li, J. J. and Jurca, R. *Incentive Mechanisms for Community Sensing*. IEEE Transaction on Computers, 63(1), 115-128, 2014
- Garcin, F. and Faltings, B. *Swissnoise: Online Polls with Game-Theoretic Incentives*. IAAI, 2014.
- Radanovic, G. and Faltings, B. *Incentives for Truthful Information Elicitation of Continuous Signals,* AAAI, 2014.

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Boi Faltings and Goran Radanovic
EPFL Artificial Intelligence Laboratory
{firstname.lastname}@epfl.ch
Radu Jurca, Google Zürich
Radu.jurca@gmail.com