Games and Economic Behavior 86 (2014) 458–474

Contents lists available at SciVerse ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

Designing incentives for online question-and-answer forums

Shaili Jain [a,*], Yiling Chen [b], David C. Parkes [b]

[a] Computer Science Department, Yale University, New Haven, CT 06520, United States
[b] SEAS, Harvard University, 33 Oxford St., Cambridge, MA 02138, United States

ARTICLE INFO

Article history:
Received 21 May 2012
Available online 14 December 2012

JEL classification:
D82
D47

Keywords:
Question and answer forums
User generated content
Contests

ABSTRACT

We provide a game-theoretic model of sequential information aggregation motivated by online question-and-answer forums. An asker posts a question and each user decides when to aggregate a unique piece of information with existing information. When the quality exceeds a certain threshold, the asker closes the question and allocates points to users. We consider the effect of different rules for allocating points on the equilibrium behavior. A best-answer rule provides a unique, efficient equilibrium in which all users respond in the first round, for substitutes valuations over information. However, the best-answer rule isolates the least efficient equilibrium for complements valuations. We demonstrate alternate scoring rules that provide an efficient equilibrium for distinct subclasses of complements valuations, and retain an efficient equilibrium for substitutes valuations. We introduce a reasonable set of axioms, and establish that no rule satisfying these axioms can achieve the efficient outcome in a unique equilibrium for all valuations.

1. Introduction

In online question-and-answer forums such as Yahoo! Answers, users can post questions and answer questions on wide variety of topics. In particular, Yahoo! Answers has 25 categories ranging from *Computers & Internet* to *Travel* to *Family & Relationships* to *Health*. Users may post discussion questions, factual questions or polls. In Yahoo! Answers, people do not exchange money for the exchange of information, but instead receive points for contributions that influence leaderboard and top-contributor designations, while also allowing users to post their own questions.[1]

We study a game-theoretic model of a problem of sequential information acquisition that is motivated by these online question-and-answer forums. In particular, our model is suitable for the study of the design of methods to assign scores to answers to factual questions, such as "What were the main causes of the Great Plague of London?". This is because we model the value of answers submitted to a question as strictly increasing over time, either because users incorporate their own private information with the information provided by earlier reports, or because the asker is able to piece together different responses and thus has an increasing value for the answers. Harper et al. (2009) have demonstrated that factual questions have a higher archival value than discussion questions, justifying the focus of this work. An example of a discussion question is, "What is your favorite movie of all time?". Our model is less suitable for these questions, where it makes less sense for subsequent answers to improve on earlier answers and to incorporate earlier responses.

* Corresponding author.
E-mail addresses: shaili.jain@yale.edu (S. Jain), yiling@eecs.harvard.edu (Y. Chen), parkes@eecs.harvard.edu (D.C. Parkes).

[1] A user is placed into one of seven levels based on her number of points. The higher the level, the greater privileges a user will get in terms of the number of questions she can ask per day. All users have a profile where the number of points the user has, the level, and the percentage of best answers are clearly displayed. In addition to the point system, the "top contributor" is displayed at the top of the page for each semantic category and sub-category, and there is a leaderboard of the top ten users.

0899-8256/$ – see front matter
http://dx.doi.org/10.1016/j.geb.2012.11.003

Whereas the established model of *contest design* (Moldovanu and Sela, 2001, 2006) considers agents with costly but independent effort and seeks to maximize the total effort exerted across all agents, in our model each agent can build on the work already contributed by other agents, and submit a solution that dominates all solutions so far submitted. In keeping the model simple, we choose not to model the cost of contributing an answer and model the asker as satisficing, with a private *quality threshold* at which she will close the question. In the model, the asker privately draws this quality threshold at random, and is satisfied with any answer with value above this threshold. The asker prefers to receive a satisfactory answer sooner rather than later, and closes the question as soon as the threshold is exceeded.

Each user holds a unique piece of information that is relevant to a question, and the strategic decision is decide when to report this information and aggregate it with previous reports. As information is reported it is aggregated into the responses, so that the value to the asker monotonically improves while the question remains open. In the case that multiple pieces of information are simultaneously revealed, we assume that the asker is able to aggregate the information, and we associate each user that contributed in the round with an answer equal in quality to that achieved through this aggregation. We study two distinct models for the way in which individual contributions contribute to an overall value of the asker. In particular, we study a *complements case*, in which each successive piece of information is worth more to the asker than the previous one, and a *substitutes case* in which each successive piece of information is worth less than the previous one.

Our interest is in characterizing the properties of subgame perfect Nash equilibria of this game under different rules for assigning points, and for users who seek to maximize their expected number of points. By delaying a contribution, a user runs the risk that the asker will be satisfied with the current answer and close the question. On the other hand, by delaying a contribution, a user can take advantage of contributions by other users and submit a better answer, thereby increasing the probability that a user's answer (if submitted before the question has been closed) will cross the quality threshold of the asker.

As a designer, we seek to understand which rules induce equilibria in which all users choose to contribute their respective pieces of information in the first round, and thus immediately and without delay. We first analyze the equilibria for a *best-answer* rule, which models the current Yahoo! Answers environment. We find that this rule is effective for substitutes valuations, where it isolates a subgame perfect Nash equilibrium in which all users reveal information in the first round. This is the efficient outcome, with the asker receiving a satisfactory answer as soon as possible for all possible quality thresholds. On the other hand, the best-answer rule is ineffective for complements information, where it isolates an equilibrium in which every user posts information in the very last round. For the case of complements information, the expected gain from an answer with higher quality, that comes from playing later and combining an answer with previous answers, is greater than the negative effect of delaying and risking that the question will close before submitting an answer.

In addressing this problem, we consider two alternative rules for assigning points to answers. The first rule is an *approval-voting rule*, parameterized by integer $k > 1$, in which the asker assigns one point to the most recent $k > 1$ answers (or some random k subset if more than k answers were received in the most recent round) upon closing the question. The approval-voting rule retains the efficient outcome in an equilibrium of the game for substitutes valuations. The approval-voting rule also enables the most efficient, all-going-first outcome in an equilibrium for complements information, under certain restrictions on the valuation function. But, the approval-voting rule also retains an equilibrium for complements information in which every user plays in the last round. More problematically, the approval-voting rule also introduces this inefficient outcome in an equilibrium of the game with substitutes valuations.

The second rule that we introduce is the *proportional-share* rule, in which the asker assigns a share of the total available points in proportion to the marginal value contributed by a user in the round in which the user participates. Like the approval-voting rule, the proportional-share rule enables the most efficient outcome in equilibrium for a large class of complements information. In addition, the proportional-share rule retains the efficient outcome as an equilibrium for substitutes valuations, and unlike the approval-voting rule, this remains the unique equilibrium.

The approval-voting and proportional-share rules both avoid the incentives for delaying to the last round in the setting of complements information by spreading the score across more users. They do this in different ways. In particular, both the approval-voting and proportional-share rules are able to achieve the efficient outcome as a subgame perfect Nash equilibrium for certain classes of complements valuations. The approval-voting rule, but not the proportional-share rule, also introduces the inefficient outcome in an equilibrium for environments with substitutes valuations. On the other hand, the approval-voting rule is a simple generalization of the best-answer rule and likely more relevant to practice because it does not require new information from the asker when a question has been closed.

A natural question is whether there can be method of assigning points that is first best, in that it isolates the efficient outcome as a unique equilibrium for all possible complements and substitutes valuation functions. We obtain a negative result in this regard—we introduce three axioms, *anonymity, monotonicity* and *time-invariance*, and show that there is no payment rule that satisfies these properties and enables the efficient outcome in a unique subgame perfect Nash equilibrium.

1.1. Related work

We believe this to be the first work studying online question-and-answer forums in a game-theoretic light. In terms of game-theoretic analysis of other systems of human computation (von Ahn and Dabbish, 2008), prior work has presented a game-theoretic analysis of the *ESP game* (Jain and Parkes, 2013) and the *PhotoSlap game* (Ho et al., 2007). These are so-called *Games with a Purpose*, games that are designed to be fun to play, with the added benefit that users are doing useful