A Case Study of Reputation and Language

Jonas Nick

Written report for the course Game Theory and Pragmatics

Abstract. The following paper will first introduce the notion of reputation in game theory, which is compared to the reputation system of the web site Stackoverflow[1]. Then, the influence of reputation on language will be examined, using questions from Stackoverflow.

1 Introduction

In former times, reputation was determined by personal relationships, hearsay or mass media. Nowadays, in the information age, interacting with strangers becomes more and more beneficial. The information retrievable about your interaction partner is typically sparse, still you want to know whether to take his information for granted or if his opinion is valuable on a blog or social media site.

When interacting on marketplaces, like ebay, trusting a stranger also involves risk. Ebay supplies a reputation system [1] since its beginnings keeping track of the feedback a seller or buyer receives. It was shown that there is a significant relation to the probability of the buyer or seller's next transaction happening without difficulties and that the price is significantly affected by reputation [2].

Market interactions have traditionally been analyzed with the help of game theory. Further, game theory has formalized the notion of reputation clarifying how rational agents should behave when reputation comes into play. These techniques are applied on the design of reputation systems such that they allow efficient behavior in an equilibrium and reduces the possibility of abuse [3], [4].

The reputation system of the question&answer site Stackoverflow will serve as a case study to show the relationship between reputation and the probability of posting invalid questions. This information is useful in deriving models for predicting the validity based on various features. Reputation turns out having a significant influence, as well as other simple text based features. The reasons why a question is closed are related to Grice's Maxims [5], which allows to specifically point out the influence of reputation on language.

2 Reputation in Game Theory

The following concepts of repetition and reputation will be illustrated using an asymmetric trading game (Table 1). In this game the row player is a seller, who

[1] `http://stackoverflow.com/`

	Buy(B)	Not Buy(N)
High Quality(H)	1, 1	-1, -1
Low Quality(L)	2, -1	0, 0

Table 1. The trading game

can produce high or low quality and the column player can decide whether to buy the product. The buyer only wants to buy high quality, but selling low quality is more desirable because the production requires less resources. In addition, the buyer regrets not buying a high quality product but he doesn't regret not buying low quality. The trading games unique Nash equilibrium at (L, N) is clearly inefficient for the participating players.

2.1 Repeated Game

Repeated games are models where the same set of agents repeatedly play the same game called the stage game (cf. [6]). In consequence, strategies σ become sequences of stage game strategies $\sigma^{(t)}$ in period t, depending on a history $h^{(t)}$ of previous stage games. The payoff to player i when each player executes her component of the stage-game strategy profile $\sigma^{(t)}(h^{(t)})$ is $g_i(\sigma^{(t)}(h^{(t)}))$. Additionally, future payoffs are discounted by a factor $\delta \in (0, 1)$, which weighs future interactions and represents a players patience. This leads to the following utility function for player i until period T in a repeated game:

$$U_i(\sigma) = \sum_{t=0}^{T} \delta^t \, g_i(\sigma^{(t)}(h^{(t)}))$$

As usual, a Nash equilibrium in a repeated game is a strategy profile where each players strategy σ is a best response given the other players strategy σ_{-i}, or formally if for all players i and strategies σ'_i,

$$U_i(\sigma) \geq U_i(\sigma'_i, \sigma_{-i})$$

The equilibria differ between finitely and infinitely repeated games. The former equilibria are found by backwards induction and in the case of our trading game the equillibrium (L, N) can be found because the players will choose to play the stage game equilibrium in the ultimate interaction regardless of their penultimate interaction, and so on. Infinitely repeated games on the other hand entail an infinite set of strategies and many of them may be part of an equilibrium. Further, playing over an infinite horizon is equivalent to the case where a player does not know when an interaction ends. For example, $(L,\ N)$ stays a Nash equilibrium in the repeated game but a so-called trigger strategy s facilitates cooperative behavior:

1. start with H

2. Stick to H, as long as buyer deviates from B or oneself deviates from H
3. once anyone deviates, charge L

The strategy profile where the seller plays s and the buyer plays analogous is a Nash equilibrium because deviation would lead to a short term gain of at most $\pi = \text{EASYWIN} - \text{WIN} = 2 - 1$, which is lower than sticking to H if the player is sufficiently patient.

The equilibrium strategy profiles can be characterized using the folk theorem. [2]

It states that any mutually beneficial outcome can be sustained in a long term relationship if players are sufficiently patient. Concretely, the set

$$V \equiv \{(x, y) : x > 0, y > -\frac{1}{3}, y \leq x, y \leq 3 - 2x\} \subset \mathbb{R}^2$$

is the set of feasible and strictly individual rational payoffs for our trading game [7] (cf. *minmax payoff*). Every payoff profile in V can be an equilibrium outcome of the repeated game according to the folk theorem. If the seller would commit to a pure strategy she would choose H because the buyer's best response to this is B. However, she could do better by committing to a mixed strategy, for example, playing $(\frac{3}{4}, \frac{1}{4})$ would ensure the buyer played B and give the seller a higher payoff.

In fact, the folk theorem does not narrow down the equilibrium strategy profile only to the efficient (H, B) outcome, but it can be enforced if both players agree on that equilibrium beforehand, via a (social) contract, or if some uncertainty about players is added.

2.2 Reputation

When speaking of the reputation of someone, we mean that he has a reputation for belonging to a specific type or showing specific behavior. For instance, in the trading situation a seller could have the reputation for her commitment to quality (a strong seller), surely influencing the buyers' behavior. When using game theory methods in the example, the probability $p^{(t)}$ as perceived by the buyer at time step t of the seller being one of a fixed set of types represents the reputation of the seller. Initially, $p^{(0)}$ is 50% and in the rest of the game the buyer tries to determine the seller's type. They differ in that the strong player plays strategy $\hat{\sigma}$ and the normal player plays $\tilde{\sigma}$. Therefore, the buyer expects the seller to play action a with probability $\bar{\sigma}^{(t)}(a) = p^{(t)}\hat{\sigma}^{(t)}(a) + (1 - p^{(t)})\tilde{\sigma}^{(t)}(a)$. The property of the game that a players type and therefore his preferences and payoffs are unknown to the other player is called incomplete information. This situation, where the buyer tries to infer the type of his counterpart, is represented by modeling the player as a bayesian agent, who revises his beliefs according to

[2] Here we restrict ourselves to the perfect monitoring case: after a stage game the players correctly observe the action of other players.

the observed action $a^{(t)}$. Application of the Bayes' rule leads to the following update rule:

$$p^{(t+1)} = \frac{Pr(a^{(t)} \cap \text{Strong})}{Pr(a^{(t)})} = \frac{p^{(t)}\hat{\sigma}^{(t)}(a^{(t)})}{\bar{\sigma}^{(t)}(a^{(t)})}$$

Hence, every action of the seller serves as a *signal* for his type. It can be shown that a bayesian agent will either learn the type of the seller and can perfectly predict their actions or both seller types end up playing the same strategy, which in turn also allows the buyer to perfectly predict the actions (Merging property [7]).

In the reputation building phase, the seller has to consider the effect of his actions in order to achieve or maintain the reputation of a strong seller. Offering low quality is always more beneficial in the short run, but the buyer would adjust the probability of a strong seller downwards, resulting in less buying in the future. In contrast, high quality today increases the probability of the strong player and therefore the likelihood of buying.

Thus, one option of the normal player is mimicking the behavior of the commitment type, acquiring a reputation of behaving like this type. Exactly how the normal seller chooses to trade off long run costs benefits and short run costs is unclear. If there was no strong seller present, the normal seller had no particular reason for expecting that high quality would necessarily be rewarded with higher sales.

In fact, a small probability of the presence of a strong seller places a lower bound on the equilibrium payoff for a normal seller, eliminating some kinds of equilibria. However, there are finite opportunities for the normal seller to convince of her commitment since her type will be known sometime according to the merging property. The exact lower bound of the equilibrium payoff of a normal seller depends on whether there are short term buyers, one long run buyer or multiple long run buyer. When looking at short term buyers the normal seller will receive a stage game payoff of $R(b)$, where

$$R(b) := 2 - b$$

and b is the probability of the strong player producing high quality ($> \frac{1}{3}$) in an equilibrium, ergo her commitment. The lower bound of equilibrium payoff is therefore 1, which excludes some of the equilibria payoffs achieved in simple repeated games, where the seller might obtain $1 > x > 0$ (Folk Theorem).

In conclusion, the reputation effect, stating that signals sent will affect current and future behavior, will discard some inefficient equilibria.

3 Case study

After examining some theoretical concepts of reputation in games, a closer look on actual reputation systems will reveal the importance of reputation in efficient communication. A reputation system aggregates feedback about the past behavior of a participant and quantifies the reputation using a specific reputation algorithm.

3.1 Stackoverflow

The question and answer website Stackoverflow (SO) can be seen as an example for such a reputation system. Users on the site pose programming related questions which others try to answer. They are encouraged to vote on the usefulness of a question or an answer, thereby directly affecting others reputation score. Because SO is collaboratively edited website, reputation directly determines the privileges of a user, ranging from voting down and editing to voting on closing or deleting questions and answers. Thus, reputation on SO is among other thing a measure of how much the community trusts a user[3].

SO offers monthly data dumps and which allows us to examine the effect of reputation on the behavior of users. The data consists of the text from users, some meta information, for example his reputation score or the tags (categories) a question was annotated with, and whether the question was closed. The latter was recently subject of a machine learning competition hosted by SO where the goal was to predict the OpenStatus of the question (open or closed and why) based on the data. Because of that, we will focus on this binary outcome of an interaction with SO. In general, about 6% of all new questions end up closed.

3.2 Stackoverflow and the trading game

It turns out that the game depicted in table 1 is a appropriate model for studying the events on SO and the influence of reputation. A user interacting with SO can either submit high or low quality content. Therefore he can be associated with the row player in the trading game. The collectively maintained institution SO is the other player in the interaction, which can either choose to accept or discard the user input, whereby the latter is achieved by closing the question. The assumption that producing high quality content is costly is crucial, because otherwise there would be no explanation for closed questions after all. Instead of showing individual users' history, the data provided by SO contains the column players action, namely the resulting OpenStatus. The action of the row player, high or low effort, underlies imperfect monitoring because the only information provided to SO is the text written by the questioner.

3.3 Influence of reputation

As we have clarified before players have a reputation *for* something. On SO, users have reputation for being reliable and trustworthy. This reputation score is comparable to the notion of reputation in game theory because the score has a relation to the perceived probability of the user being the trustworthy type. Figure 1 depicts a section of the density of reputation of questioners in July 2012. We see that there are many users with low reputation, higher reputation is getting more and more uncommon. It looks like the distribution is following a power law, which is strengthened by Figure 2 showing that the distribution is

[3] `http://meta.stackoverflow.com/questions/7237/how-does-reputation-work`

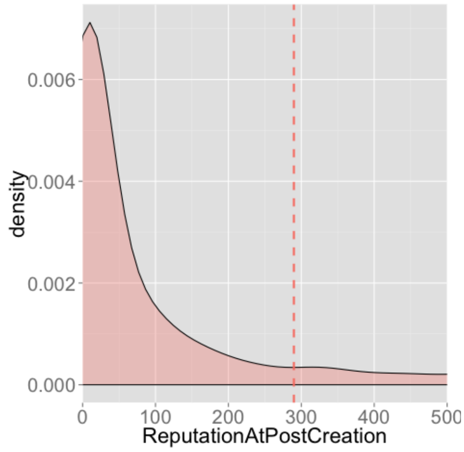

Fig. 1. Density of reputation (section). Dashed line represents mean.

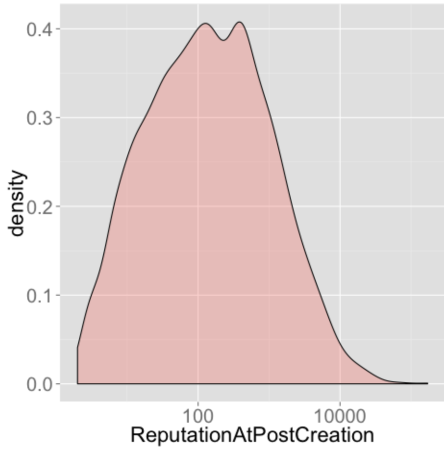

Fig. 2. The density of reputation on a log scale, without reputation = 1.

approximately log-normal, when users with reputation scores equal to one are excluded. If those are included there is a peak near ReputationAtPostCreation= 1, which is then also the maximum. The exceeding fraction users with a reputation equal to one is explained by initial reputation of new users on Stackoverflow - they start out with this reputation value.

When investigating reputations given the question was closed or left open we can see that closed questions are posed mainly by users with low reputation (Figure 3 and 4). One interpretation is that a user with low reputation belongs to one of two different user categories, whose members have an incentive to choosing low effort. Those are users who have low reputation because they are not trustful and new users who discount the future immensely because they have a single specific question.

The inverse argument, that questions posed by users with higher reputation have a lower probability of posing questions that end up closed is strengthened using a logistic regression model[4] (Figure 5). Also, the decision boundary is shown in Figure indicating the reputation where the model estimates a 50% probability of a closed question. The result is that reputation is a significant influence and this model alone has an accuracy of 59.44% on test data.

[4] For an introduction to logistic regression refer to [8]

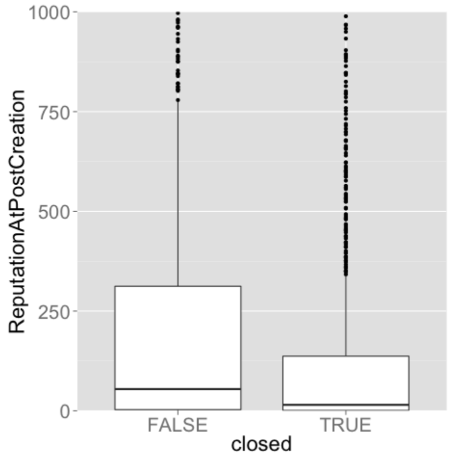

Fig. 3. Boxplot of the reputation of users that have posed a closed or accepted question (section).

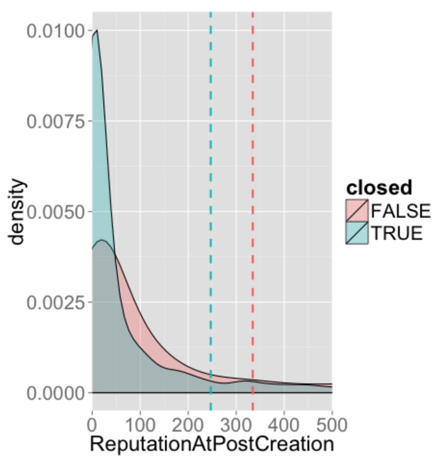

Fig. 4. Density of reputation of users that have posed a closed or accepted question (section). Dashed line represents means.

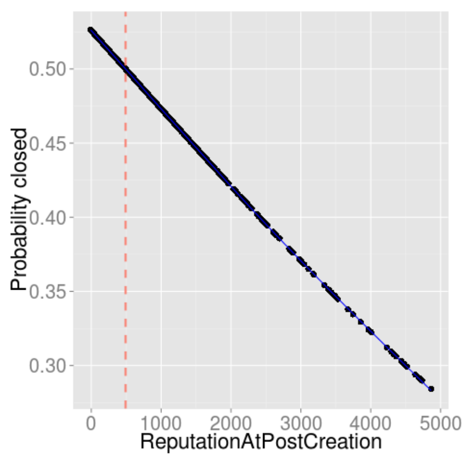

Fig. 5. Predicted probability obtained by logistic regression of a closed question given the user's reputation. Dashed line at $x \approx 491$ represents decision boundary.

3.4 Reputation and language

Naturally, the decision of SO whether to close the question is based on the text of the question. Therefore, features extracted from the language of questions of Stackoverflow are able to to improve the accuracy of the logistic model to 66.8%. The additional extracted features were the number of characters in the title and text (Figure 6) and the fraction of stopwords[5].

Fig. 6. Density of the characters in a question given closed. Dashed lines represent the mean.

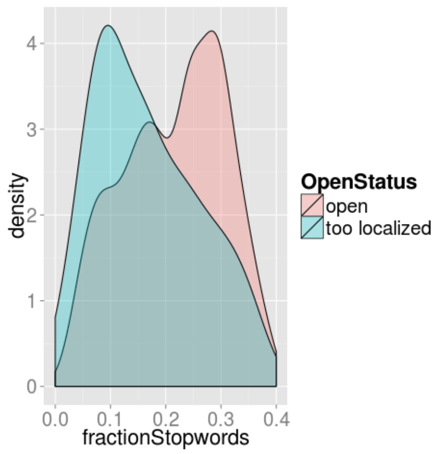

Fig. 7. Density of the fraction of stopwords in a question given the question remained open or was closed because of being too localized.

When closing a question a moderator specifies a reason for doing so, namely *off topic*, *not constructive*, *not a real question*, or *too localized*.[6]

There is a relation to Gricean Maxims[7], which are presented as guidelines of how to communicate successfully, but also convey presumptions of speaker and listener while communicating [9].

Questions labeled *off topic*(not related to programming) and *too localized*(unlikely to help future visitors) clearly violate the maxim of relevance. *Not a real question* are those that are ambiguous, vague, incomplete, overly broad, or rhetorical, hence the maxims of manner and quantity are both violated. The maxim of quality is violated by questions labeled *not constructive* because the question are not supported by facts. Rather, it would solicit debate, since there is no true answer.

[5] Stopwords are words that do convey little information like articles conjunctions and prepositions

[6] `http://stackoverflow.com/faq\#close`

[7] The maxims are quality (be truthful), quantity (make your contribution as informative as is required, but not more), relation (be relevant) and manner (be clear)

For example, the influence of simple language features like the fraction of stop-words on the probability of obedience to the maxim of relevance (too localized) can be examined (Figure 7).

A deeper analysis reveals that reputation influences which maxims are violated. Users with a high reputation are more likely violating the maxim of quality than the maxim of quantity or manner.(Figure 8)

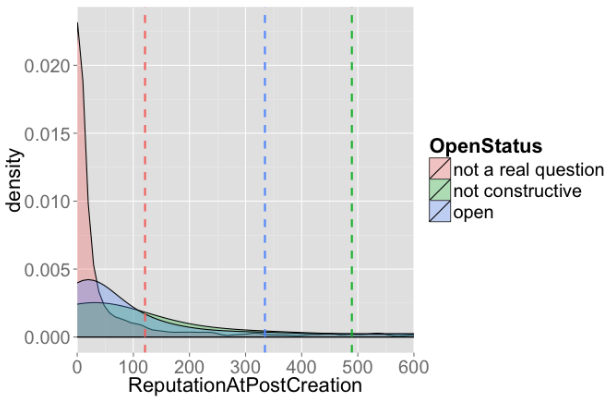

Fig. 8. Density of reputation of users that have posed an accepted question or one which was flagged "not a real question" or "not constructive" (section). Dashed line represents means.

Most questions that are incomplete are posed by low reputation users, while controversial questions are posed by high reputation users.

3.5 Conclusion

The reputation on Stackoverflow is log-normally distributed, except that there are overly many new users with reputation = 1. They are responsible for most of the closed questions, specifically by being not concise and clear. In contrast, questions that lack factual support are posed by users with higher reputation.

When it comes to the task of predicting whether a question will be closed, we come to the conclusion, that reputation is an important feature: It facilitates efficient communication in online communities theoretically as well as practically, because it carries information about the probability of the user's reliability. However, the dataset was studied on the surface in the previous work, compared to potential findings related to reputation, specifically when focusing on the interactions between language and reputation in a statistical model.

4 Analysis Specifics

The R [10] code for analysis and plots is available under http://github.com/jonasnick/Gricean-Classifier/.

References

1. Paul Resnick, Ko Kuwabara, Richard Zeckhauser, and Eric Friedman. Reputation systems. *Commun. ACM*, 43(12), December 2000. http://doi.acm.org/10.1145/355112.355122.
2. Daniel Houser and John Wooders. *Reputation in Auctions: Theory and Evidence from eBay.* Mimeo, 2001.
3. K. Aberer and Z. Despotovic. On reputation in game theory. 2004.
4. L. Mui, M. Mohtashemi, and A. Halberstadt. A computational model of trust and reputation for e-businesses. In *Proceedings of the 35th Annual Hawaii International Conference on System Sciences (HICSS'02)-Volume 7 - Volume 7*, HICSS '02, pages 188–, Washington, DC, USA, 2002. IEEE Computer Society.
5. H. P. Grice. Logic and conversation. In Peter Cole and Jerry L. Morgan, editors, *Syntax and semantics*, volume 3. New York: Academic Press, 1975.
6. Kandori Michihiro. repeated games. In Steven N. Durlauf and Lawrence E. Blume, editors, *The New Palgrave Dictionary of Economics.* Palgrave Macmillan, Basingstoke, 2008.
7. Martin W. Cripps. reputation. In Steven N. Durlauf and Lawrence E. Blume, editors, *The New Palgrave Dictionary of Economics.* Palgrave Macmillan, Basingstoke, 2008.
8. H. Baayen. *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge University Press, 2008.
9. Michael Franke. *Signal to Act: game theory in Pragmatics.* Institute for Logic, Language and Computation, 2009.
10. R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2012. ISBN 3-900051-07-0.