MIT Sloan School of Management

Working Paper 4296-03
March 2003

THE DIGITIZATION OF WORD-OF-MOUTH:
PROMISE AND CHALLENGES OF
ONLINE FEEDBACK MECHANISMS

Chrysanthos Dellarocas

THE DIGITIZATION OF WORD-OF-MOUTH: PROMISE AND CHALLENGES OF ONLINE FEEDBACK MECHANISMS

Chrysanthos Dellarocas

Sloan School of Management

Massachusetts Institute of Technology

Cambridge, MA 02139

dell@mit.edu

ABSTRACT[1]:

Online feedback mechanisms harness the bi-directional communication capabilities of the Internet in order to engineer large-scale word-of-mouth networks. Best known so far as a technology for building trust and fostering cooperation in online marketplaces, such as eBay, these mechanisms are poised to have a much wider impact on organizations. Their growing popularity has potentially important implications for a wide range of management activities, such as brand building, customer acquisition and retention, product development, and quality assurance. This paper surveys our progress in understanding the new possibilities and challenges that these mechanisms represent. It discusses some important dimensions in which Internet-based feedback mechanisms differ from traditional word-of-mouth networks and surveys the most important issues related to their design, evaluation, and use. It provides an overview of relevant work in game theory and economics on the topic of reputation. It discusses how this body of work is being extended and combined with insights from computer science, management science, sociology, and psychology in order to take into consideration the special properties of online environments. Finally, it identifies opportunities that this new area presents for OR/MS research.

---

1 INTRODUCTION

One of the most important new capabilities of the Internet relative to previous mass communication technologies is its bi-directionality. Through the Internet, not only can organizations reach audiences of unprecedented scale at a low cost, but also, for the first time in human history, individuals can make their personal thoughts, reactions, and opinions easily accessible to the global community of Internet users.

Word-of-mouth, one of the most ancient mechanisms in the history of human society, is being given new significance by this unique property of the Internet. *Online feedback mechanisms*, also known as *reputation systems* (Resnick et al., 2000), are using the Internet's bi-directional communication capabilities in order to artificially engineer large-scale word-of-mouth networks in which individuals share opinions and experiences on a wide range of topics, including companies, products, services, and even world events. Table 1 lists several noteworthy examples of such mechanisms in use today.

Perhaps the best-known application of online feedback mechanisms to date has been their use as a technology for building trust in electronic markets. This has been motivated by the fact that many traditional trust-building mechanisms, such as state-enforced contractual guarantees, tend to be less effective in large-scale online environments (Kollock 1999). Online feedback mechanisms have emerged as a viable mechanism for fostering cooperation among strangers in such settings by ensuring that the behavior of a trader towards any other trader becomes publicly known and may therefore affect the behavior of the entire community towards that trader in the future. Knowing this, traders have an incentive to behave well towards each other, even if their relationship is a one-time deal. As I discuss in Section 3, a growing body of empirical evidence seems to demonstrate that these systems have managed to provide remarkable stability in otherwise very risky trading environments.

Web Site	Category	Summary of feedback mechanism	Format of solicited feedback	Format of published feedback
BBC	World news	BBC's Talking Point: Readers forum on world events	Readers send their opinions on specific topics selected by Talking Forum staff in free text format; readers propose possible topics of interest	Selective list of (potentially edited) opinions submitted by readers; no quantitative statistics provided
Citysearch	Entertainment guide	Users rate restaurants, bars, clubs, hotels and shops.	Users rate multiple aspects of reviewed items from 1-10 and answer a number of yes/no questions; readers rate reviews as "useful", "not useful", etc.	Weighted averages of ratings per aspect reflecting both user and editorial ratings; user reviews can be sorted according to "usefulness"
eBay	Online auction house	Buyers and sellers rate one another following transactions	Positive, negative or neutral rating plus short comment; ratee may post a response	Sums of positive, negative and neutral ratings received during past 6 months (see Section 3)
eLance	Professional services marketplace	Contractors rate their satisfaction with subcontractors	Numerical rating from 1-5 plus comment; ratee may post a response	Average of ratings received during past 6 months
Epinions	Online opinions forum	Users write reviews about products/services; other members rate the usefulness of reviews	Users rate multiple aspects of reviewed items from 1-5; readers rate reviews as "useful", "not useful", etc.	Averages of item ratings; % of readers who found a review "useful"
Google	Search engine	Search results are ordered based on how many sites contain links that point to them (Brin and Page, 1998)	A Web page is rated based on how many links point to it , how many links point to the pointing page, etc.	No explicit feedback scores are published; ordering acts as an implicit indicator of reputation
Slashdot	Online discussion board	Postings are prioritized or filtered according to the ratings they receive from readers	Readers rate posted comments	

Table 1. Some noteworthy examples of online feedback mechanisms (in use as of March 2003).

The application of feedback mechanisms in online marketplaces is particularly interesting because many of these marketplaces would probably not have come into existence without them. It is, however, by no means the only possible application domain of such systems. Internet-based feedback mechanisms are appearing in a surprising variety of settings. For example, Epinions.com encourages Internet users to rate practically any kind of brick-and-mortar business, such as airlines, telephone companies, resorts, etc. Moviefone.com solicits and displays user feedback on new movies alongside professional reviews, and Citysearch.com does the same for restaurants, bars and performances. Even news sites, perhaps the best embodiment of the unidirectional mass media of the previous century, are now soliciting and publishing reader feedback on world events alongside professionally written news articles (see, for example, BBC's Talking Point web forum.)

The proliferation of online feedback mechanisms is already changing people's behavior in subtle but important ways. Anecdotal evidence suggests that people now increasingly rely on opinions posted on such systems in order to make a variety of decisions ranging from what movie to watch to what stocks to invest on (Guernsey, 2000). Only five years ago the same people would primarily base those decisions on advertisements or professional advice.

Such trends have important repercussions for OR/MS. Managers of today's networked organizations need to understand how the growing popularity of large-scale online feedback mechanisms affects a wide range of activities within their organizations. Some examples include:

- *Brand building and customer acquisition.* Online feedback mechanisms can serve as a low-cost and, potentially, very effective channel for acquiring and retaining customers, complementary to advertising (Mayzlin, 2003). At the same time they quickly disseminate bad news that can potentially harm brand equity.

- *Product development and quality control.* Online feedback networks can assist an organization better understand consumer reactions to its current product line. At the same time, they reveal this information to competitors and they also accelerate the dissemination of information about product defects.

- *Supply chain quality assurance.* Industry-wide feedback mechanisms can assist organizations better assess prospective first-time suppliers; they can also act as a powerful disciplining mechanism that ensures fulfillment of contractual obligations and can potentially lower the legal costs of doing business.

There is currently very little work studying these and other related possibilities. The rising importance of online feedback mechanisms, thus, not only invites, but also necessitates rigorous OR/MS research on their functioning and consequences. How do such mechanisms affect the behavior of participants in the communities where they are introduced? To what extent can their operators and participants manipulate them? How can communities protect themselves from such potential abuse? What mechanism designs work best in what settings? This is just a small subset of questions that invite exciting and valuable research.

The traditional tools of OR/MS can be of value in tackling these questions. In common with other Internet technologies, however, online feedback mechanisms intensify the interdependencies between organizations, their customers, their partners, and their competitors. Managers will therefore find that proper decision-making related to the implementation and use of feedback mechanisms requires careful consideration, not only of their own actions, but also of the likely responses of other players interconnected through them. Accordingly, the tools of game theory play a prominent role in the study of these mechanisms[2].

This paper surveys our progress so far in understanding the new possibilities and challenges that these mechanisms represent. Section 2 discusses some important dimensions in which Internet-based feedback mechanisms differ from traditional word-of-mouth networks. Section 3 presents an overview of eBay's feedback mechanism, perhaps the best-studied online feedback mechanism to date. It summarizes initial field evidence on the mechanism's properties and formulates the most important open questions relating to designing, evaluating and using such mechanisms. The next two sections survey our progress in developing a systematic discipline that can help answer those questions. First, Section 4 provides an overview of relevant past work

---

[2] OR/MS is not alone in realizing that the higher degree of organizational interdependence brought about by the Internet increases the need to incorporate game theoretic concepts and techniques in system design methodologies. Papadimitriou (2001) provides an insightful discussion of how the properties of the Internet have generated substantial interest among computer scientists in incorporating game theory into algorithm and computer system design.

in game theory and economics. Section 5 then discusses how this stylized body of work is being extended in order to take into consideration the special properties of online environments. Finally, Section 6 summarizes the main points of the paper and discusses the opportunities that this new area presents for OR/MS research.

2   ONLINE FEEDBACK MECHANISMS: AN ANCIENT CONCEPT IN A MODERN SETTING

Word-of-mouth networks constitute an ancient solution to a timeless problem of social organization: the elicitation of good conduct in communities of self-interested individuals who have short-term incentives to cheat one another. The historical appeal of these networks has been their power to induce cooperation without the need for costly enforcement institutions. Before the establishment of formal law and centralized systems of contract enforcement backed by the sovereign power of a state, most ancient and medieval communities relied on word-of-mouth as the primary enabler of economic and social activity (Benson, 1989; Greif, 1993; Milgrom, North and Weingast, 1990). Many aspects of social and economic life still do so today (Klein, 1997).

What makes online feedback mechanisms different from the word-of-mouth networks of the past is the combination of (a) their unprecedented scale, achieved through the exploitation of the Internet's low-cost, bi-directional communication capabilities, (b) the ability of their designers to precisely control and monitor their operation through the introduction of automated feedback mediators, and (c) new challenges introduced by the unique properties of online interaction, such as the volatile nature of online identities and the almost complete absence of contextual cues that would facilitate the interpretation of what is, essentially, subjective information.

Scale enables new applications. Scale is essential to the effectiveness of word-of-mouth networks. In an online marketplace, for example, sellers care about buyer feedback primarily to the extent that they believe that it might affect their future profits; this can only happen if feedback is provided by a sufficient number of current customers and communicated to a significant portion of future prospects. Theory predicts that a minimum degree of participation in word-of-mouth communities is required before reputation effects can induce any cooperation. Once this threshold is reached, however, the power of reputation immediately springs to life and high levels of cooperation emerge in a discontinuous fashion (Bakos and Dellarocas, 2002).

6

Therefore, the vastly increased scale of Internet-based feedback mechanisms is likely to render them powerful institutions in environments where traditional word-of-mouth networks were heretofore considered ineffective devices. The social, economic and perhaps even political consequences of such a trend deserve careful study.

Information technology enables systematic design. In pre-Internet societies word of mouth emerged naturally and evolved in ways that were difficult to control or model. The Internet allows this powerful social force to be precisely measured and controlled through proper engineering of the information systems that mediate online feedback communities. Such automated *feedback mediators* specify who can participate, what type of information is solicited from participants, how it is aggregated and what type of information is made available to them about other community members. Through the proper design of these mediators, mechanism designers can exercise precise control over a number of parameters that are very difficult or impossible to influence in brick-and-mortar settings. For example, feedback mediators can replace detailed feedback histories with a wide variety of summary statistics; they can apply filtering algorithms to eliminate outlier or suspect ratings; they can weight ratings according to some measure of the rater's trustworthiness, etc. Such degree of control can impact the resulting social outcomes in non-trivial ways (see Sections 5.2-5.4). Through the use of information technology, what had traditionally fallen within the realm of the social sciences is, to a large extent, being transformed to an engineering design problem. Understanding the full space of design possibilities and the impacts of specific design choices on the resulting social outcomes is an important research challenge introduced by these new systems.

Online interaction introduces new challenges. The disembodied nature of online environments introduces several challenges related to the interpretation and use of online feedback. Some of these challenges have their roots in the subjective nature of feedback information. Brick-and-mortar settings usually provide a wealth of contextual cues that assist in the proper interpretation of opinions and gossip (such as familiarity with the person who acts as the source of that information, the ability to draw inferences from the source's facial expression or mode of dress, etc.). Most of these cues are absent from online settings. Readers of online feedback are thus faced with the task of evaluating the opinions of complete strangers. Other challenges to feedback interpretation have their root in the ease with which online identities can be changed.

This opens the door to various forms of strategic manipulation. For example, community members can build a good reputation, milk it by cheating other members and then disappear and reappear under a new online identity and a clean record (Friedman and Resnick, 2001). They can use fake online identities to post dishonest feedback and thus try to inflate their reputation or tarnish that of their competitors (Dellarocas, 2000). Finally, the mediated nature of online feedback mechanisms raises questions related to the trustworthiness of their operators. An important prerequisite for the widespread acceptance of online feedback mechanisms is, therefore, a better understanding of how such systems can be compromised, as well as the development of adequate defenses.

3   A CASE STUDY: EBAY'S FEEDBACK MECHANISM

eBay's feedback mechanism is, arguably, the best-studied online feedback mechanism to date. This section summarizes initial field evidence on the mechanism's properties and motivates the need for a systematic discipline of online feedback mechanism design and evaluation.

Founded in September 1995, eBay is the leading online marketplace for the sale of goods and services by a diverse community of individuals and businesses. At the beginning of 2003 the eBay community numbered 49.7 million registered users, and was the most popular shopping site on the Internet when measured by total user minutes[3].

One of the most remarkable aspects of eBay is that the transactions performed through it are not backed up by formal contractual guarantees. Instead, cooperation and trust are primarily based on the existence of a simple feedback mechanism. This mechanism allows eBay buyers and sellers to rate one another following transactions and makes the history of a trader's past ratings public to the entire community. For an overview of eBay's feedback mechanism the reader is referred to (Resnick and Zeckhauser, 2002).

Summary of empirical evidence

eBay's impressive commercial success seems to indicate that its feedback mechanism has succeeded in achieving its primary objective: generate sufficient trust among buyers to persuade

---

[3] Source: MediaMetrix.

them to assume the risk of transacting with complete strangers (Ba and Pavlou, 2002). Since *sufficiently* does not necessarily mean *efficiently*, eBay's success has generated substantial interest in better understanding how its feedback mechanism works, how much it contributes to its success, and how its success can be replicated in other environments.

A first set of results comes from empirical studies. Even a surface analysis of a representative eBay data set can uncover some interesting properties (Resnick and Zeckhauser, 2002):

- Most trading relationships are one-time deals: 89% of all buyer-seller pairs conducted just one transaction during the five-month period covered by the data set.

- Buyers left feedback on sellers 52.1% of the time; sellers on buyers 60.6% of the time.

- Feedback is overwhelmingly positive; of feedback provided by buyers, 99.1% of comments were positive, 0.6% were negative and 0.3% were neutral.

A number of studies have delved deeper into eBay data sets in order to uncover additional properties. Resnick, Zeckhauser, Swanson and Lockwood (2002) provide a comprehensive survey and methodological critique of these works. The author is aware of fourteen such studies, as summarized in Tables 2 and 3. All follow a similar logic, though the details vary in important ways. Apart from one laboratory experiment, each is an observational study of a particular category of items.

Shorthand	Citation	Items sold	Remarks
BP	Ba and Pavlou, 2002	Music, Software, Electronics	Positive feedback increased estimated price, but negative feedback did not have an effect
BH	Bajari and Hortacsu, 2003	Coins	Both positive and negative feedback affect probability of modeled buyer entry into the auction, but only positive feedback had a significant effect on final price
DH	Dewan and Hsu, 2001	Stamps	Higher net score increases price
E	Eaton, 2002	Electric guitars	Negative feedback reduces probability of sale, but not price of sold items
HW	Houser and Wooders, 2000	Pentium chips	Positive feedback increases price; negative feedback reduces it
KM	Kalyanam and McIntyre, 2001	Palm Pilot PDAs	Positive feedback increases price; negative feedback reduces price
KW	Kauffman and Wood, 2000	Coins	No significant effects, but negative feedback seems to increase price (!) in univariate analysis
LIL	Lee, Im and Lee, 2000	Computer monitors and printers	Negative feedback reduces price, but only for used items
L	Livingston, 2002	Golf clubs	Positive feedback increases both likelihood of sale and price; effect tapers off once a record is established
LBPD	Lucking-Reiley et al., 2000	Coins	No effect from positive feedback; negative feedback reduces price
MA	Melnik and Alm, 2002	Gold coins	Positive feedback increases price; negative feedback decreases price
MS	McDonald and Slawson, 2002	Dolls	Higher net score (positives -negatives) increases price
RZ	Resnick and Zeckhauser, 2002	MP3 players, Beanie babies	Both forms of feedback affect probability of sale but not price contingent on sale
RZSL	Resnick Zeckhauser, Swanson and Lockwood, 2002	Vintage postcards	Controlled field experiment; established seller commands higher prices than newcomers; among newcomers, small amounts of negative feedback have little effect

Table 2 Summary of principal results [4]

---

[4] Adapted and expanded from (Resnick, Zeckhauser, Swanson and Lockwood, 2002). The following papers are cited on Table 2 but are not yet published in a refereed journal, conference proceedings, or official working paper series.
- Dewan, S. and Hsu, V. (2001) Trust in Electronic Markets: Price Discovery in Generalist Versus Specialty Online Auctions. Working paper (available at: http://databases.si.umich.edu/reputations/bib/papers/Dewan&Hsu.doc.)
- Livingston, J. (2002) How Valuable is a Good Reputation? A Sample Selection Model of Internet Auctions. Working Paper, University of Maryland (available at: http://www.wam.umd.edu/~kth/reputation1.pdf.)
- Lucking-Reiley, D., D. Bryan, et al. (2000) Pennies from eBay: the Determinants of Price in Online Auctions, Working Paper, University of Arizona (available at: http://eller.arizona.edu/~reiley/papers/PenniesFromEBay.html.)
- Resnick, P., Zeckhauser, R., Swanson, J. and Lockwood, K. (2002) The Value of Reputation on eBay: A Controlled Experiment. Working Paper (available at: http://www.si.umich.edu/~presnick/papers/postcards/index.html.)

Question considered	Studies
How does a seller's feedback profile affect prices?	All
How does a seller's feedback profile affect the probability of sale?	BH, E, L, RZ
Does feedback matter more for riskier transactions/more expensive products?	BP, LIL
How do prices on eBay compare to prices in a more conventional channel?	DH, KW
What components of eBay's feedback profile better explain buyer behavior?	DH

Table 3 Summary of principal questions

The following points summarize the principal conclusions derived from a collective reading of these works:

- Feedback profiles seem to affect both prices and the probability of sale. However, the precise effects are ambiguous; different studies focus on different components of eBay's complex feedback profile and often reach different conclusions.

- The impact of feedback profiles on prices and probability of sale is relatively higher for riskier transactions and more expensive products.

- Among all different pieces of feedback information that eBay publishes for a member, the components that seem to be most influential in affecting buyer behavior are the overall number of positive and negative ratings, followed by the number of recently (last 7 days, last month) posted negative comments.

Towards a systematic discipline of feedback mechanism design

The initial evidence provided by empirical studies, though useful, does not help answer the most important underlying question: how well does eBay's mechanism work? In fact, these studies raise a whole new set of interesting questions. For example, why is the fraction of negative feedback so low? Is this an indication of the mechanism's poor functioning (buyers are reluctant to express their true opinions fearing retaliatory bad ratings from sellers), or perhaps a consequence of the mechanism's success (sellers are induced to behave well and therefore, there

are simply very few dissatisfied buyers)? Why is the relationship between feedback and prices ambiguous? Is this an indication that the mechanism is not well designed, or perhaps that many users do not yet understand how to best process the information it provides?

In the author's opinion, the two most concrete evaluation criteria of a feedback mechanism's performance ought to be (a) the expected payoffs of the outcomes induced by the mechanism for the various classes of stakeholders over the entire time horizon that matters for each of them and (b) the robustness of those outcomes against different assumptions about the participants' behavior[5].

Calculation of payoffs requires an understanding of how eBay's mechanism affects the behavior of buyers and sellers and how these behaviors evolve over time if all players are simultaneously pursuing their own interests. The tools of game theory are, thus, instrumental in developing conceptual models of such systems.

Robustness considerations are especially important on eBay since the essence of feedback mechanisms relies on voluntary elicitation of behavior and this, in turn, relies on a number of assumptions about human rationality and beliefs. Two issues stand out as particularly important. First, since feedback provision is currently voluntary, the impact of incomplete or untruthful feedback needs to be better understood. Second, the vulnerability of the system against strategic manipulation and online identity changes must be carefully studied.

Once we have sufficiently understood the properties and performance of eBay's current mechanism, the next obvious question is: how can this mechanism be improved? Answering this question requires a better understanding of the unique design possibilities of online feedback mechanisms. Here are a few examples of a much larger set of possibilities:

- Online feedback mechanisms can precisely control the form of information they solicit: eBay asks users to rate transactions as *positive, negative*, or *neutral*. Would it have been better to ask users to rate transactions on a scale from 1-5 (which is what Amazon does)? Could a question with different phrasing lead to even higher efficiency?

---

[5] Other plausible, but currently less well understood evaluation criteria include inducing outcomes that are perceived as "fair" by the majority of players and ensuring the privacy of the participants.

- Feedback mechanisms control how information gets aggregated and what information is publicly available in feedback profiles. Currently, eBay's feedback profile is a relatively complex artifact that includes the entire history of ratings together with a number of summary statistics. Since different users pay attention to different subsets of this information, this complicates the modeling and predictions of the induced outcomes. Would it be better to hide some parts of this information (for example, the detailed feedback history?) Would some other summary statistics (e.g., the fraction of negative ratings) lead to even more efficient outcomes? Would it be desirable to implement some sort of automated filtering of ratings that fail to satisfy some criteria?

- Feedback submission is currently voluntary on eBay. Furthermore, there is currently no quality control of submitted feedback. Could eBay introduce a carefully designed system of buyer fees and rewards that elicits complete participation and truthful feedback?

A third set of questions revolves around how online feedback mechanisms compare with more established institutions for achieving similar outcomes, such as formal contracts and advertising. These comparisons are important; their outcome will help determine how wide an impact these mechanisms will ultimately have.

An objective of any discipline of design is to eventually abstract from the study of specific cases and reach some general principles and guidelines. In the case of feedback mechanisms, this objective translates to recognizing general classes of settings where feedback mechanisms may be usefully applied, identifying important families of feedback mechanism architectures, and understanding what architectures are best suited to what settings.

Finally, design involves a responsibility for detail; this creates a need to deal with complications. In the service of design, several established OR/MS paradigms, such as decision theory and simulation, as well as empirical and experimental studies, are natural complements to game theory, both for qualifying these models, as well as for adapting them to account for the complexities of the "real world" and the bounded rationality of actual human behavior[6].

---

[6] See Roth (2002) for a broad discussion of the new methodological challenges introduced by the increasing use of economics not only for analyzing markets but also for designing them.

The rest of the paper provides a survey of past work that can serve as a starting point for answering the above questions in a systematic way.

4   REPUTATION IN GAME THEORY AND ECONOMICS

Given the importance of word-of-mouth networks in human society, reputation formation has been extensively studied by economists using the tools of game theory. This body of work is perhaps the most promising foundation for developing an analytical discipline of online feedback mechanism design. This section surveys past work on this topic, emphasizing the results that are most relevant to the design of online feedback mechanisms. Section 5 then discusses how this stylized body of work is being extended to address the unique properties of online environments.

4.1   Basic Concepts

According to Wilson (1985), reputation is a concept that arises in repeated game settings when there is uncertainty about some property (the "type") of one or more players in the mind of other players. If "uninformed" players have access to the history of past stage game outcomes, reputation effects then often allow informed players to improve their long-term payoffs by gradually convincing uninformed players that they belong to the type that best suits their interests. They do this by repeatedly choosing actions that make them appear to uninformed players as if they were of the intended type, thus "acquiring a reputation" for being of that type.

The existence of some initial doubt in the mind of uninformed players regarding the type of informed players is crucial in order for reputation effects to occur. To see this, consider a repeated game between a long-run player and a sequence of short-run (one-shot) opponents. In every stage game, the long-run player can choose one of several actions but cannot credibly commit to any of those actions in advance. If there is no uncertainty about the long-run player's type[7], rational short-run players will then always play their stage-game Nash equilibrium response. Such behavior typically results in inefficient outcomes.

---

[7] In other words, if short-run players are convinced that the long-run player is a rational utility-maximizing player whose stage-game payoffs are known with certainty.

Consider, for example, the following stylized version of a repeated "online auction" game. A long-lived seller faces an infinite sequence of sets of identical one-time buyers in a marketplace where there are only two kinds of products:

- low-quality products that cost 0 to the seller and are worth 1 to the buyers, and

- high-quality products that cost 1 to the seller and are worth 3 to the buyers.

Each period the seller moves first, announcing the quality of the product he promises to buyers. Since high quality products are more profitable, the seller will always promise high quality. Buyers then compete with one another on a Vickrey auction and therefore bid amounts equal to their expected valuation of the transaction outcome. The winning bidder sends payment to the seller. The seller then has the choice of either "cooperating" (delivering a high quality good) or "cheating" (delivering a low quality good). It is easy to see that this game has a unique subgame perfect equilibrium. In equilibrium the seller always cheats (delivers low quality), buyers each bid 1, each buyer's expected payoff is zero and the seller's expected payoff is 1.

The ability to build a reputation allows the long-run player to improve his payoffs in such settings. Intuitively, a long-run player who has a track record of playing a given action (e.g. cooperate) often enough in the past acquires a reputation for doing so and is "trusted" by subsequent short-run players to do so in the future. However, why would a profit-maximizing long-term player be willing to behave in such a way and why would rational short-term players use past history as an indication of future behavior?

To explain such phenomena, Kreps, Milgrom, Roberts and Wilson (1982), Kreps and Wilson (1982), and Milgrom and Roberts (1982) introduced the notion of "commitment" types. Commitment types are long-run players who are locked into playing the same action[8]. An important subclass of commitment types are Stackelberg types: long-run players who are locked into playing the so-called Stackelberg action. The Stackelberg action is the action to which the long-run player would credibly commit if he could. In the above "online auction" example the

---

[8] Commitment types are sometimes also referred to as "irrational" types because they follow fixed, "hard-wired" strategies as opposed to "rational" profit-maximizing strategies. An alternative way to justify such players is to consider them as players with non-standard payoff structures such that the "commitment" action is their dominant strategy given their payoffs.

Stackelberg action would be to cooperate; cooperation is the action that maximizes the seller's lifetime payoffs if the seller could credibly commit to an action for the entire duration of the game[9]. Therefore, the Stackelberg type in this example corresponds to an "honest" seller who never cheats. In contrast, an "ordinary" or "strategic" type corresponds to an opportunistic seller who cheats whenever it is advantageous for him to do so.

Reputation models assume that short-run players know that commitment types exist, but are ignorant of the type of the player they face. An additional assumption is that short-run players have access to the entire history of past stage game outcomes[10]. A player's reputation at any given time then consists of the conditional posterior probabilities over that player's type, given a short-run player's prior over types and the repeated application of Bayes' rule on the history of past stage game outcomes.

In such a setting, when selecting his next move, the informed player must take into account not only his short-term payoff, but also the long-term consequences of his action based on what that action reveals about his type to other players. As long as the promised future gains due to the increased (or sustained) reputation that comes from playing the Stackelberg action offset whatever short-term incentives he might have to play otherwise, the equilibrium strategy for an "ordinary" informed player will be to try to "acquire a reputation" by masquerading as a Stackelberg type (i.e. repeatedly play the Stackelberg action with high probability).

In the "online auction" example, if the promised future gains of reputation effects are high enough[11], ordinary sellers are induced to overcome their short-term temptation to cheat and to try to acquire a reputation for honesty by repeatedly delivering high quality. Expecting this, buyers then place high bids, thus increasing the seller's long-term payoffs.

In general, reputation effects benefit the most patient player in the game: the player who has the longest time horizon (discounts future payoffs less) is usually the one who is able to reap the

---

[9] If the seller could commit to cooperation (production of high quality), buyers would then each bid 2 and the seller's expected per period payoff would be 2.

[10] The traditional justification for this assumption is that past outcomes are either publicly observable or explicitly communicated among short-run players. The emergence of online feedback mechanisms provides, of course, yet another justification (but see discussion of complications arising from the *private* observability of outcomes in such systems in Section 5.2).

[11] In this type of game this requires that the remaining horizon of the seller is long enough and that the profit margin of a single transaction is high enough relative to the discount factor.

benefits of reputation. Fudenberg and Levine (1992) show that this result holds even when players can observe only noisy signals of each other's actions, so that the game has imperfect public monitoring. They prove that, if short-run players assign positive prior probability to the long-run player being a Stackelberg type, and if that player is sufficiently patient, then an ordinary long-run player achieves an average discounted payoff close to his commitment payoff (i.e., his payoff if he could credibly commit to the Stackelberg action). In order to obtain this payoff, the ordinary player spends long periods of time choosing the Stackelberg action with high probability[12].

4.2 Reputation dynamics

In most settings where reputation phenomena arise, equilibrium strategies evolve over time as information slowly leaks out about the types of the various players. In general, the derivation of closed-form solutions in repeated games with reputation effects is complicated. Nevertheless, a small number of specific cases have been studied. They provide interesting insight into the complex behavioral dynamics introduced by reputational considerations.

Initial phase

In most cases, reputation effects begin to work immediately and in fact are strongest during the initial phase, when players must work hard to establish a reputation. Holmstrom (1999) discusses an interesting model of reputational considerations in the context of an agent's "career" concerns. Suppose that wages are a function of an employee's innate ability for a task. Employers cannot directly observe an employee's ability. However, they can keep track of the average value of her past task outputs. Outputs depend both on ability and labor. The employee's objective is to maximize her lifetime wages while minimizing the labor she has to put in. At equilibrium, this provides incentives to the employee to work hard right from the beginning of her career in order to build a reputation for competence. In fact these incentives are strongest at the very beginning of her career when observations are most informative.

---

[12] This result also requires that the stage game is either a simultaneous move game, or, in a sequential-move game, that the short-run players always observe whether or not the Stackelberg strategy has been played.

During the initial phase of a repeated game, it is common that some players realize lower or even negative profits, while the community "learns" their type. In those cases players will only attempt to build a reputation if the losses from masquerading as a Stackelberg type in the current round are offset by the present value of the gains from their improved reputation in the later part of the game. In trading environments, this condition usually translates to the need for sufficiently high profit margins for "good quality" products so that the promise of future gains from sustaining a reputation is persuasive enough to offset the short-term temptation to cheat. This was first pointed out in Klein and Leffler (1981) and explored more formally in Shapiro (1983).

Another case where reputation effects may fail to work is when short-run players are "too cautious" vis-à-vis the long-run player and therefore update their beliefs too slowly in order for the long-run player to find it profitable to try to build a reputation. Such cases may occur when, in addition to Stackelberg ("good") types, the set of commitment types also includes "bad" or "inept" types: players who always play the action that the short-run players like least. In the "online auction" example, a "bad" type corresponds to a player who always cheats (because, for example, he lacks the capabilities that would enable him to deliver high quality.) If short-run players have a substantial prior belief that the long-run player may be a "bad" type, then the structure of the game may not allow them to update their beliefs fast enough to make it worthwhile for the long-run player to try to acquire a reputation.

Diamond's (1989) analysis of reputation formation in debt markets presents an example of such a setting. In Diamond's model there are three types of borrowers: safe borrowers, who always select safe projects (i.e. projects with zero probability of default); risky borrowers, who always select risky projects (i.e. projects with higher returns if successful but with nonzero probability of default); and strategic borrowers who will select the type of project that maximizes their long term expected payoff. The objective of lenders is to maximize their long term return by offering competitive interest rates, while at the same time being able to distinguish profitable from unprofitable borrowers. Lenders do not observe a borrower's choice of projects, but they do have access to her history of defaults. In Diamond's model, if lenders believe that the initial fraction of risky borrowers is significant, then, despite the reputation mechanism, at the beginning of the game, interest rates will be so high that strategic players have an incentive to select risky projects. Some of them will default and will exit the game. Others will prove lucky and will

begin to be considered as safe players. It is only after lucky strategic players have already acquired some initial reputation (and therefore begin to receive lower interest rates) that it becomes optimal for them to begin "masquerading" as safe players by consciously choosing safe projects in order to sustain their good reputation.

Steady state (or lack thereof)

In their simplest form, reputation games are characterized by an equilibrium in which the long-run player repeatedly plays the Stackelberg action with high probability and the player's reputation converges to the Stackelberg type.

The existence of such steady states crucially depends on the ability to perfectly monitor the outcomes of individual stage games. For example, consider the "online auction" game that serves as an example throughout this section with the added assumption that buyers perfectly and truthfully observe and report the seller's action. In such a setting, the presence of even a single negative rating on a seller's feedback history reveals the fact that the seller is not honest. From then on, buyers will always choose the low bid in perpetuity. Since such an outcome is not advantageous for the seller, reputation considerations will induce the seller to cooperate forever.

The situation changes radically if monitoring of outcomes is imperfect. In the online auction example, imperfect monitoring means that even when the seller produces high quality, there is a possibility that a buyer will post a negative rating, and, conversely, even when the seller produces low quality, the buyer may post a positive rating. A striking result is that in such "noisy" environments reputations cannot be sustained indefinitely: if a strategic player stays in the game long enough, short-run players will eventually learn his true type and the game will inevitably revert to one of the static Nash equilibria (Cripps, Mailath and Samuelson, 2002).

To see the intuition behind this result, note that reputations under perfect monitoring are typically supported by a trigger strategy. Deviations from the equilibrium strategy reveal the type of the deviator and are punished by a switch to an undesirable equilibrium of the resulting complete-information continuation game. In contrast, when monitoring is imperfect, individual deviations neither completely reveal the deviator's type nor trigger punishments. A single deviation has only a small effect on the beliefs of the short-term players. As a result, a player of

normal type trying to maintain a reputation as a Stackelberg type incurs only a small cost (in terms of altered beliefs) from indulging in occasional deviations from Stackelberg play. In fact, it is clear that always playing the Stackelberg action cannot be an equilibrium strategy, because if the short-term players expect long-term players of normal type to behave that way, then they can actually deviate at no cost, since any bad outcome will be interpreted by the short-run players as a result of imperfect monitoring. But the long-run effect of many such small deviations from the commitment strategy is to drive the equilibrium to full revelation.

Holmstrom's paper provides an early special case of this striking result: the longer an employee has been on the market, the more "solid" the track record she has acquired and the less important her current actions in influencing the market's future assessment of her ability. This provides diminishing incentives for her to keep working hard. Cripps, Mailath and Samuelson's result then states that, if the employee stays on the market for a very long time, these dynamics will lead to an eventual loss of her reputation.

If one tries to reconcile the above result with Fudenberg and Levine's 1992 result on long-term payoffs induced by reputation, even more interesting phenomena come to the surface. If players eventually lose their reputation, in order for them to achieve average long-term payoffs that are close to their Stackelberg payoff, they must realize payoffs higher than their Stackelberg payoff during some stage of the game. This makes the dynamics of reputation formation in environments with imperfect monitoring quite complex indeed: an initial stage of reputation formation (with potentially suboptimal payoffs) is followed by a stage where the long-run player is able to occasionally "fool" short-run players and realize payoffs above his Stackelberg payoff, followed by a stage where short-run players eventually learn the truth and the game reverts to its static Nash equilibrium.

These dynamics have important repercussions for systems like eBay. According to the Cripps, Mailath and Samuelson result, if eBay makes the entire feedback history of a seller available to buyers and if an eBay seller stays on the system long enough, once he establishes an initial reputation for honesty he will be tempted to cheat buyers every now and then. In the long term, this behavior will lead to an eventual collapse of his reputation and therefore of cooperative

behavior. The conclusion is that, if buyers pay attention to a seller's entire feedback history, eBay's current mechanism fails to sustain long-term cooperation.

Endgame considerations

Since reputation relies on a tradeoff between current "restraint" and the promise of future gains, in finitely repeated games, incentives to maintain a reputation diminish and eventually disappear as the end of the game comes close.

One solution to this problem is to introduce community membership rules that elicit good behavior throughout the game (Ba, 2001). For example, online communities can levy a sufficiently high entrance fee that is refundable subject to maintaining a good reputation upon exit.

Another solution is to assign some post-mortem value to reputation, so that players find it optimal to maintain it throughout the game. For example, reputations can be viewed as assets that can be bought and sold in a market for reputations. Tadelis (1998) shows that a market for reputations is indeed sustainable. Furthermore, the existence of such a market provides "old" agents and "young" agents with equal incentives to exert effort (Tadelis, 2002). However, the long-run effects of introducing such a market can be quite complicated since good reputations are then likely to be purchased by "inept" agents for the purpose of depleting them (Mailath and Samuelson, 2001; Tadelis, 2002). Further research is needed in order to fully understand the long-term consequences of introducing markets for reputation as well as for transferring these promising concepts to the online domain.

5   NEW OPPORTUNITIES AND CHALLENGES OF ONLINE MECHANISMS

In Section 2, I discussed a number of differences between online feedback mechanisms and traditional word-of-mouth networks. This section surveys our progress in understanding the opportunities and challenges that these special properties imply.

5.1 Understanding the impact of scalability

Bakos and Dellarocas (2002) model the impact of information technology on online feedback mechanisms in the context of a comparison of the social efficiency of litigation and online feedback. They observe that online feedback mechanisms provide linkages between otherwise disconnected smaller markets (each having its own informal word-of-mouth networks) in which a firm operates. This, in turn, is equivalent to increasing the discount factor of the firm when considering the future impacts of its behavior on any given transaction. In trading relationships, a minimum discount factor is necessary to make reputation effects effective *at all* in inducing cooperative behavior (this is an alternative way to interpret Klein and Leffler, 1981 and Shapiro, 1983). Once this threshold is reached, however, the power of reputation springs to life in a discontinuous fashion and high levels of cooperation can be supported. Thus the vastly increased potential scale of Internet-based feedback mechanisms and the resulting ability to cover a substantial fraction of economic transactions are likely to render these mechanisms into powerful quality assurance institutions in environments where the effectiveness of traditional word-of-mouth networks has heretofore been limited. The social, economic and perhaps even political consequences of such a trend deserve careful study. For example, Bakos and Dellarocas show how, under certain conditions, sufficiently scalable feedback mechanisms can be a more socially efficient institution for inducing honest trade than the threat of litigation.

5.2 Eliciting sufficient and honest feedback

Most game theoretic models of reputation formation assume that stage game outcomes (or imperfect signals thereof) are *publicly* observed. Online feedback mechanisms, in contrast, rely on *private* monitoring of stage game outcomes and voluntary feedback submission. This introduces two important new considerations: (a) ensuring that sufficient feedback is, indeed, provided, and (b) inducing truthful reporting.

Economic theory predicts that voluntary feedback will be underprovided. There are two main reasons for this. First, feedback constitutes a public good: once available, everyone can costlessly benefit from it. Voluntary provision of feedback leads to suboptimal supply, since no individual takes account of the benefits that her provision gives to others. Second, provision of feedback presupposes that the rater will assume the risks of transacting with the ratee. Such risks are

highest for new products. Prospective consumers may, thus, be tempted to wait until more information is available. However, unless somebody decides to take the risk of becoming an early evaluator, no feedback will ever be provided.

Avery, Resnick and Zeckhauser (1999) analyze mechanisms whereby early evaluators are paid to provide information and later evaluators pay in order to balance the budget. They conclude that, of the three desirable properties for such a mechanism (voluntary participation, no price discrimination and budget balance), any two can be achieved, but not all three.

Since monitoring is private and assessments usually subjective, an additional consideration is whether feedback is honest. Miller, Resnick and Zeckhauser (2002) propose mechanisms for eliciting honest feedback based on the technique of proper scoring rules. A scoring rule is a method for inducing decision makers to reveal their true beliefs about the distribution of a random variable by rewarding them based on the actual realization of the random variable and their announced distribution. A proper scoring rule has the property that the decision maker maximizes the expected score when he truthfully announces his belief about the distribution.

Such mechanisms work as long as raters are assumed to act independently. Collusive behavior can defeat proper scoring rules. Unfortunately, online environments are particularly vulnerable to collusion. The development of effective mechanisms for dealing with collusive efforts to manipulate online ratings is currently an active area of research. Dellarocas (2000; 2001a) explores the use of robust statistics in aggregating individual ratings as a mechanism for reducing the effects of coordinated efforts to bias ratings. To this date, however, there is no effective solution that completely eliminates the problem.

5.3   Exploiting the information processing capabilities of feedback mediators

Most game theoretic models of reputation assume that short-run players have access to the entire past history of stage game outcomes and update their prior beliefs by repeated application of Bayes' rule on that information.

Online feedback mediators completely control the amount and type of information that is made available to players. This opens an entire range of new possibilities. For example, feedback mediators can hide the detailed history of past feedback from short-term players and replace it

with a summary statistic (such as the sum, mean or median of past ratings) or with any other function of the feedback history. They can filter outlying or otherwise suspect ratings. They can offer *personalized* feedback profiles, that is, present different information about the same long-run player to different short-run players.

Such information transformations can have non-trivial effects in the resulting equilibria and can often allow online feedback mechanisms to induce outcomes that are difficult or impossible to attain in standard settings. The following are two examples of what can be achieved.

As discussed in Section 4.2, in environments with imperfect monitoring, traditional reputation models predict that reputations are not sustainable. Once firms build a reputation, they are tempted to "rest on the laurels"; this behavior, ultimately, leads to a loss of their reputation. Economists have used a variety of devices to construct models that do not exhibit this undesirable behavior. For instance, Mailath and Samuelson (1998) assume that in every period there is a fixed, exogenous probability that the type of the firm might change. Horner (2002) proposes a model in which competition among firms induces them to exert sustained effort.

Online feedback mediators provide yet another, perhaps much more tangible, approach to eliminating such problems: by *designing* the mediator to only publish recent feedback, firms are given incentives to constantly exert high effort. In the context of eBay, this result argues for the elimination of the detailed feedback history from feedback profile and the use of summaries of recent ratings as the primary focal point of decision-making. Dellarocas (2003) studies the equilibria induced by a variation of eBay's feedback mechanism in which the only information available to buyers is the sum of positive and negative ratings posted on a seller during the most recent N transactions. He finds that, in trading environments with opportunistic sellers, imperfect monitoring of a seller's effort level, and two possible transaction outcomes (corresponding to "high" and "low" quality respectively), such a mechanism induces high levels of cooperation that remain stable over time. Furthermore, the long-run payoffs are independent of the size of the window N. A mechanism that only publishes the *single* most recent rating is just as efficient as a mechanism that summarizes larger numbers of ratings.

A second example of improving efficiency through proper mediator design can be found in Dellarocas (2002), which studies settings in which a monopolist sells products of various

qualities and announces the quality of each product. The objective of a feedback mechanism in such settings is to induce truthful announcements. Once again, Cripps, Mailath and Samuelson's result predicts that, in noisy environments, a mechanism that simply publishes the entire history of feedback will not lead to sustainable truth-telling. Dellarocas proposes a mechanism that acts as an intermediary between the seller and the buyers. The mechanism does not publish the history of past ratings. Instead, it keeps track of discrepancies between past seller quality announcements and corresponding buyer feedback and then punishes or rewards the seller by "distorting" the seller's subsequent quality announcements so as to compensate him for whatever "unfair" gains or losses he has realized by misrepresenting the quality of his items. If consumers are risk-averse, at equilibrium this induces the seller to truthfully announce quality throughout the infinite version of the game.

5.4 Coping with easy name changes

In online communities it is usually easy for members to disappear and re-register under a completely different online identity with zero or very low cost. Friedman and Resnick (2001) refer to this property as "cheap pseudonyms." This property hinders the effectiveness of feedback mechanisms. Community members can build a reputation, milk it by cheating other members and then vanish and re-enter the community with a new identity and a clean record.

Friedman and Resnick discuss two classes of approaches to this issue: either make it more difficult to change online identities, or structure the community in such a way that exit and re-entry with a new identity becomes unprofitable. The first approach makes use of cryptographic authentication technologies and is outside the scope of this paper. The second approach is based on imposing an upfront cost to each new entrant, such that the benefits of "milking" one's reputation are exceeded by the cost of subsequent re-entry. This cost can be an explicit entrance fee or an implicit cost of having to go through a reputation building (or *"dues paying"*) stage with low or negative profits. Friedman and Resnick show that, although dues paying approaches incur efficiency losses, such losses constitute an inevitable consequence of easy name changes.

Dellarocas (2003) shows how such a "dues paying" approach can be implemented in trading environments where feedback is binary (i.e. transactions are rated as "good" or "bad") and mediators only publish the sum of recent ratings. He proves that, in the presence of easy name

changes, the design that results in optimal social efficiency is one where the mechanism sets the initial profile of new members to correspond to the "worst" possible reputation[13]. Dellarocas further demonstrates that, although this design incurs efficiency losses relative to the case where identity changes is not an issue, in settings with two possible transaction outcomes and opportunistic sellers, if players can costlessly change their identities its efficiency is the highest possible attainable by any mechanism.

5.5   Exploring alternative architectures

The preceding discussion has assumed a centralized architecture in which feedback is explicitly provided and a single trusted mediator controls feedback aggregation and distribution. Though the design possibilities of even that simple architecture are not yet fully understood, centralized feedback mechanisms do not nearly exhaust the new possibilities offered by information technology.

In recent years the field of multi-agent systems (Jennings, Sycara, and Wooldridge, 1998) has been actively researching online feedback systems as a technology for building trust and inducing good behavior in artificial societies of software agents. Two lines of investigation stand out as particularly novel and promising:

Reputation formation based on analysis of "implicit feedback." In our networked society, several traces of an agent's activities can be found on publicly accessible databases. Instead of (or in addition to) relying on explicitly provided feedback, automated feedback mechanisms can then potentially infer aspects of an agent's attributes, social standing and past behavior through collection and analysis of such "implicit feedback" information.

Perhaps the most successful application of this approach to date is exemplified by the Google search engine. Google assigns a measure of reputation to each web page that matches the keywords of a search request. It then uses that measure to rank order search hits. Google's page reputation measure is based on the number of links that point to a page, the number of links that

---

[13] For example, if the mechanism summarizes the most recent 10 ratings, newcomers would begin the game with a profile that indicates that all 10 recent ratings were negative. An additional assumption is that buyers cannot tell how long a given seller has been on the market and therefore cannot distinguish between newcomers with "artificially tarnished" profiles and dishonest players who have genuinely accumulated many negative ratings.

point to the pointing page, and so on (Brin and Page, 1998). The underlying assumption is that if enough people consider a page to be important enough in order to place links to that page from their pages, and if the pointing pages are "reputable" themselves, then the information contained on the target page is likely to be valuable. Google's success in returning relevant results is testimony to the promise of that approach.

Pujol , Sangüesa and Delgado (2002) apply network flow techniques in order to propose a generalization of the above algorithm that "extracts" the reputation of nodes in a general class of social networks. Sabater and Sierra (2002) describe how direct experience, explicit and implicit feedback can be combined into a single reputation mechanism.

Basing reputation formation on implicit information is a promising solution to problems of eliciting sufficient and truthful feedback. Careful modeling of the benefits and limitations of this approach is needed in order to determine in what settings it might be a viable substitute or complement of voluntary feedback provision.

Decentralized feedback architectures. Our discussion of feedback mechanisms so far has implicitly assumed the honesty of feedback mediators. However, mediators are also designed and operated by parties whose interests may sometimes diverge from those of society at large.

Decentralizing the sources of feedback is a promising approach for achieving robustness in the presence of both potentially dishonest mediators and privacy concerns. A number of decentralized feedback mechanisms have recently been proposed (Zacharia, Moukas and Maes, 2000; Mui, Szolovits and Ang, 2001; Sen and Sajja, 2002; Yu and Singh, 2002). Though novel and intriguing, none of these works provides a rigorous analysis of the behavior induced by the proposed mechanisms, nor an explicit discussion of their advantages relative to other alternatives. More collaboration is needed in this promising direction between both computer scientists, who better understand the new possibilities offered by technology, and management scientists, who better understand the tools for evaluating the potential impact of these new systems.

5.6 Accounting for bounded rationality

The ambition of a discipline of online feedback mechanism design is the inducement of social outcomes with a degree of precision that approaches that of engineering design. This, in turn, requires precise modeling not only of the technological components of those systems but also of the human users.

It is now well known that human behavior does not conform to the traditional economics assumptions of rational maximization of well-defined utility functions[14]. Two recent laboratory experiments provide some initial insight into human behavior vis-à-vis feedback mechanisms. Bolton, Katok and Ockenfels (2002) compare trading in a market with (automatically generated) feedback to a market without, as well as to a market in which the same people interact with each other repeatedly (partners market). They find that, while the feedback mechanism induces a substantial improvement in trading efficiency, it falls short of the efficiency achieved in the partners market. Keser (2002) reports the results of a repeated trust game among strangers with and without the ability to provide feedback. She finds that the presence of a feedback mechanism significantly increases the levels of trust and trustworthiness. Furthermore, efficiency is slightly higher if trading partners are informed of the entire distribution of each other's previous ratings than if they are informed of each other's most recent rating only.

6 CONCLUSIONS: OPPORTUNITIES FOR OR/MS RESEARCH

Online feedback mechanisms harness the remarkable ability of the Web to not only disseminate, but also collect and aggregate information from large communities at very low cost, in order to artificially construct large-scale word-of-mouth networks. Best known so far as a technology for building trust and fostering cooperation in online marketplaces, these mechanisms are poised to have a much wider impact on organizations. As discussed in the Introduction, their popularity has potentially important implications for a wide range of management activities, such as customer acquisition and retention, brand building, product development, and quality assurance. The study of such topics clearly falls within the domain of OR/MS.

---

[14] For an excellent survey of psychological findings relevant to economics the reader is strongly encouraged to consult (Rabin, 1996).

The design of online feedback mechanisms can greatly benefit from the insights produced by more than twenty years of economics and game theory research on the topic of reputation. These results need to be extended to take into account the unique new properties of online environments, such as their unprecedented scalability, the ability to precisely design the type of feedback information that is solicited and distributed, and the volatility of online identities. The following list contains what the author considers to be the most important open areas of research in feedback mechanism design:

- Scope and explore the design space and limitations of online feedback mechanisms. Understand what set of design parameters work best in what settings. Develop models and prototype implementations of such systems.

- Develop effective solutions to the problems of sufficient participation, easy identity changes and strategic manipulation of online feedback.

- Conduct theory-driven experimental and empirical research that sheds more light on buyer and seller behavior vis-à-vis such mechanisms.

- Compare the relative efficiency of feedback mechanisms to the efficiency of more established mechanisms for dealing with information asymmetries (such as state-backed contractual guarantees and advertising); develop theory-driven guidelines for deciding which set of mechanisms to use when.

- Understand how decision-makers must adapt their strategies to react to the presence of such mechanisms in areas such as marketing, product development, and customer service.

The power of online feedback mechanisms has its roots in the strategic side effects brought about by the increased interdependencies they create among firms, their customers, their partners, and their competitors. As the mathematical study of interaction of self-interested agents, game theory is the natural foundation for the study of online feedback mechanisms. Game theoretic analyses, however, are often based on stylized abstractions of real systems and are fundamentally grounded on the assumption of rational, utility maximizing economic agents.

The traditional paradigms of OR/MS can play an important role in translating the conceptual insights of game theoretic models into concrete guidelines for building (and reacting to) large-scale feedback mechanisms that can influence the dynamics of entire industries or societies. For example, computational methods can help analyze games that may be too complex to solve analytically. Laboratory experiments can inform about how people will behave when confronted with these mechanisms, both when they are inexperienced and as they gain experience. Finally, game-theoretic models can often be approximated by generally more tractable, decision-theoretic approaches[15].

Large-scale online feedback mechanisms have the potential to influence the organizational dynamics of the 21st century in a similarly powerful way in which the ability to mass broadcast affected business and society in the 20th century. OR/MS has a lot to offer in helping organizations understand how these mechanisms can improve efficiency in a variety of practical settings, as well as how companies should adapt their strategies to react to their presence.

References

Avery, C., Resnick, P. and Zeckhauser, R. (1999) The Market for Evaluations. *American Economics Review*, 89 (3), pp. 564-584.

Ba, S. (2001) Establishing online trust through a community responsibility system. *Decision Support Systems*, 31, pp. 323-336.

Ba, S. and Pavlou, P. (2002). Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior. *MIS Quarterly*, 26 (3).

Bajari, P. and Hortacsu, A. (2003). Winner's Curse, Reserve prices and Endogenous Entry: Empirical Insights from eBay Auctions. *Rand Journal of Economics* (forthcoming).

Bakos, Y. and Dellarocas, C. (2002) Cooperation without Enforcement? A Comparative Analysis of Litigation and Online Reputation as Quality Assurance Mechanisms. *Proceedings of the 23rd International Conference on Information Systems* (ICIS 2002), Barcelona, Spain.

---

[15] See Zacharia et al. (2001) for an example of using simulation modeling to study the effects of feedback mechanisms on markets with dynamic pricing. See Dellarocas (2001b) and Shapiro (1982) for examples of how the adoption of "reasonable" (common-sense or empirically-driven) assumptions about the behavior of other players permits the use of decision-theoretic approaches to model reputation phenomena.

Benson, B. (1989). The Spontaneous Evolution of Commercial Law. *Southern Economic Journal* 55, pp. 644-61. Reprinted in *Reputation: Studies in the Voluntary Elicitation of Good Conduct*, edited by Daniel B. Klein, pp. 165-89. Ann Arbor: University of Michigan Press, 1997.

Bolton, G. E., Katok, E. and Ockenfels, A. (2002) How Effective are Online Reputation Mechanisms? An Experimental Investigation. Discussion Paper 25-2002, Max Planck Institute for Research into Economic Systems, Jena, Germany (available at: ftp://ftp.mpiew-jena.mpg.de/pub/esi/discussionpapers/2002-25.pdf)

Brin S. and Page L. (1998). The Anatomy of a Large-Scale Hypertextual Web Search Engine. *Computer Networks and ISDN Systems*, 30 (1-7), pp. 107-117.

Cripps, M., Mailath, G. and Samuelson., L. (2002) Imperfect Monitoring and Impermanent Reputations, Penn Institute for Economic Research Working Paper 02-021 (available at: http://www.econ.upenn.edu/Centers/pier/Archive/02-021.pdf)

Dellarocas, C. (2000) Immunizing online reputation reporting systems against unfair ratings and discriminatory behavior. *Proceedings of the 2nd ACM Conference on Electronic Commerce*, Minneapolis, MN, October 17-20, 2000, pp. 150-157.

Dellarocas. C. (2001a) Building Trust On-Line: The Design of Reliable Reputation Mechanisms for Online Trading Communities. MIT Sloan Working Paper No. 4180-01 (available at: http://ebusiness.mit.edu/research/papers/101%20Dellarocas,%20Trust%20Management.pdf)

Dellarocas, C. (2001b) Analyzing the economic efficiency of eBay-like online reputation reporting mechanisms. *Proceedings of the 3rd ACM Conference on Electronic Commerce*, Tampa, FL, October 14-16, 2001, pp. 171-179.

Dellarocas C. (2002) Goodwill Hunting: An economically efficient online feedback mechanism in environments with variable product quality. In *Agent-Mediated Electronic Commerce IV. Designing Mechanisms and Systems,* Padget, J., Shehory, O., Parkes, D., Sadeh, N., and Walsh, W. E. (Eds.) *Lecture Notes in Computer Science* 2351, Springer Verlag, pp. 238-252.

Dellarocas C. (2003) Efficiency and Robustness of Binary Feedback Mechanisms in Trading Environments with Moral Hazard. MIT Center for eBusiness Working Paper #170 (available at: http://ebusiness.mit.edu/research/papers/170_Dellarocas_Moral_Hazard.pdf)

Diamond, D. (1989). Reputation Acquisition in Debt Markets. *Journal of Political Economy* 97, 4, pp. 828-862.

Eaton, D. H. (2002). Valuing Information: Evidence from Guitar Auctions on eBay. Working paper, Murray State University, Department of Economics Working Paper #0201 (available at: http://campus.murraystate.edu/academic/faculty/david.eaton/workpaper0201.pdf)

Friedman, E. and Resnick, P. (2001). The Social Cost of Cheap Pseudonyms. *Journal of Economics and Management Strategy* 10 (1), pp. 173-199.

Fudenberg, D., and Levine, D. (1992). Maintaining a Reputation when Strategies are Imperfectly Observed, *Review of Economic Studies* 59, pp. 561–579.

Greif, A. (1993). Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition. *American Economic Review* 83, pp. 525-548.

Guernsey, L. (2000). Suddenly, Everybody's an Expert on Everything. *The New York Times*, February 3, 2000.

Holmstrom, B. (1999). Managerial Incentive Problems: A Dynamic Perspective. *Review of Economic Studies*, 66 (1), pp. 169-182.

Horner, J. (2002) Reputation and Competition, *American Economic Review* 92, pp. 644-663.

Houser, D. and Wooders, J. (2000) Reputation in Auctions: Theory and Evidence from eBay. Working Paper, The University of Arizona Department of Economics Working Paper 00-01 (available at: http://info-center.ccit.arizona.edu/~econ/working_papers/Internet_Auctions.pdf)

Jennings, N., Sycara, K., and Wooldridge, M. (1998) A Roadmap of Agent Research and Development. *Autonomous Agents and Multi-Agent Systems*, 1, 275–306.

Kalyanam, K. and S. McIntyre (2001). Returns to Reputation in Online Auction markets. Santa Clara University Retail Workbench Working Paper W-RW01-02.

Kauffman, R. J., and C. Wood (2000). Running up the Bid: Modeling Seller Opportunism in Internet Auctions, *Americas Conference on Information Systems* (AMCIS), Long Beach, CA, August 2000.

Keser, C. (2002) Trust and Reputation Building in E-Commerce. Scientific Publication 2002s-75, Center for Interuniversity Research and Analysis on Organizations (CIRANO), Montreal, Canada (available at: http://www.cirano.qc.ca/pdf/publication/2002s-75.pdf)

Klein, B. and Leffler, K. (1981) The Role of Market Forces in Assuring Contractual Performance. *Journal of Political Economy*. 89 (4), pp. 615-641.

Klein, D. ed. (1997) *Reputation: Studies in the Voluntary Elicitation of Good Conduct.* University of Michigan Press, Ann Arbor, MI.

Kollock, P. (1999) The Production of Trust in Online Markets. In *Advances in Group Processes* (Vol. 16), eds. E.J. Lawler, M. Macy, S. Thyne, and H.A. Walker, Greenwich, CT: JAI Press.

Kreps, D., Milgrom, P., Roberts, J and Wilson, R. (1982). Rational Cooperation in the Finitely Repeated Prisoners' Dilemma. *Journal of Economic Theory*, 27, pp. 245-52.

Kreps, D. and Wilson, R. (1982). Reputation and Imperfect Information. *Journal of Economic Theory*, 27, pp. 253-79.

Lee, Z., I. Im, et al. (2000). The Effect of Negative Buyer Feedback on Prices in Internet Auction Markets. *Proceedings of the 21st International Conference on Information Systems*, Brisbane, Australia.

Mailath, G. J., and Samuelson, L. (1998) Your Reputation Is Who You're Not, Not Who You'd Like to Be, Center for Analytic Research in Economics and the Social Sciences (CARESS) Working Paper 98-11, University of Pennsylvania (available at: http://www.ssc.upenn.edu/~gmailath/wpapers/rep-is-sep.html)

Mailath, G. J., and Samuelson, L. (2001) Who Wants a Good Reputation? *Review of Economic Studies,* 68, pp. 415-441.

Mayzlin, Dina (2003) Promotional Chat on the Internet. Yale School of Management Working Paper #MK-14, January 31, 2003.

McDonald, C. G. and V. C. Slawson, Jr. (2002). Reputation in an Internet Auction Market, *Economic Inquiry*, 40 (3), October 2002, pp. 633-650.

Melnik, M. I. and J. Alm (2002). Does a Seller's Reputation Matter? Evidence from eBay Auctions. *Journal of Industrial Economics*, September 2002, 50 (3), pp. 337-349.

Milgrom, P. R., North, D., and B. R. Weingast. (1990). The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs. *Economics and Politics* 2, pp. 1-23.

Milgrom, P., and Roberts, J. (1982). Predation, Reputation and Entry Deterrence, *Journal of Economic Theory* 27, pp. 280–312.

Miller, N., Resnick, P., and Zeckhauser, R. (2002) Eliciting Honest Feedback in Electronic Markets, Harvard Kennedy School Research Working Paper RWP02-039 (available at: http://www.si.umich.edu/~presnick/papers/elicit/index.html)

Mui, L., Szolovits, P. and Ang, C. (2001) Collaborative sanctioning: applications in restaurant recommendations based on reputation. *Proceedings of the fifth international conference on Autonomous agents*, Montreal, Quebec, Canada, pp. 118-119.

Papadimitriou, C. (2001) Algorithms, Games, and the Internet. *Proceedings of the thirty-third annual ACM symposium on Theory of computing*, July 2001, Hersonissos, Greece, pp. 749-753.

Pujol, J. M., Sanguesa, R. and Delgado, J. (2002) Extracting reputation in multi agent systems by means of social network topology. *Proceedings of the first international joint conference on Autonomous agents and multiagent systems*, Bologna, Italy, pp. 467-474.

Rabin, M. (1998) Psychology and Economics. *Journal of Economic Literature*, 36, pp. 11-46.

Resnick, P., Zeckhauser, R., Friedman, E., Kuwabara, K. (2000) Reputation Systems. *Communications of the ACM*, 43 (12), pp. 45-48.

Resnick, P., and Zeckhauser, R. (2002) Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System. In *The Economics of the Internet and E-Commerce*. Michael R. Baye, editor. JAI Press.

Roth, A. (2002) The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics. *Econometrica,* 70 (4), pp. 1341-1378

Sabater, J. and Sierra, C. (2002) Reputation and social network analysis in multi-agent systems. *Proceedings of the first international joint conference on Autonomous Agents and Multiagent Systems* (AAMAS), Bologna, Italy, pp. 475-482.

Sen, S. and Sajja, N. (2002) Robustness of reputation-based trust: boolean case. *Proceedings of the first international joint conference on Autonomous Agents and Multiagent Systems* (AAMAS), Bologna, Italy, pp. 288-293.

Shapiro, C. (1982) Consumer Information, Product Quality, and Seller Reputation. *Bell Journal of Economics* 13 (1), pp 20-35.

Shapiro, C. (1983). Premiums for High Quality Products as Returns to Reputations. *The Quarterly Journal of Economics*, 98 (4), pp. 659-680.

Tadelis, S. (1999). What's in a Name? Reputation as a Tradeable Asset. *The American Economic Review,* 89 (3), pp. 548-563.

Tadelis, S. (2002). The Market for Reputations as an Incentive Mechanism. *Journal of Political Economy*, 92 (2), pp. 854-882

Wilson, R. (1985). Reputations in Games and Markets. In *Game-Theoretic Models of Bargaining*, edited by Alvin Roth, Cambridge University Press, pp. 27-62.

Yu, B. and Singh, M. (2002) An evidential model of distributed reputation management. *Proceedings of the first international joint conference on Autonomous Agents and Multiagent Systems* (AAMAS), Bologna, Italy, pp. 294-301.

Zacharia, G., Moukas, A. and Maes, P. (2000) Collaborative Reputation Mechanisms in Electronic Marketplaces. *Decision Support Systems*, 29 (4), pp. 371-388.

Zacharia, G., Evgeniou, T., Moukas, A., Boufounos, P., Maes, P. (2001) Economics of Dynamic Pricing in a Reputation Brokered Agent Mediated Marketplace. *Electronic Commerce Research* 1(1), pp. 85-100.