Center for eBusiness@MIT
http://ebusiness.mit.edu

A research and education initiative at the MIT Sloan School of Management

THE DIGITIZATION OF WORD-OF-MOUTH:

PROMISE AND CHALLENGES OF ONLINE FEEDBACK MECHANISMS

Paper 173

Chrysanthos Dellarocas

October 2003

For more information,
please visit our website at **http://ebusiness.mit.edu**
or contact the Center directly at **ebusiness@mit.edu or 617-253-7054**

MITSloan

The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms

Chrysanthos Dellarocas
Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
dell@mit.edu

Online feedback mechanisms harness the bidirectional communication capabilities of the Internet 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 to take into consideration the special properties of online environments. Finally, it identifies opportunities that this new area presents for operations research/management science (OR/MS) research.
(*Reputation Mechanisms; Online Feedback; Electronic Markets; Trust; Internet; Game Theory*)

1. Introduction

One of the most important capabilities of the Internet relative to previous mass communication technologies is its bidirectionality. 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 bidirectional communication capabilities 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 toward any other trader becomes

0025-1909/03/4910/1407
1526-5501 electronic ISSN

Table 1 Some Noteworthy Examples of Online Feedback Mechanisms (in Use as of March 2003)

Website	Category	Summary of feedback mechanism	Format of solicited feedback	Format of published feedback
BBC	World news	BBC's Talking Point: Reader 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," and so on.	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 §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 and 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," and so on.	Averages of item ratings; percent 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, and so on.	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.	No explicit feedback scores are published; ordering acts as an implicit indicator of reputation.

publicly known and may, therefore, affect the behavior of the entire community toward that trader in the future. Knowing this, traders have an incentive to behave well toward each other, even if their relationship is a one-time deal. As I discuss in §3, a growing body of empirical evidence seems to demonstrate that these systems have managed to provide remarkable stability in otherwise risky trading environments.

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, and so on. 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 to make a variety of decisions ranging from what movie to watch to what stocks to invest in (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, 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 to 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 to 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 little work studying these and other related possibilities. The rising importance of online feedback mechanisms 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? Which mechanism designs work best in which settings? This is just a small subset of questions that invite exciting and valuable research.

In common with other Internet technologies, 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.[1]

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. Section 4 provides an overview of relevant past work in game theory and economics. Section 5 then discusses how this stylized body of work is being extended to take into consideration the special properties of online environments. Finally, §6 summarizes the main points of this 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

[1] 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.

communities relied on word of mouth as the primary enabler of economic and social activity (Benson 1989, Greif 1993, Milgrom et al. 1990), and 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 (1) their unprecedented scale, achieved through the exploitation of the Internet's low-cost, bidirectional communication capabilities, (2) the ability of their designers to precisely control and monitor their operation through the introduction of automated feedback mediators, and (3) 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). 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 (IT) 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 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 nontrivial ways (see §§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, and so on). 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 roots 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 good reputations, milk it by cheating other members, and then disappear and reappear under new online identities and clean records (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 and 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.[2]

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 them to assume the risk of transacting with complete strangers (Ba and Pavlou 2002). Because *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 to uncover additional properties. Resnick et al. (2002) provide a comprehensive survey and methodological critique of these works. The author is aware of 14 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.

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 seven days, last month) posted negative comments.

Towards a Systematic Discipline of Feedback Mechanism Design. The initial evidence provided

[2] Source: Media Metrix (2001).

Table 2 Summary of Principal Results

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 (2002)	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 price.
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 et al. (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 and 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 et al. (2002)	Vintage postcards	Controlled field experiment; established seller commands higher prices than newcomers; among newcomers, small amounts of negative feedback have little effect.

Note. Adapted from Resnick et al. (2002).

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 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?

Table 3 Summary of Principal Questions

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 or 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

Note. Shorthand is defined in Table 2, Columns 1 and 2.

In the author's opinion, the two most concrete evaluation criteria of a feedback mechanism's performance ought to be (1) 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 (2) the robustness

of those outcomes against different assumptions about the participants' behavior.[3]

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 because 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, because 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. A few examples of a much larger set of possibilities follow.

- 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?
- 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. Because 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, and empirical and experimental studies, are natural complements to game theory, both for qualifying these models and for adapting them to account for the complexities of the "real world" and the bounded rationality of actual human behavior.[4] The rest of this paper provides a survey of past work that can serve as a starting point for answering the above questions in a systematic way.

[3] 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.

[4] 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.

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 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,[5] rational short-run players will then always play their stage game Nash equilibrium response. Such behavior typically results in inefficient outcomes.

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:

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

(2) 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. High-quality products are more profitable, so the seller will always promise high quality. Buyers then compete with one another in 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 0, 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 with 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 et al. (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.[6] 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

[5] 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.

[6] 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 nonstandard payoff structures such that the "commitment" action is their dominant strategy given their payoffs.

the action to which the long-run player would credibly commit if he could. In the "online auction" example, the 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.[7] 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 player they face. An additional assumption is that short-run players have access to the entire history of past stage game outcomes.[8] 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,[9] 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 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). To obtain this payoff, the ordinary player spends long periods of time choosing the Stackelberg action with high probability.[10]

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

[7] 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.

[8] 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 the discussion of complications arising from the *private* observability of outcomes in such systems in §5.2).

[9] 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.

[10] 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.

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 past task outputs. Outputs depend on both ability and labor. The employee's objective is to maximize lifetime wages while minimizing his labor. At equilibrium, this provides incentives to the employee to work hardest early on in her career to build a reputation for competence.

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 more formally explored 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 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; that in, players who always play the action that the short-run players least like. 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: (1) safe borrowers who always select safe projects (i.e., projects with zero probability of default), (2) risky borrowers who always select risky projects (i.e., projects with higher returns if successful but with nonzero probability of default), and (3) 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 (1989) 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 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. Such an outcome is not advantageous for the seller, so reputation considerations will induce the seller to cooperate forever.

The situation radically changes 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 indefinitely sustained: 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 et al. 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, because 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 (1999) 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. The Cripps et al. (2002) result then states that, if the employee stays on the market for a long time, these dynamics will lead to an eventual loss of her reputation.

Even more interesting phenomena come to the surface if one tries to reconcile the above result with Fudenberg and Levine's (1992) result on long-term payoffs induced by reputation. If players eventually lose their reputation, 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 et al. (2000) 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. Because reputation relies on a trade-off 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 approaches.

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 postmortem 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 (1999) 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 because 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 to fully understand the long-term consequences of introducing markets for reputation and for transferring these promising concepts to the online domain.

5. New Opportunities and Challenges of Online Mechanisms

In §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 IT 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.[11] 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 (2002) 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.

[11] This is an alternative way to interpret the results of Klein and Leffler (1981) and Shapiro (1983).

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: (1) ensuring that sufficient feedback is, indeed, provided, and (2) 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, because 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 et al. (1999) analyze mechanisms whereby early evaluators are paid to provide information and later evaluators pay 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.

Because monitoring is private and assessments usually subjective, an additional consideration is whether feedback is honest. Miller et al. (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, 2004) explores the use of robust statistics in aggregating individual ratings as a mechanism for reducing the effects of coordinated efforts to bias ratings. To 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 gamey 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 nontrivial 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 §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, the Cripps et al. (2002) result predicts that, in noisy environments, a mechanism that simply publishes the entire history of feedback will not lead to sustainable truth telling. Dellarocas (2002) 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 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 reregister under a completely different online identity with zero or 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 (2001) 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 reentry 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 reentry. 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 (2001) 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.[12] Dellarocas (2003) further demonstrates that, although this design incurs efficiency losses relative to the case where identity change 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 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 IT.

In recent years, the field of multi-agent systems (Jennings et al. 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 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

[12] 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.

hits. Google's page reputation measure is based on the number of links that point to a page, the number of links that 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 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 et al. (2002) apply network flow techniques 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 and 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 to determine in what settings it might be a viable substitute for, or complement to, voluntary feedback provision.

Decentralized Feedback Architectures. 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 et al. 2000, Mui et al. 2001, Sen and Sajja 2002, Yu and Singh 2002).

The emergence of peer-to-peer networks provides a further motivation for developing decentralized feedback mechanisms. In such networks, feedback mechanisms represent a promising approach for inducing cooperative behavior among the participant nodes. Initial attempts to develop reputation mechanisms for peer-to-peer networks are reported in Aberer and Despotovic (2001), Kamvar et al. (2003), and Xiong and Liu (2003).

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.[13] Two recent laboratory experiments provide some initial insight into human behavior vis-à-vis feedback mechanisms. Bolton et al. (2002) compare trading in a market with (automatically generated) feedback to a market without and 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

[13] For an excellent survey of psychological findings relevant to economics, the reader is strongly encouraged to consult Rabin (1998).

collect and aggregate information from large communities at low cost, 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.

The design of online feedback mechanisms can benefit greatly from the insights produced in more than 20 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 some 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 enforced contractual guarantees and advertising); develop theory-driven guidelines for deciding which set of mechanisms to use and 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.

Other established 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 analytically solve. 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.[14]

There is much work to be done. But it is important that research be conducted now, in the formative phases of this technology and the social practices surrounding it. There are likely to be path-dependent effects in the deployment and use of online feedback mechanisms, so it is important that researchers develop insights into the functioning and impacts of these systems, while they can still have a large impact on practice.

References

Aberer, K., Z. Despotovic. 2001. Managing trust in a peer-2-peer information system. *Proc. Tenth Internat. Conf. Inform. Knowledge Management.* Association for Computing Machinery, Atlanta, GA, 310–317.

Avery, C., P. Resnick, R. Zeckhauser. 1999. The market for evaluations. *Amer. Econom. Rev.* **89**(3) 564–584.

Ba, S. 2001. Establishing online trust through a community responsibility system. *Decision Support Systems* **31**(3) 323–336.

[14] 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 (2001) 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.

——, P. Pavlou. 2002. Evidence of the effect of trust building technology in electronic markets: Price premiums and buyer behavior. *MIS Quart.* **26**(3) 243–268.

Bajari, P., A. Hortacsu. 2003. Winner's curse, reserve prices and endogenous entry: Empirical insights from eBay auctions. *RAND J. Econom.* **34**(2) 329–355.

Bakos, Y., C. Dellarocas. 2002. Cooperation without enforcement? A comparative analysis of litigation and online reputation as quality assurance mechanisms. L. Applegate, R. Galliers, J. I. DeGross, eds. *Proc. 23rd Internat. Conf. Inform. Systems* (ICIS 2002). Association for Information Systems, Barcelona, Spain, 127–142.

Benson, B. 1989. The spontaneous evolution of commercial law. *Southern Econom. J.* **55**(January) 644–661. Reprinted in D. B. Klein, ed. 1997. *Reputation: Studies in the Voluntary Elicitation of Good Conduct*. University of Michigan Press, Ann Arbor, MI, 165–189.

Bolton, G. E., E. Katok, A. Ockenfels. 2002. How effective are online reputation mechanisms? An experimental investigation. Discussion paper 25-2002, Max Planck Institute for Research into Economic Systems, Jena, Germany, ftp://ftp.mpiew-jena.mpg.de/pub/esi/discussionpapers/2002-25.pdf.

Brin, S., L. Page. 1998. The anatomy of a large-scale hypertextual web search engine. *Comput. Networks ISDN Systems* **30**(1–7) 107–117.

Cripps, M., G. Mailath, L. Samuelson. 2003. Imperfect monitoring and impermanent reputations. *Econometrica*. Forthcoming.

Dellarocas, C. 2000. Immunizing online reputation reporting systems against unfair ratings and discriminatory behavior. *Proc. 2nd ACM Conf. Electronic Commerce*. Association for Computing Machinery, Minneapolis, MN, 150–157.

——. 2001. Analyzing the economic efficiency of eBay-like online reputation reporting mechanisms. *Proc. 3rd ACM Conf. Electronic Commerce*. Association for Computing Machinery, Tampa, FL, 171–179.

——. 2002. Goodwill hunting: An economically efficient online feedback mechanism in environments with variable product quality. J. Padget, O. Shehory, D. Parkes, N. Sadeh, W. Walsh, eds. *Agent-Mediated Electronic Commerce IV. Designing Mechanisms and Systems*. Lecture Notes in Computer Science 2351, Springer Verlag, Berlin, Germany, 238–252.

——. 2003a. Building trust on-line: The design of robust reputation mechanisms for online trading communities. G. Doukidis, N. Mylonopoulos, N. Pouloudi, eds. *Social and Economic Transformation in the Digital Era*. Idea Group Publishing, Hershey, PA.

——. 2003b. Efficiency and robustness of binary feedback mechanisms in trading environments with moral hazard. Working paper 4297-03, Massachusetts Institute of Technology, Cambridge, MA, http://ssrn.com/abstract=393043.

Dewan, S., V. Hsu. 2002. Adverse selection in reputations-based electronic markets: Evidence from online stamp auctions. Working paper, Graduate School of Management, University of California, Irvine, CA, http://web.gsm.uci.edu/~sdewan/Home%20Page/Adverse%20Selection%20in%20Reputations.pdf.

Diamond, D. 1989. Reputation acquisition in debt markets. *J. Political Econom.* **97**(4) 828–862.

Eaton, D. H. 2002. Valuing information: Evidence from guitar auctions on eBay. Working paper #0201, Department of Economics, Murray State University, Murray, KY, http://campus.murraystate.edu/academic/faculty/david.eaton/workpaper0201.pdf.

Friedman, E., P. Resnick. 2001. The social cost of cheap pseudonyms. *J. Econom. Management Strategy* **10**(1) 173–199.

Fudenberg, D., D. Levine. 1992. Maintaining a reputation when strategies are imperfectly observed. *Rev. Econom. Stud.* **59**(3) 561–579.

Greif, A. 1993. Contract enforceability and economic institutions in early trade: The Maghribi Traders' Coalition. *Amer. Econom. Rev.* **83**(June) 525–548.

Guernsey, L. 2000. Suddenly, everybody's an expert on everything. *The New York Times* (February 3).

Holmstrom, B. 1999. Managerial incentive problems: A dynamic perspective. *Rev. Econom. Stud.* **66**(1) 169–182.

Horner, J. 2002. Reputation and competition, *Amer. Econom. Rev.* **92**(3) 644–663.

Houser, D., J. Wooders. 2000. Reputation in auctions: Theory and evidence from eBay. Working paper 00-01, Department of Economics, University of Arizona, Tucson, AZ, http://infocenter.ccit.arizona.edu/~econ/working_papers/Internet_Auctions.pdf.

Jennings, N., K. Sycara, M. Wooldridge. 1998. A roadmap of agent research and development. *Autonomous Agents Multi-Agent Systems* **1**(1) 275–306.

Kalyanam, K., S. McIntyre. 2001. Returns to reputation in online auction markets. Working paper W-RW01-02, Retail Workbench, Santa Clara University, Santa Clara, CA.

Kamvar, S. D., M. T. Schlosser, H. Garcia-Molina. 2003. The Eigentrust algorithm for reputation management in P2P networks. *Proc. Twelfth Internat. Conf. World Wide Web*. Association for Computing Machinery, Budapest, Hungary, 640–651.

Kauffman, R. J., C. Wood. 2000. Running up the bid: Modeling seller opportunism in Internet auctions. *Proc. Amer. Conf. Inform. Systems* (AMCIS 2000). Association for Information Systems, Long Beach, CA, 929–935.

Keser, C. 2002. Trust and reputation building in e-commerce. Scientific publication 2002s-75, Center for Interuniversity Research and Analysis on Organizations (CIRANO), Montreal, Quebec, Canada, http://www.cirano.qc.ca/pdf/publication/2002s-75.pdf.

Klein, B., K. Leffler. 1981. The role of market forces in assuring contractual performance. *J. Political Econom.* **89**(4) 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. E. J. Lawler, M. Macy, S. Thyne, H. A. Walker, eds. *Advances in Group Processes*, Vol. 16. JAI Press, Greenwich, CT.

Kreps, D., R. Wilson. 1982. Reputation and imperfect information. *J. Econom. Theory* **27**(2) 253–279.

——, P. Milgrom, J. Roberts, R. Wilson. 1982. Rational cooperation in the finitely repeated prisoners' dilemma. *J. Econom. Theory* **27**(2) 245–252.

Lee, Z., I. Im, S. J. Lee. 2000. The effect of negative buyer feedback on prices in internet auction markets. W. J. Orlikowski, S. Ang, P. Weill, H. C. Krcmar, J. I. DeGross, eds. *Proc. 21st Internat. Conf. Inform. Systems* (ICIS 2000). Association for Information Systems, Brisbane, Australia, 286–287.

Livingston, J. 2002. How valuable is a good reputation? A sample selection model of Internet auctions. Working paper, University of Maryland, College Park, MD, http://www.wam.umd.edu/~kth/reputation1.pdf.

Lucking-Reiley, D., D. Bryan, N. Prasad, D. Reeves. 2000. Pennies from eBay: The determinants of price in online auctions. Working paper, University of Arizona, Tucson, AZ, http://eller.arizona.edu/~reiley/papers/PenniesFromEBay. html.

Mailath, G. J., L. Samuelson. 1998. Your reputation is who you're not, not who you'd like to be. Working paper 98-11, Center for Analytic Research in Economics and the Social Sciences (CARESS), University of Pennsylvania, Philadelphia, PA, http://www.ssc.upenn.edu/~gmailath/wpapers/rep-is-sep.html.

——, ——. 2001. Who wants a good reputation? *Review Econom. Stud.* **68**(2) 415–441.

Mayzlin, D. 2003. Promotional chat on the Internet. Working paper #MK-14, Yale School of Management, New Haven, CT.

McDonald, C. G., V. C. Slawson, Jr. 2002. Reputation in an internet auction market. *Econom. Inquiry* **40**(3) 633–650.

Melnik, M. I., J. Alm. 2002. Does a seller's reputation matter? Evidence from eBay auctions. *J. Indust. Econom.* **50**(3) 337–349.

——, D. North, B. R. Weingast. 1990. The role of institutions in the revival of trade: The law merchant, private judges, and the champagne fairs. *Econom. Politics* **2**(1) 1–23. Reprinted in D. B. Klein, ed. 1997. *Reputation: Studies in the Voluntary Elicitation of Good Conduct*. University of Michigan Press, Ann Arbor, MI, 243–266.

Milgrom, P., J. Roberts. 1982. Predation, reputation and entry deterrence. *J. Econom. Theory* **27**(2) 280–312.

Miller, N., P. Resnick, R. Zeckhauser. 2002. Eliciting honest feedback in electronic markets, research. Working paper RWP02-039, Harvard Kennedy School, Cambridge, MA, http://www.si.umich.edu/~presnick/papers/elicit/index. html.

Mui, L., P. Szolovits, C. Ang. 2001. Collaborative sanctioning: Applications in restaurant recommendations based on reputation. *Proc. 5th Internat. Conf. Autonomous Agents*. Association for Computing Machinery, Montreal, Quebec, Canada, 118–119.

Papadimitriou, C. 2001. Algorithms, games, and the Internet. *Proc. 33rd Annual ACM Sympos. Theory Comput.* Association for Computing Machinery, Hersonissos, Greece, 749–753.

Pujol, J. M., R. Sangüesa, J. Delgado. 2002. Extracting reputation in multi agent systems by means of social network topology. *Proc. First Internat. Joint Conf. Autonomous Agents Multiagent Systems*. Association for Computing Machinery, Bologna, Italy, 467–474.

Rabin, M. 1998. Psychology and economics. *J. Econom. Literature* **36**(1) 11–46.

Resnick, P., R. Zeckhauser. 2002. Trust among strangers in Internet transactions: Empirical analysis of eBay's reputation system. M. R. Baye, ed. *The Economics of the Internet and E-Commerce. Advances in Applied Microeconomics*, Vol. 11. JAI Press, Greenwich, CT.

——, ——, E. Friedman, K. Kuwabara. 2000. Reputation systems. *Comm. ACM* **43**(12) 45–48.

——, ——, J. Swanson, K. Lockwood. 2002. The value of reputation on eBay: A controlled experiment. Working paper, University of Michigan, Ann Arbor, MI, http://www.si.umich.edu/~presnick/papers/postcards/index.html.

Roth, A. 2002. The economist as engineer: Game theory, experimentation, and computation as tools for design economics. *Econometrica* **70**(4) 1341–1378.

Sabater, J., C. Sierra. 2002. Reputation and social network analysis in multi-agent systems. *Proc. First Internat. Joint Conf. Autonomous Agents and Multiagent Systems.* Association for Computing Machinery, Bologna, Italy, 475–482.

Sen, S., N. Sajja. 2002. Robustness of reputation-based trust: Boolean case. *Proc. First Internat. Joint Conf. Autonomous Agents Multiagent Systems*. Association for Computing Machinery, Bologna, Italy, 288–293.

Shapiro, C. 1982. Consumer information, product quality, and seller reputation. *Bell J. Econom.* **13**(1) 20–35.

——. 1983. Premiums for high quality products as returns to reputations. *Quart. J. Econom.* **98**(4) 659–680.

Tadelis, S. 1999. What's in a name? Reputation as a tradeable asset. *Amer. Econom. Rev.* **89**(3) 548–563.

——. 2002. The market for reputations as an incentive mechanism. *J. Political Econom.* **92**(2) 854–882.

Wilson, R. 1985. Reputations in games and markets. A. Roth, ed. *Game-Theoretic Models of Bargaining*. Cambridge University Press, Cambridge, U.K., 27–62.

Xiong, L., L. Liu. 2003. A reputation-based trust model for peer-to-peer eCommerce communities. *Proc. IEEE Internat. Conf. E-Commerce (CEC'03)*. Institute of Electrical and Electronics Engineers, Newport Beach, 275–284.

Yu, B., M. Singh. 2002. An evidential model of distributed reputation management. *Proc. First Internat. Joint Conf. Autonomous Agents Multiagent Systems*. Association for Computing Machinery, Bologna, Italy, 294–301.

Zacharia, G., A. Moukas, P. Maes. 2000. Collaborative reputation mechanisms in electronic marketplaces. *Decision Support Systems* **29**(4) 371–388.

——, T. Evgeniou, A. Moukas, P. Boufounos, P. Maes. 2001. Economics of dynamic pricing in a reputation brokered agent mediated marketplace. *Electronic Commerce Res.* **1**(1) 85–100.

Accepted by Arthur M. Geoffrion and Ramayya Krishnan; received December 2001. This paper was with the author 19 months for 6 revisions.