The Dynamics of Reputation and Its Origins in Economic and Social Contexts

Thesis submitted for the degree of "Doctor of Philosophy"

By AMIR BAN

Submitted to the Senate of the Hebrew University of Jerusalem
July 2014

This work was carried out under the supervision of:

Nati Linial

Acknowledgements

I am grateful to Nati Linial for his wise guidance and support, and for teaching me many things.

I dedicate this work to the memory of my father, E.G. Ban (*eg-ban.org*). I still follow his footsteps.

Abstract

Reputation is a measure of a society's belief about the quality, or expertise, of an entity, which may be an expert, a product, a restaurant and so on. Often, the only sources of information about quality are reputation and personal experience. Can we expect that the entity with the highest reputation is the one with the highest quality? At any specific point in time? In the long run? Our study of the dynamics involved answers both these questions in the negative. Our investigations show that favorable initial conditions (an advantage in reputation) often outweigh a superiority in quality. In this thesis this advantage is established in a quantitative manner. It turns out that the dynamics closely resemble those of recommendation systems and search engines. Also, in a multi-product market, market share is closely related to reputation. But then, if reputation and quality do not faithfully reflect each other, is it rational to use reputation as an indicator for quality? It is shown here that the answer to this question is affirmative.

If reputation, rather than being common knowledge, is propagated by a social or geographical network, new questions arise. Most studies of social learning predict conformity as a steady-state, which is at odds with the observed quasi-stable diversity or heterogeneity of the real world. A simple model of reputation that we introduce is able to accommodate the observed diversity.

A Letter of Contribution

The papers submitted in this thesis have all been authored by my instructor Nati Linial and by me, and are all signed by both of us. In every one of the papers, the individual contribution of each of us was substantial.

Most of the papers were reviewed by various reviewers, presented to many audiences and discussed with several colleagues. These raised questions, found faults, suggested lines of investigation, and so on. None of these reached the level of contribution, with the following possible exception: An anonymous reviewer for Games and Economic Behavior sketched the line of proof we use for the main theorems of "Market Share Indicates Quality" in place of our former more elaborate proof. He graciously declined our offer of co-authorship status.

Contents

Preface

The motivation for this thesis stems from the desire to understand how society arrives at a consensus regarding the truth or quality of a subject matter (e.g. "Einstein was right", "Van Gogh was a great painter"). The objective truth or quality of these statements are insufficient to explain their universal acceptance, not least because historically their acceptance was a gradual process. That Einstein and Van Gogh are universally acclaimed is now in large part self-sustaining, and ultimately resides on the word of *authorities* who attest to their truth or excellence, and who are trusted by the public. But this begs the question: How does the public know an authority's word can be trusted and by what process does the authority gain its position of influence?

A related phenomenon is "virality": The rapid spread of popularity or acceptance of a cultural item, as opposed to the gradual and limited spread of other cultural items.

Public perception of excellence, as well as the ability of authorities to influence, is called "reputation" in informal speech. The term is often used in game theory and in economics, and has several meanings, many similar but some different from the meaning used here.

"The Dynamics of Reputation Systems" was published as extended abstract in the proceedings of *TARK '11*. In this paper we formulate a model that captures the essence of these questions. Reputations systems and recommendation systems on the Internet and others such as citation networks closely model the pertinent questions. Search engines seem especially relevant, in particular Google's *page rank*, in which a document's relevance and influence arises out of the structure of the network itself. The question is

distilled in this paper to the following: Suppose there are several *experts*, each able to provide a service that the public seeks. The experts differ in their level of *expertise*, defined as the probability for providing satisfactory service. The public approaches experts mainly based on their *reputation*, which is formed for each expert by selective reporting of his past successes and failures.

The question posed in this paper is this: How well does the rank-order of reputation fit the rank-order of expertise? In other words, does reputation reflect expertise? The alternative is that reputation is determined by initial conditions (a high reputation is self-sustaining, while a highly-skilled but novice expert might take much time to build a reputation). The paper shows that both outcomes are possible, and in the long-run the outcome that occurs depends on simple inequality relations between the expertise levels of the experts.

The steady-state reputation order of several experts, given their expertise and initial reputation order, turned out to hinge on a question on partially ordered sets. Despite being elementary, it turned out not to have been addressed in the literature. Therefore, after solving it, and proving that the final order is unique given the initial order, it was published in the *Journal of Combinatorial Theory Series A (JCTA)* as "Strong convergence of posets".

The reputation model posits that users use reputation as a proxy for expertise, e.g. when a user approaches experts serially in order of descending reputation. Is this model assumption rational or *ad hoc*? The results in "The Dynamics of Reputation Systems" make this question even more poignant: If higher reputation does not always reflect higher expertise, why should a rational user assume that it does? And might there be a better strategy for users than the model assumption? The question turned out to be difficult, but the answer was reassuring: Reputation is indeed an indication for expertise, under very mild assumptions on the behavior of users. The model, therefore, may arise out of the autonomous decisions of benefit-maximizing agents.

This result is formulated in the context of *market share* rather than reputation, whence the title "Market Share Indicates Quality". A recurring

observation in this thesis is that reputation and market share are closely allied concepts: Whereas reputation is the sum of endorsements of an expert, market share is the total consumption of a product. The product quality is a cognate of an expert's expertise. Assuming customers are at least partially aware of market share, and of their own experiences with products, dynamics arise that resemble those described above for reputation. In this context, the result states that, based on market share information only, there is at least a 50% probability that the product with higher market share has the higher quality. At the time of writing, this paper is under review for publication.

The papers so far mentioned treat reputation (and market share) as a *global* attribute of a society: The individuals of the society are undifferentiated according to place and position, and reputation is common knowledge. An extension of the model to include social and geographical relations suggests itself: In such a model, society is a social network where vertices represent individuals. Reputation is local and is indirectly propagated across the network, where edges signify social or geographical proximity. The local model facilitates the study of the dynamics and statics of *social choices*. A social choice is an individual's choice among several possibilities, which is formed under the influence of the choices of that individual's neighborhood: E.g. choice of native language (Irish Gaelic vs. English) or a preference for wine over beer. A geographical network is a specialized network in which individuals are placed on a map, and connections signify proximity (e.g. individuals who live less than 10 km from each other are connected by an edge).

The dynamics of reputation as adapted to this localized model constitute a model of social learning. It is described in "Why Social Learning May Lead to Diversity". Preferences (a high reputation) of social choices evolve and propagate in the network. Choices are not necessarily of equal merit: The expertise, or quality attribute makes some choices objectively better than others, but as seen in the global model, an advantage in expertise does not necessarily imply an advantage in reputation. Studies of social learning in the literature commonly conclude that society reaches conformity, except perhaps where social choices have equal merit. This conclusion seems to be at

odds with observed reality, where a quasi-stable diversity of social choices is commonly found. In contrast, it was immediately apparent that the localized reputation model has diverse and stable solutions even when merits of social choices are unequal.

A good part of this thesis revolves around the question whether diverse steady-state solutions are feasible. This turned out to hinge on open questions in graph theory and in geometry. In graph theory, the question (for two equal-merit choices) may be phrased as following: When is it possible to partition a simple undirected graph in two, so that each vertex has most of its neighbors in its own part? In geometry, the question is: When is it possible to partition a 2-(or higher-) dimensional set such that at least half of the neighborhood (defined as points nearer than some positive given radius by the Euclidean or other metric) is on its own side of the partition?

The graph theory aspect is analyzed in "Internal Partitions of Regular Graphs". This paper is currently in review. The geographical/geometric aspect is analyzed in "Why Social Learning May Result in Diversity". This paper contains several intriguing conjectures and some approaches toward their proof. Our intention is to send it to publication once these mysteries are resolved.

The Dynamics of Reputation Systems

Published:
Ban, Amir, and Nati Linial. "The dynamics of reputation systems." In Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge, pp. 91-100. ACM, 2011.

Market Share Indicates Quality

Published in an electronic format:
Ban, Amir, and Nati Linial. "Market Share Indicates Quality." arXiv preprint arXiv:1407.3641 (2014).

Strong Convergence in Posets

Published:
Ban, Amir, and Nati Linial. "Strong convergence in posets." Journal of Combinatorial Theory, Series A 119, no. 6 (2012): 1299-1301.

Internal Partitions of Regular Graphs

Published in an electronic format:
Ban, Amir, and Nati Linial. "Internal Partitions of Regular Graphs." arXiv preprint arXiv:1307.5246 (2013).

Why Social Learning May Result in Diversity

Unpublished

Conclusion

The questions I sought to investigate in this Ph.D. thesis, as described in the Preface, were ambitious. The results of the investigation are detailed here in the form of three papers dealing directly with the social and economic subject matter, and two papers dealing with theoretical questions that arose out of the research.

In aggregate, the resulting papers show how to model the situations described, and reach significant results. Some of these neatly capture home truths, or our cynical suspicions about the world we live in, while others are surprising and thought-provoking. They also show that this line of investigation is fruitful and may be continued in various directions.

The success of the research is in large part due to the model itself. An examination of the three non-mathematical papers will show that the models used are adaptations of the same model concepts, which, in turn, differs from models used in the literature to examine related questions. The common thread of these models is:

- Expertise or quality - An objective but hidden quantity of a service provider, signifying the probability for providing satisfaction to a user

- Reputation or market-share - The public's shared assessment of expertise/quality, aggregated from individual user actions

The benefit of a service to a user, assuming satisfaction, is fixed: If a user receives satisfactory service, the quality of the provider is of no importance: Quality is important *ex ante* but not *ex post*. Since users do not wish to waste

resources on unsatisfactory trials, they are motivated to try the highest-quality experts first, and since quality is not publicly known, users rely on reputation or market-share as a proxy for quality or expertise.

In this model, the race for reputation is a race for being given first chance to perform. This is crucial for success: Given a chance, an expert has a fair chance of serving a customer satisfactorily, even if his expertise (probability to satisfy) is far from perfect, while a much better but lower-ranked expert has little chance to prove his worth. Furthermore, the results of the research show that this anomaly is often not temporary: An initially low reputation (or market share) might prevent gaining a higher one at any time in the future.

Models used in the literature to capture related questions, such as *preferential attachment*, are significantly different in their details and in their results. In preferential attachment models, users randomize their choices with a distribution based on the providers' sizes. As a result "the rich get richer", as in the reputation model, but there is no underlying expertise parameter and therefore no plausible scenario for the "poor" and the "rich" to change places, unlike in the reputation model. Furthermore, as individual behavior it is difficult to justify the randomization of choices inherent to preferential attachment, since a rational individual would mix between possible actions only when indifferent about their outcomes. If a user believes reputation is a signal for expertise, she has no reason to mix between higher and lower reputation providers.

The anomalous results of the reputation model apparently send a message to users to take reputation, or market share, with a grain of salt. Possibly, it might be argued, if they do so, and pursue strategies which are more subtle than mere reputation seeking, different dynamics will ensue, with no or fewer anomalies. However, the results of "Market Share Indicates Quality" show that such subtler strategies may not be available: Even when users are aware that reputation is not a fool-proof guide to expertise, they may have no better strategy than to assume that it is.

The result in "Market Share Indicates Quality" also shows that the behavior of users in the reputation model is rational given their information, and

that therefore the dynamics of reputation, with all their noted anomalies, will arise out of the rational behavior of autonomous, benefit-maximizing agents. When presenting the reputation model and its results, we have often encountered critical comments on these anomalous results. These comments would often be accompanied by helpful suggestions how to change the model so as to eliminate these anomalies. This, however, misses the motivation of the research which is not concerned with mechanism design to ensure that reputation faithfully represents expertise. Indeed it would be relatively simple to alter the model to ensure this. For example, if users reported the results of all encounters with experts, whether satisfactory or non-satisfactory, simpler dynamics would be the result and anomalies would largely disappear. However, such behavior is commonly not observed in the real world (market share, for example, counts consumption events, while non-consumption events are neither counted nor even reported), and the motivation for a user behaving this way would need to come from artificial exogenous incentives.

Some broad conclusions that may be drawn from the research:

- Reputation is important. It is worth obtaining and defending.
- It is rational to work with reputation systems. Therefore such systems may arise out of the voluntary cooperation of independent agents.
- All is not what it seems.

The research into the social and geographical aspects of reputation has diverged into various directions, some of which were unanticipated. The object was to investigate a model of localized reputation, which is (in effect) propagated across a network representing social or geographical connections. The local model was constructed from the reputation model of "The Dynamics of Reputation Systems", but with local node reputation values accumulated from neighbors' feedback, replacing global reputation values accumulated from an entire community's feedback.

This constituted a model of social learning. The ranking of reputation represents the preference for one social choice over another. E.g., some nodes of the network might prefer (attach higher reputation to) expert A over

expert B, while in other nodes the preference may be reversed. The basic question asked was: Will a connected network eventually display uniform preferences, or will the network converge on a diverse steady state, one in which different parts of the network have different preferences? In other words, is stable diversity of preferences likely, or even possible?

The question, stated in terms of graph theory is: Can a simple undirected graph be partitioned into two (assuming the simple case where the social choice is between just two possibilities) components, such that every vertex has at least as many edges from its own component as from any other (with edges coming from different components being weighed differently, to reflect that social choices have different intrinsic merit).

The question, it turned out, was basically unanswered in the literature, although it was previously stated and some special cases answered. However, a result by Stiebitz (1996) shows that the answer is affirmative when some "inertia" is assumed, i.e., if a node gives sufficient weight to its current preference. With no inertia, the question is difficult. The answer is often positive, but some classes of exceptions are known. In "Internal Partitions of Regular Graphs" we found several more classes of exceptions, proven partitionability for some other cases, and made conjectures regarding the partitionability in the general case.

These results support the existence of a stable diversity, if not in all, then in many social networks, and in all of them if some inertia is modeled. It should be noted that the existence of a diverse steady state does not mean that such a steady state is the forced or even the likely conclusion of any initial state of society.

The question is particularly interesting in a geographical setting, as it investigates the diversity of social choices across geographical boundaries: Preferences for Pepsi-Cola vs. Coca-Cola are known to follow stable geographical patterns; the domains where Celtic languages are spoken have shrunk since ancient times, etc. The network here is a specialized social network: Individuals are connected if they reside in geographical proximity: Their mutual distance is below some threshold. Distance, in this context, is not necessarily as-the-crow-flies, but may be a metric, measuring, e.g. the

cost of traveling from one location to another. A further idealization is to make the network into a map: Instead of nodes (representing individuals) densely filling an area, we imagine the area as uniformly, continuously filled. The graph-theoretical problem is then transformed into a geometric one: Can a given map be partitioned into regions, so that a circle of given radius drawn around any point on the map covers a greater part in its own region than in the other region (with areas suitably weighed to reflect that social choices have different intrinsic merit)?

The model is developed in "Why Social Learning may Result in Diversity". In this paper a conjecture is put forward for the following most general result: Any set S in m-dimensional space ($m \geq 1$), and given a metric μ and an influence radius $r > 0$, and $0 < q < 1$, allows a partition into two parts A, B, such that the "neighborhood" (i.e. intersection of S and a μ-ball of radius r) of any $x \in A$ has at least q of its area in A, while the "neighborhood" of any $x \in B$ has at least $1-q$ of its area in B. If true, this would imply that stable diversity is possible (though not inevitable) for every conceivable map, every definition of neighborhood, and every balance of strengths between social choices. The attempted proof in "Why Social Learning may Result in Diversity" is however incomplete (it constructs a series of increasingly fine-grained discrete solutions, and misses a proof that this series converges). At the time of writing I attempt to close the gap in the proof.