Strategy and Dynamics in Reputation Systems

Amir Ban

Center for the Study of Rationality, Hebrew University, Jerusalem, Israel

Nati Linial

School of Computer Science and Engineering, Hebrew University, Jerusalem, Israel

Abstract

Online reputation systems collect, maintain and disseminate reputations as a summary numerical score of past interactions of an establishment with its users. As reputation systems, including web search engines, gain in popularity and become a common method for people to select sought services, a dynamical system unfolds: *Experts'* reputation attracts the potential *customers.* The experts' *expertise* affects the probability of satisfying the customers. This rate of success in turn influences the experts' reputation. We consider here several models where each expert has innate, constant, but unknown level of expertise and a publicly known, dynamically varying, reputation.

The specific model depends on (i) The way that experts' reputation affects customers' preferences, (ii) How experts' reputation is modified as a result of their success/failure in satisfying the customers' requests.

We investigate several such models and elucidate some of the key characteristics of reputation in such a market of experts and customers.

Keywords: Reputation, expert, social learning, search engine, reputation system

Email addresses: amirban@netvision.net.il (Amir Ban), nati@cs.huji.ac.il (Nati Linial)

Preprint submitted to Elsevier *March 8, 2011*

1. Introduction

1.1. How to Find a Good Expert

How do you find a good restaurant to celebrate that special occasion? How to find a lawyer, mechanic or medical specialist when you need one? Where do you find the citation you need in your paper for a subject that is peripheral to your own work? How to find a good movie, a book, a travel destination or a cool YouTube video? Traditional methods of answering these questions are more and more being augmented, or even replaced, by online reputation systems.

According to Friedman et. al.[5], online reputation systems collect, maintain and disseminate *reputations*, aggregated records of past interactions of each participant in a community. While reputation systems have been used to track the trustworthiness of a participant, the property being tracked by a reputation system is more generally a measure of the quality or ability of a participant in a certain class of interactions. Online systems for grading restaurants, travel destinations, movies as well as academic papers and researchers are examples.

We use "expertise" as a catch-all term for the attribute on which a reputation system is designed to report. *Expertise* may be a measure of the quality of services (of a certain kind) rendered, ability to perform a task well, ability to correctly answer questions or predict events of a certain kind, or trustworthiness in interactions. Participants in a reputation system may be divided into *experts*, who provide a service and have an aggregated reputation record, and *users*, who use the expert's services and the system's disseminated reputations. In so-called *peer-to-peer* reputation systems each participant serves in both roles.

We define **expertise** to be the probability that an expert's service will meet a user's requirements, with the probability meant to cover the great variability in this user-expert interaction: variations between the requirements of members of the community, variations of requirements on different occasions, and variations in the skill of the expert according to his[1] own circumstances. We make the

[1] Throughout this paper we refer to experts as masculine and to members of the public as feminine.

assumption that this expertise probability is an objective and fixed attribute of an expert, which is however hidden and can only be estimated by trials. Hence the value of the reputation system in providing a community of users information that is helpful in estimating expertise.

We treat expertise as an innate ability of an expert, not a strategic choice. We assume that an expert cannot perform better than his fixed expertise and has no cost in doing his best (and therefore has no motive in performing worse).

1.2. Reputation

The reputation provided by a reputation system may consist of the aggregated record of reported past interactions. More commonly it is a numerical score derived from that record, enabling experts rated by the system to be ranked. The reputations disseminated by a reputation system help users estimate the objective expertise (defined as probability for success) of each expert. Often, it is enough if the reputations assist users in *ranking* the estimated expertise levels, since a user's strategy typically consists of selecting the best available expert.

A reputation system may be called effective if, given the way past interactions are reported and aggregated, a higher reputation score more often than not reflects a higher expertise. Consequently the effectiveness of a reputation system depends not only on its design, but on its sources of information as well.

When all past interactions, successful or not, are recorded by a reputation system, designing it for effectiveness is simple: The success rate of each expert (the ratio of successes to trials) is an unbiased estimator of expertise and therefore may be used as the reputation score. Users may make the natural and simple choice of the highest-score expert, or may engage in longer-range and ultimately superior k-armed bandit strategies[6].

However, full reporting, or even representative reporting of interactions is not realistic in most real-world situations. Often, the motivations of users who provide reports to a reputation system lead to one-sided reporting, of successes only, or, less commonly, of failures only. Some systems, such as search engines (where the "endorsements" consist of links), or academic importance (where the "endorsements" consist of citations), simply have no scope for negative mention. Statistics made of sales figures, such as for books or movies, also reflect only positive choices, while a customer's decision not to buy is a silent, unrecorded act. Nor is the problem solely a question of the reputation system's design: Users seem to be inclined to report positive interactions, while generally glossing over negative ones. For example, a system of "thumbs up/down" introduced to rate YouTube videos shows that "thumbs up" occurrences consistently outnumber "thumbs down" occurrences so heavily that it is doubtful that their relative proportion measures viewer approval/disapproval.

1.3. Expertise vs. Reputation

We aim to investigate the relation between expertise and reputation in a reputation system, with two key questions in mind. The first of these is to determine under what conditions a reputation system is effective, in the sense given above, i.e. what behavior of the system and its users will make its reputation scores a positive signal of expertise. The implication of an effective system is that choosing an expert by reputation is a correct strategic choice for each of its users in maximizing their chances of obtaining satisfactory service, and therefore furthermore, that the reputation system is viable by eliciting the collaboration and participation of autonomous users following their own interests.

We shall demonstrate that, under mild and practically unrestrictive conditions, a reputation system that relies on (possibly selective) user feedback is effective, and reliance on its reputation scores is correct strategy for its users.

A related, but different, question that we investigate is: Does reputation, the public's perception of expertise, reflect real expertise in the long run? Can we be assured that, between several experts of varying skill, eventually the expert of highest expertise will have the highest reputation? Or, alternatively, might reputation be self-perpetuating, with a high reputation merely reflecting a favorable head start?

Both possibilities are plausible. On the one hand, an expert's reputation is reinforced by a customer's positive experience, which becomes more likely the higher his expertise. On the other hand, the number of his customers depends on his current reputation. Both expertise and current reputation therefore contribute positively to future reputation and either may conceivably dominate in the long run.

The expertise-reputation contest is well-known to the corporate world: Entry into an established market is generally difficult and costly. Having an excellent product may not suffice, and the newcomer may need to spend a lot to close his reputation gap: For example, in the mid-1990's Netscape ruled the web browser world, until Microsoft's Internet Explorer managed (with considerate effort) to sideline it. On the other hand, the success of Mozilla's Firefox, a non-profit open source browser, teaches that entry against an entrenched leader is possible, and on merit alone.

Nobody in their right mind expects a better-tasting but no-name cola drink to supplant Coca-Cola and Pepsi-Cola. Nor is it commonly believed that the unrivaled supremacy of these mega brands rests on the unrivaled quality of their soft drinks.

The situation is well-known in the cultural world, where being "in vogue" is in large part self-sustaining: The people who flock to see a Van Gogh exhibition, a Rolling Stones concert or a performance of Verdi's Aida seem to be driven at least in part by the respective artists being acknowledged as "all-time greats". Indeed Van Gogh's wretched career during his own lifetime indicates that there is nothing inevitable about his posthumous fame.

In arts, music and literature there is still an intangible but definite "expertise", but when considering the "reputation" of TV personalities, movie stars, supermodels and so on, it becomes less and less clear what it is that sustains them in their elevated position in the face of the hordes of wannabes who would love to take their place and seem just as qualified. This has led a cynic to quip "a celebrity is someone who is famous for being well-known".

We will show that being ahead in reputation confers a quantifiable advantage that may balance inferiority in expertise ("No. 2 Tries Harder"). In a system where reputation is enhanced positively, a steady-state order is always reached, and is unique given initial conditions. On the other hand, when negative feedbacks are included, reputation orders become chaotic.

1.4. Search Engines as Reputation Systems

Web search engines, and in particularly the ubiquitous Google, play the part of universal "managers" of reputation. The Google search engine is easily the world's most popular reputation system. Google famously employs the Page-Rank algorithm [3] to rank the importance of pages. Briefly, the importance of a page is the sum of the importance of each page that hyperlinks to it, plus a constant self-importance.

In the Appendix we demonstrate a close relationship between page-rank values and the value of reputation as defined in our model.

Google and the Page-Rank algorithm exemplify well the interaction of "expertise" and "reputation": In response to a search query, Google ranks pages in order of their "reputation" (in fact, their page rank), which is indicative of their "expertise" (in fact, their likelihood to be what was searched for). A page found in a search is likely to acquire new links (for example, when looking for a travel destination, which will later be mentioned in the traveler's blog). It is tacitly assumed (by Google and its users) that this procedure ultimately causes the objectively best pages to be ranked high. Whether or not this is indeed the likely outcome is the subject of our investigation.

1.5. Reputation in Economics and Game Theory

The subject of reputation is extensively discussed in the literature of game theory. It was introduced by Selten in the "chain-store paradox" [11] to mean the belief of players in games that another player takes actions that fall within a certain class, e.g. "aggressive". Kreps and Wilson [8] showed how reputation may affect behavior in Bayesian games where there is uncertainty about players' payoff structure. Reputation is usually used in order to capture strategic choices that a player selects at will (such as honesty or aggresiveness), rather than intrinsic attributes, such as quality of service or expertise in a field, which a player cannot choose at will.

The concept of brand as a carrier of a firm's reputation was put forward by Kreps [7], in the context of moral hazard. Cabral [4] discusses firm reputation as a posterior belief of its customers of the firm's quality level given the firm's history of performance, in the context of whether a firm with a strong brand is well-advised to use the same brand for a new product.

Tadelis[12] as well as Mailath and Samuelson[9] consider reputation as a tradable asset: A firm's reputation is a noisy signal of its competence or effort observable by customers. Firms may trade in reputations, and such trades are only partially observable by customers.

Information cascades [2] study situations in which it is optimal for an individual, having observed the actions of those ahead of her, to follow the behavior of the preceding individual without regard to her own information. Information cascades have been advanced as an explanation of the localized conformity of behavior and the fragility of mass behavior.

1.6. Organization of this paper

Section 2 presents our model.

In Section 3 we analyze the dynamics of the reputation model and the long-term relation between expertise and reputation.

In Section 4 we discuss the rationality of user behavior in reputation systems, investigating whether the behavior ascribed to users in our model conforms to their best strategy.

In Section 5 we summarize and discuss future work.

In the Appendix the reader will find a demonstration of the relation between our model's notion of reputation and the measures used by web search engines to rank search results, as well as supplementary details on the behavior of the model.

2. The Model

We use a model wherein N users repeatedly seek the services of n experts, each expert having a publicly known reputation. The reputations evolve dynamically and represent the aggregate feedback of the users based on their satisfaction with the services they received. The probability of providing satisfactory service to any user request is modeled as the expert's expertise, a fixed but hidden quantity characterizing each expert.

When seeking service, users engage experts according to some selection order, stopping when they are satisfied with the service (or when they have exhausted all experts). The selection order may be simply the experts ordered by descending reputation (the "Reputation Ordered Scheme"), or may also take into account a user's previous experience with experts ("Loyalty Schemes").

2.1. Rounds

Time is discrete: At each integral time a **round** takes place, in each of which each user seeks a service that may be provided by any of the experts. The service provided by an expert may either succeed or fail. Each expert has his own **reputation** and **expertise**. The reputation of expert i is a real-valued function of the discrete time, $r_i = r_i(t)$ representing accumulated user feedback (at round t) on expert i's success rate. The expertise of expert i is his actual success probability $\epsilon_i \in [0, 1]$ in satisfying users requests. At any given round t, the current reputation values $r_1(t), \ldots, r_n(t)$ are common knowledge to all users. On the other hand, the expertise values $\epsilon_1, \ldots, \epsilon_n$ are unknown to the experts and the users.

2.2. Success and Failure Probabilities

The outcome of expert i's service is a random Bernoulli event with probability ϵ_i for success and probability $1 - \epsilon_i$ for failure. This outcome is independent of any other user-expert interaction in any round. However, a repeat service request to an expert in the same round produces the same result as the first request. Therefore there is no point in seeking the service of the same expert more than once in a round.

2.3. Selection Order

The order in which experts are asked in each round is called the selection order. Users follow their selection order until success, or until no more experts are available. The selection order may be based on the reputations publicly known at the start of the round, or on the user's previous experience, or both. A user has no direct information of other users' experience. The selection order may be non-deterministic.

A selection scheme that depends only on the order of experts' reputations (and not, e.g., on actual reputation values) is called **order-based**.

A user that queries an expert in some particular round, is called a **customer** of the expert in that round.

2.4. Reputation Update Rule

At the end of each round, each expert's reputation is updated according to his customers' experience in that round. For every successful service (i.e., for every satisfied customer) in the current round, the expert's reputation is incremented by β, and for every failed service (i.e., a dissatisfied customer) in the current round, it is decremented by $1 - \beta$. The total of all reputation updates for an expert in a round is called the expert's **feedback**. The parameter $0 \leq \beta \leq 1$ is called the **reward/penalty factor**. Note that for $\beta = 1$, only successes are rewarded, while for $\beta = 0$, only failures are penalized.

At the beginning of each round, the previous round's reputation is multiplied by a persistence (discount) factor $0 \leq \alpha \leq 1$. No discounting corresponds to $\alpha = 1$.

We denote the number of customers of expert i in round t by $c_i(t)$, and his feedback in round t by $w_i(t)$. Let $S_i(t) \subset [N], U_i(t) \subset [N]$ be i's set of satisfied and unsatisfied customers, respectively. Then $|S_i(t) \cup U_i(t)| = c_i(t)$, and the full reputation update rule is:

$$w_i(t) = |S_i(t)|\beta - |U_i(t)|(1 - \beta) \tag{2.1}$$

$$r_i(t + 1) = \alpha r_i(t) + w_i(t) \tag{2.2}$$

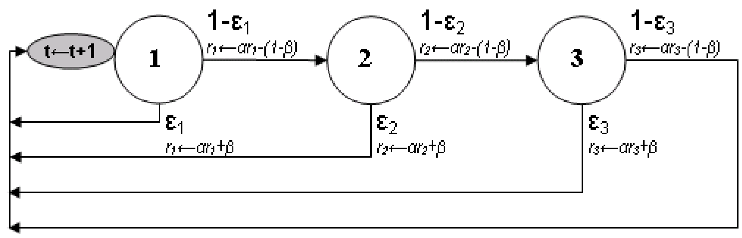

Figure 1: Flow, timing and feedback in the reputation-order scheme

It is easy to express the expectation of expert i's feedback in round t. If he has $c_i(t)$ customers at that round, then

$$\mathbb{E}[w_i(t)] = \beta\epsilon_i \, \mathbb{E}[c_i(t)] - (1-\beta)(1-\epsilon_i) \, \mathbb{E}[c_i(t)] \Rightarrow \tag{2.3}$$

$$\mathbb{E}[w_i(t)] = (\epsilon_i + \beta - 1) \, \mathbb{E}[c_i(t)] \tag{2.4}$$

It follows that the expected feedback of an expert is positive if and only if his expertise ϵ is $\geq 1-\beta$, regardless of the selection order or any other detail.

2.5. Selection Order as a Markov Chain

Figure 1 demonstrates the flow of a user in a round in the reputation-ordered scheme with 3 experts, shown as a Markov chain with feedback side-effects.

The 3 experts are ranked by reputation so that $r_1 > r_2 > r_3$. The diagram represents the flow so long as this order is stable (an order change, if it happens, is noted only at the start of a round, at the $t \leftarrow t+1$ oval). Each of the three states (circles) represents a possible query of one of the experts, with two emanating edges representing failure (emanating right), and success (emanating downwards), with their respective feedback updates.

Similarly, Figure 2 demonstrates the flow of a user in a round in the loyalty scheme with 3 experts, shown as a Markov chain with feedback side-effects.

3. Behavior of the Model

We explore the behavior of the model under various selection schemes. The main question that will interest us is whether the rank order of expert repu-

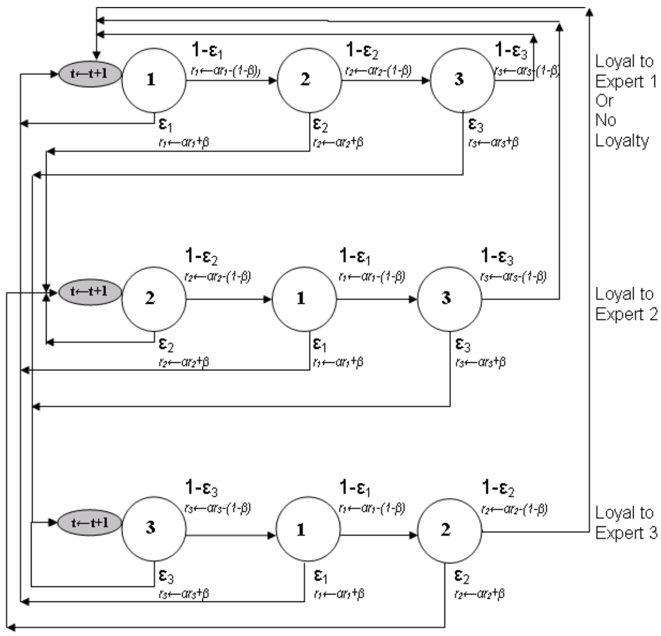

Figure 2: Flow, timing and feedback in the loyalty scheme

tations reaches a steady state in the long run, and if so, to what extent the reputation ranking in the long run reflects the experts' objective levels of expertise $\epsilon_1, \ldots, \epsilon_n$, vs. the initial order of reputations $r_1(0) > \ldots > r_n(0)$.

Since the model is stochastic in nature, the notion of *steady state* should be clarified. To this end we define the notion of *quasi-stability*:

3.1. Quasi-Stability

Definition 1. *A **setting** is a combination of a reputation ranking $r_1(t) > \ldots > r_n(t)$, of experts with expertise levels $\epsilon_1, \ldots, \epsilon_n$ and a selection scheme $\mathcal{S}$ (with reward-penalty factor β).*

*The i'th **pair** within a setting consists of a **leader**, the expert at index i, and the pair's **follower** - the expert with index $i+1$.*

*The i'th pair within a setting is said to be **quasi-stable** (at round t) if the expected reputation of the pair's leader given the setting (as defined in* (2.1)*) is greater or equal to the expected reputation of the pair's follower, at all rounds* *starting at t, i.e.:*

$$\mathbb{E}[r_i(T)] \geq \mathbb{E}[r_{i+1}(T)], \qquad \forall T \geq t \tag{3.1}$$

where the probability space for the expectation is taken over all possible user-expert interactions in rounds after t, ***on the assumption that the expert order does not change****.*

If a pair is not quasi-stable, it is called ***unstable****.*

For order-based schemes, feedback expectations stay constant so long as the reputation order does not change. Therefore in view of (2.2) the definition of quasi-stability is equivalent to:

$$\mathbb{E}[w_i(t)] \geq \mathbb{E}[w_{i+1}(t)] \tag{3.2}$$

Quasi-stability encompasses the notion of stochastic stability of the order between the leader and the follower in a pair: The leader is expected to retain his position in subsequent rounds, perhaps indefinitely. Since the feedback is a random variable, this is not guaranteed: An unlikely lucky streak may close the gap and put the follower in the lead. In the context of our model, a "lucky streak" for an expert would be to have a higher rate of success than his true level of expertise. The probability for this happening vanishes as the number of users N grow without limit, as by Law of Large Numbers the (relative) variance in reputation feedback per round vanishes.

We use the term WITH HIGH PROBABILITY or w.h.p. for short to describe an attribute that, in the scope of a given time frame, is expected to be true with probability $1 - o(1)$ as $N \to \infty$. Thus an equivalent to saying that a pair is quasi-stable is the statement that it is w.h.p. stable, or that the leader's reputation is w.h.p. greater than the follower's.

In contrast, an unstable pair is guaranteed to experience an order change between leader and follower, in a time frame that does not depend on the number of users.

We use the following additional definitions on stability:

Definition 2. *A setting is called **quasi-stable** if all pairs within it are quasi-stable.*

*A setting that is not quasi-stable, i.e. if any pair within it is unstable, is called **unstable**.*

*A pair within a setting is called **two-sided quasi-stable** if it is quasi-stable and if it would be quasi-stable in a setting wherein the pair's leader and follower have traded places.*

*A pair within a setting is called **two-sided unstable** or **chaotic** if it is unstable and if it would be unstable in a setting wherein the pair's leader and follower have traded places.*

Two-sided quasi-stability describes a situation wherein whoever leads retains the lead. Two-sided instability describes a situation wherein no lead is stable.

3.2. The Reputation-Ordered Scheme

In the reputation ordered scheme all users select experts in descending order of their current reputations.

The scheme is "memoryless": Users have no recall of what happened in previous rounds. Therefore it is irrelevant whether they are the same users each round: In memoryless schemes, there is no significance to users' identity, and all users employ the same selection order for experts.

Let us analyze the dynamics of this scheme. Assume the experts are indexed by their reputation order (at round t) $r_1(t) > r_2(t) > \ldots > r_n(t)$ and have respective expertise $\epsilon_1, \ldots, \epsilon_n$. All N users will be the 1st expert's customers, $c_1(t) = N$.

By expectation, $\epsilon_1 c_1(t)$ of the customers will be satisfied, and $(1 - \epsilon_1)c_1(t)$ will be dissatisfied. 1st expert's dissatisfied customers will become 2nd expert's customers. In general, i'th expert's dissatisfied customers will become $(i+1)$'th expert's customers. Formally (and marking by $c_{n+1}(t)$ the number of users who failed with **all** experts):

$$\mathbb{E}[c_{i+1}(t)] = (1 - \epsilon_i)\,\mathbb{E}[c_i(t)], \qquad \forall i \in [n] \tag{3.3}$$

Therefore:

$$\mathbb{E}[c_i(t)] = N \prod_{j=1}^{i-1} (1 - \epsilon_j), \qquad \forall i \in [n+1] \tag{3.4}$$

Under what conditions will the reputation order be quasi-stable?

Theorem 1. *Let n experts be indexed by their reputation order (at round t) $r_1(t) > r_2(t) > \ldots > r_n(t)$ and have respective expertise $\epsilon_1, \ldots, \epsilon_n$. Let n_1 be the smallest index for which $\epsilon_{n_1} = 1$, or let $n_1 = n$ if no such index exists. Then the order is quasi-stable under the* ***reputation-ordered scheme*** *if and only if the following equivalent inequalities apply for each $i \in [n_1 - 1]$:*

1. LEADER'S ADVANTAGE: $\frac{\beta}{1-\epsilon_i}\epsilon_i \geq \epsilon_{i+1}$
2. FOLLOWER'S HANDICAP: $\epsilon_i \geq \frac{1}{\beta+\epsilon_{i+1}}\epsilon_{i+1}$
3. RECIPROCAL DIFFERENCE: $\frac{1}{\epsilon_i} - \frac{\beta}{\epsilon_{i+1}} \leq 1$

PROOF. The expected feedback of each expert (see (2.4)) is:

$$\mathbb{E}[w_i(t)] = (\epsilon_i + \beta - 1)\,\mathbb{E}[c_i(t)], \qquad \forall i \in [n] \tag{3.5}$$

Since this scheme is order-based, it is quasi-stable if $\forall i \in [n-1] \quad \mathbb{E}[w_i(t)] \geq \mathbb{E}[w_{i+1}(t)]$. If $i \geq n_1$ the inequality holds trivially, as both expectations are zero. Otherwise the condition translates to:

$$(\epsilon_i + \beta - 1)\,\mathbb{E}[c_i(t)] \geq (\epsilon_{i+1} + \beta - 1)\,\mathbb{E}[c_{i+1}(t)] \tag{3.6}$$

By (3.3):

$$(\epsilon_i + \beta - 1)\,\mathbb{E}[c_i(t)] \geq (\epsilon_{i+1} + \beta - 1)(1 - \epsilon_i)\,\mathbb{E}[c_i(t)] \tag{3.7}$$

Dividing both sides by $\mathbb{E}[c_i(t)]$ (a positive number as $i < n_1$):

$$\epsilon_i + \beta - 1 \geq (\epsilon_{i+1} + \beta - 1)(1 - \epsilon_i) \tag{3.8}$$

Which, by rearrangement, leads to each of the three equivalent inequalities.

Corollary 1. *Under the reputation-ordered scheme, the quasi-stability of a pair within a setting depends only on the expertise of the pair's leader and follower.*

Under the reputation-ordered scheme, rank certainly has its privileges: Considering the situation where experts get positive feedback only ($\beta = 1$), Corollary 2 shows that being ahead in reputation confers on an expert with expertise ϵ an advantage factor of $1/(1-\epsilon)$ in expertise over followers:

Corollary 2. *Under the conditions of Theorem 1 with a reward-only scheme ($\beta = 1$), quasi-stability requires the following equivalent inequalities for each $i \in [n_1 - 1]$:*

1. LEADER'S ADVANTAGE: $\frac{1}{1-\epsilon_i}\epsilon_i \geq \epsilon_{i+1}$
2. FOLLOWER'S HANDICAP: $\epsilon_i \geq \frac{1}{1+\epsilon_{i+1}}\epsilon_{i+1}$
3. RECIPROCAL DIFFERENCE: $\frac{1}{\epsilon_i} - \frac{1}{\epsilon_{i+1}} \leq 1$

It is worth noting that this advantage becomes insurmountable when $\epsilon > 1/2$, for in this case $\frac{1}{1-\epsilon}\epsilon > 1$, and therefore the pair is quasi-stable against any follower.

A different perspective on this is to consider the disadvantage of being second: Corollary 2 shows that this inflicts a handicap factor of $1/(1+\epsilon)$ on an expert with expertise ϵ. In other words, an expert can reasonably expect to overtake a leader over whom his advantage in expertise is greater than his handicap. Again, it is worth noting that $\frac{1}{1+\epsilon}\epsilon \leq 1/2$. That is, a follower can never expect to overtake a leader with expertise of more than 1/2.

A pair whose leader and follower have the same expertise is always two-sided quasi-stable, i.e. between equals, the order determined by initial conditions is preserved. Generally, by Corollary 2 (3), two-sided quasi-stability exists iff:

$$-1 \leq \frac{1}{\epsilon_i} - \frac{1}{\epsilon_{i+1}} \leq 1 \tag{3.9}$$

On the other hand, when negative feedback applies ($\beta = 0$), chaos reigns, as spelled out by Corollary 3:

Corollary 3. *Under the conditions of Theorem 1 with a penalty-only scheme ($\beta = 0$), quasi-stability is possible only if $\epsilon_1 = 1$ or $\epsilon_i = 0$ for all but the first expert. (Substituting $\beta = 0$ in Theorem 1 (3) we derive either $\epsilon_{i+1} = 0$ or $\epsilon_i \geq 1$ which is impossible as we assumed $i < n_1$).*

As negative feedback is gradually added (i.e. as β gradually decreases below 1), the above situation changes in several respects. By Theorem 1:

- The leader's advantage decreases from $1/(1-\epsilon)$ to $\beta/(1-\epsilon)$, disappearing (i.e. equaling 1) at the critical point $\beta = 1-\epsilon$.
- A leader with expertise greater than $1/(1+\beta)$ has an unassailable position.
- The follower's handicap decreases from $1+\epsilon$ to $\beta+\epsilon$, disappearing (i.e. equaling 1) at the critical point $\beta = 1-\epsilon$.
- Note (see (2.4)) that experts with expertise above the critical $1-\beta$ have positive feedback expectation, while experts with expertise below the critical value have negative feedback expectation.
- A pair with leader expertise of less than $1-\beta$ and follower expertise of at least $1-\beta$ is always unstable, while in reverse order it is always quasi-stable.
- Two-sided quasi-stability is possible only if both follower and leader have expertise greater or equal to $1-\beta$.

Corollary 4. *Under the conditions of Theorem 1, two-sided instability, i.e. chaos, is possible only if an expert pair exists such that* $\epsilon_i < 1-\beta$ *as well as* $\epsilon_{i+1} < 1-\beta$.

t

Proof. By Theorem 1 two-sided instability requires both of the following inequalities to be true:

$$\frac{1}{\epsilon_i} - \frac{\beta}{\epsilon_{i+1}} > 1 \tag{3.10}$$

$$\frac{1}{\epsilon_{i+1}} - \frac{\beta}{\epsilon_i} > 1 \tag{3.11}$$

Multiplying both sides of (3.11) by β and adding (3.10) results in:

$$
\begin{aligned}
& & \frac{1}{\epsilon_i} - \frac{\beta}{\epsilon_{i+1}} + \frac{\beta}{\epsilon_{i+1}} - \frac{\beta^2}{\epsilon_i} > 1 + \beta \\
\Rightarrow & & \frac{1}{\epsilon_i}(1-\beta^2) > 1 + \beta \\
\Rightarrow & & \frac{1}{\epsilon_i}(1-\beta) > 1 \\
\Rightarrow & & 1 - \beta > \epsilon_i
\end{aligned}
$$

Similarly, multiplying both sides of (3.10) by β and adding (3.11) leads to $1 - \beta > \epsilon_{i+1}$.

Note, though, that Corollary 4 states a necessary, but not a sufficient condition for two-sided instability. For example, $\beta = 0.5, \epsilon_1 = 0.4, \epsilon_2 = 0.1$ is quasi-stable, and becomes chaotic only for $\beta < 0.15$.

3.3. Steady State Orders

3.3.1. Definitions

Previously in this section we defined the notion of quasi-stability and instability of a setting, and formulated criteria for it in reputation-ordered settings. Assuming unstable pairs will eventually flip their leader-follower order, and assuming quasi-stable pairs to be stable[2], we ask:

1. Will the setting converge to a steady state, i.e. to a quasi-stable setting?
2. If so, given the setting, what will the steady state setting be?
3. Given an initial setting, is the steady state unique?

We will answer these questions for the reputation-ordered scheme and for other schemes, but first we need some definitions:

Definition 3. *A selection scheme is called* ***regular*** *if in settings that employ it the quasi-stability of a pair depends only on the expertise values of the leader and of the follower.*

[2] Meaning that we neglect the $o(1)$ (as $N \to \infty$) probability for their experiencing an order change

By Corollary 1 the reputation-ordered scheme is regular.

Definition 4. *The notation* $\epsilon_1 \prec_s \epsilon_2$ *means that a pair with leader expertise* ϵ_1 *and with follower expertise* ϵ_2 *is unstable under the regular selection scheme* $\mathcal{S}$.

The notation $\epsilon_1 \not\prec_s \epsilon_2$ *means that a pair with leader expertise* ϵ_1 *and with follower expertise* ϵ_2 *is quasi-stable under the regular selection scheme* $\mathcal{S}$.

A setting is called ***chaotic*** *with respect to selection scheme* $\mathcal{S}$ *if it includes a* ***chaotic pair****: a pair of experts with expertise* ϵ_1, ϵ_2 *such that* $\epsilon_1 \prec_s \epsilon_2$ *and* $\epsilon_2 \prec_s \epsilon_1$.

The instability operator is transitive: $\epsilon_1 \prec_s \epsilon_2$ and $\epsilon_2 \prec_s \epsilon_3 \Rightarrow \epsilon_1 \prec_s \epsilon_3$. This follows from the definition of instability (Definition 1).

Recall that in a regular setting the quasi-stability or instability of an expert pair depends only on their respective expertise levels. Therefore, in a non-chaotic setting, the instability operator is a *partial order* on the levels of expertise.

3.3.2. Existence and Uniqueness

Starting at some initial setting in which the experts are arranged by descending reputation and numbered from 1 to n, $r_1(0) > \ldots > r_n(0)$ with respective expertise $\epsilon_1, \ldots \epsilon_n$, the initial order is described by the permutation $(1, 2, \ldots, n)$ of [n]. If this setting is not quasi-stable, then at some future round, two neighboring experts in an unstable pair will trade places. E.g. if the leader and follower at pair i trade places, the resulting permutation is $(1, \ldots, i-1, i+1, i, i+2, \ldots, n)$.

Let us mark the expert order at round t by the permutation $\pi(t) = (\pi_1(t), \pi_2(t), \ldots, \pi_n(t))$. $\pi_i(t)$ is the expert at the i'th position at round t.

The meaning of reaching a steady state is that there exists a round T with a quasi-stable setting, i.e.: $r_{\pi_1(T)}(T) > r_{\pi_2(T)}(T) > \ldots > r_{\pi_n(T)}(T)$ and $\epsilon_{\pi_1(T)} \not\prec_s \epsilon_{\pi_2(T)} \not\prec_s \ldots \not\prec_s \epsilon_{\pi_n(T)}$.

The path from the initial setting to the steady-state setting consists of successive swapping of the leader and follower in unstable pairs, until no more such swapping is possible.

We now derive a general result in the theory of partially ordered sets (=posets). This result will be applied below to the instability operator, on the way to answering our questions regarding steady states.

Let $(P, \prec)$ be a finite poset. If $x, y \in P$ and either $x \prec y$ or $y \prec x$ holds, we say that x, y are **comparable**. Otherwise we say they are incomparable and write $x \parallel y$. Let $\pi = (x_1, x_2, \ldots, x_n)$ be an ordering of P's elements. A **swap** changes this permutation to $(x_1, x_2, \ldots, x_{i-1}, x_{i+1}, x_i, x_{i+2}, \ldots, x_n)$ for some index i. This swap is **permissible** if $x_i \prec x_{i+1}$. We say that a permutation σ of P's elements is *reachable* from π if it is possible to move from π to σ through a sequence of permissible swaps. A permutation of P's elements is called *terminal* if no swap is permissible. It is an easy observation that starting from any permutation of P, any series of permissible swaps is finite, since every two elements can be swapped at most once.

Theorem 2. *For every permutation π of the elements of a finite poset $(P, \prec)$ there is exactly one terminal permutation reachable from π.*

Proof. Let τ be a terminal permutation that is reachable from π. The uniqueness of τ is proved by providing a criterion, depending only on π, as to which pairs of elements appear in the same order in π and τ and which are reversed.

An (x, y)-*fence* in π is a sequence $x = z_1, z_2, \ldots z_k = y$ that appear in this order (not necessarily consecutively) in π such that $z_\alpha \parallel z_{\alpha+1}$ for every $\alpha \in [k-1]$.

Clearly, if $x \parallel y$ no permissible swap can change the relative order of x and y. Consequently:

- No sequence of permissible swaps can change the relative order of x and y if an (x, y)-*fence* exists.
- No sequence of permissible swaps can create or eliminate an (x, y)-*fence*.

We say that (x, y) is a *critical pair* in π if (i) $x \prec y$, (ii) x precedes y in π and, (iii) there is no (x, y)-*fence* in π.

We now assert and prove the criterion for whether any two elements x and y in π, with x preceding y, preserve or reverse their relative order in a terminal permutation:

1. If $y \prec x$, the order is preserved.
2. If there exists an (x, y)-*fence*, the order is preserved.
3. Otherwise, i.e. if (x, y) is a critical pair, the order is reversed.

The first element of the criterion is trivial and the second has already been dealt with. It remains to show the third and last element: Since an (x, y)-*fence* cannot be created or eliminated by permissible swaps, an equivalent statement to this claim is that a permutation τ with a critical pair cannot be terminal. We prove this by induction on the number of elements in τ separating x and y:

At the base of induction, if x and y are neighbor elements, the assertion is true since as $x \prec y$ the permutation is not terminal. Now let k be the number of elements separating x and y, and the induction hypothesis is that if the number of elements separating a pair is less than k it cannot be critical.

Let z be an element between x in y in τ. Assume $x \prec z$. Then by the induction hypothesis there exists an (x, z)-*fence*. Now consider the relation between z and y: $z \parallel y$ is impossible, because then y could be concatenated to the (x, z)-*fence* to form a (x, y)-*fence*, contrary to the assumption that (x, y) is a critical pair. Similarly, $z \prec y$ would by the induction hypothesis prove the existence of a (z, y)-*fence*, but this is impossible as it could be concatenated to the (x, z)-*fence* to form an (x, y)-*fence*. This leaves $y \prec z$ as the only possibility.

In summary $x \prec z \Rightarrow y \prec z$.

By similar reasoning $z \prec y \Rightarrow z \prec x$.

Furthermore, the possibility $x \parallel z$ together with $z \parallel y$ can be dismissed as constituting an (x, y)-*fence*, leaving just two possible scenarios satisfied by each z between x and y:

- z is "small", i.e. $z \prec x$ and $z \prec y$.
- z is "large", i.e. $x \prec z$ and $y \prec z$.

Since τ is terminal, x's immediate neighbor must be "small", and y's immediate neighbor must be "large". Between these two, there must exist two consecutive elements z_1, z_2 such that z_1 is "small" and z_2 is "large". But this leads to a contradiction as $z_1 \prec x \prec z_2 \Rightarrow z_1 \prec z_2$, which is not terminal. Therefore (x, y) cannot be critical, completing the proof by induction that a terminal permutation cannot have a critical pair.

Thus the demonstration of the criterion for the terminal order of any element pair in π is completed, thereby also showing that the terminal permutation is unique.

Armed with this general result on posets, we derive a general theorem regarding the existence and uniqueness of steady-state reputation orders:

Theorem 3. *Given a setting with a regular selection scheme:*

1. *If the setting is not chaotic it will converge to a quasi-stable setting.*
2. *If the setting converges to a quasi-stable setting, it will w.h.p. converge to the same setting.*

Proof. The theorem is an immediate consequence of Theorem 2 by noting:

1. The instability relation under a non-chaotic, regular setting defines a partial order on the levels of expertise.
2. Order changes in a setting are w.h.p. between some unstable expert pair.
3. In a steady-state order all expert pairs are quasi-stable, i.e. none are unstable.

As consequence of Theorem 3 there exists a simple algorithm to determine the steady-state order arising out of any given setting: Calculate the quasi-stability or instability of all pairs in a setting. Switch the order of any unstable pair. Repeat until reaching a setting with no unstable pair.

Example 1. *In the reward-only reputation- ordered scheme, let $n = 4$, with initial setting*

$$(\epsilon_1 = \frac{1}{4}, \epsilon_2 = \frac{1}{3}, \epsilon_3 = \frac{1}{2}, \epsilon_4 = 1)$$

This initial setting is already quasi-stable. A slightly different initial setting

$$(\epsilon_1 = \frac{1}{4}, \epsilon_2 = \frac{1}{2}, \epsilon_3 = \frac{1}{3}, \epsilon_4 = 1)$$

eventually settles on the quasi-stable

$$(\epsilon_2 = \frac{1}{2}, \epsilon_4 = 1, \epsilon_1 = \frac{1}{4}, \epsilon_3 = \frac{1}{3})$$

If β is lowered to $\frac{3}{4}$, the behavior changes: Both initial settings converge to the naturally-ordered sequence:

$$(\epsilon_4 = 1, \epsilon_2 = \frac{1}{2}, \epsilon_3 = \frac{1}{3}, \epsilon_1 = \frac{1}{4})$$

3.4. Indifference to Expert Delays

Our model (Section 2) posits that all queries and replies occur within a single integral unit of time: a round.

We now wish to expand the model to allow flexible timing: Let each expert i have a delay of δ_i between query and answer (or, between query time and the time at which satisfaction or disappointment of the user manifests itself).

By example, Figure 3 is the flow diagram and Markov chain of the 3-expert reputation-ordered scheme with generalized delays. It is a generalization of the basic model's flow diagram given in Figure 1 in which $\delta_1 = 1, \delta_2 = \delta_3 = 0$.

Clearly, a similar Markov chain and flow diagram exists for every selection order scheme: For every possible and relevant history of the scheme, there is a subchain of n expert nodes, ordered by the selection order. Each expert node, for an expert with expertise ϵ, has two emanating edges in the chain: First, a failure edge with probability $1 - \epsilon$, leading to the next expert in the selection order (or, for the last expert, back to the pre-round delay). Second, a success edge with probability ϵ leads to the pre-round delay of the updated user history.

An example is given in Figure 2 for the 3-expert loyalty scheme, in which there are 3 expert nodes for each of the 3 relevant histories: the 3 distinct user loyalty states.

In the general case, there are $M = mn$ expert nodes in the flow diagram, where m is the number of distinct user histories, and M delays, one per expert node.

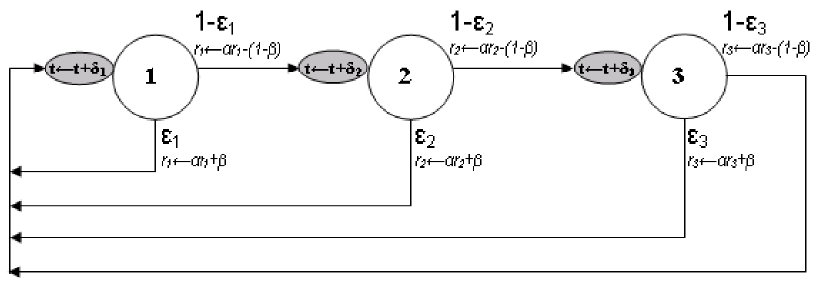

Figure 3: Flow, timing and feedback in the reputation-ordered scheme with generalized delays

So long as the reputation order of the experts does not change, each participating user may be seen as taking a random walk through the Markov chain. An order change changes the Markov chain.

We now claim that the delays are of no importance to the behavior of the model:

Theorem 4. *Let there be n experts ordered by their reputations $r_1 > \ldots > r_n$ and M expert nodes in the selection order's Markov chain. Let $\Delta_1 = (\delta_{1,1}, \ldots, \delta_{1,M})$ be a set of expert delays for each expert node and $\Delta_2 = (\delta_{2,1}, \ldots, \delta_{2,M})$ be another such set. Then the values of reputation expert feedback under the two delay sets are proportional. Specifically, there exist constants for each delay set, $C_{\Delta_1}, C_{\Delta_2}$ such that for each expert i:*

$$w_i(t;\Delta_1)C_{\Delta_1} = w_i(t;\Delta_2)C_{\Delta_2} \tag{3.12}$$

Where the notation $w_i(t;\Delta)$ generalizes (2.1)*: The feedback of expert i at round t with set of expert delays Δ.*

Proof. Assume that in addition to node delays, there is a fixed edge delay $\delta > 0$. Consider the probability of finding the user in some particular edge, conditional on her being at any edge: This probability depends only on the Markov chain's graph and transition probabilities, and is independent of the node delays (or of δ).

With each edge is associated a reputation feedback. Per unit time, the expected reputation feedback from any particular edge is the edge feedback multiplied by the probability of finding the user at that particular edge during a unit of time. Since the edge feedback is constant, and the edge probabilities are in fixed proportions to each other, the theorem follows for any particular δ. In particular, it holds while $\delta \to 0$, and so holds in the limit, with no edge delays.

Corollary 5. *The quasi-stability, instability and other behavioral aspects of a setting are independent of expert delays. The behavior under different sets of delays is identical with a suitable scaling of the time.*

3.5. The Loyalty Scheme

We now introduce a selection scheme in which users have memory:

Each round, each user remembers the expert that succeeded for her in the previous round (if there was such an expert), and selects him first in the current round. If this expert fails, the user will revert to the reputation-ordered scheme, i.e. the second expert selected will be the highest-reputation so far unselected, etc.

We call it the **loyalty scheme**. The loyalty scheme attributes limited recall to the user: She remembers what worked out for her last time, but not more. Clearly the loyalty scheme models many real-world situations better than the memoryless reputation-ordered scheme. Our analysis will show that this limited recall significantly affects the quantitative and qualitative behavior of the model.

However, to avoid some complexities of the loyalty scheme, we first concentrate on a variation that is "well-behaved", which we will call the **dual loyalty scheme**. In the dual loyalty scheme, as in the loyalty scheme, the user first queries the expert who succeeded for her last time, and if this expert fails, will revert to the reputation-ordered scheme, querying all experts in order of descending reputation, but **without skipping over the first expert** to which the user was loyal. This necessarily means that the user may query this expert twice. So, for this modified scheme, we waive our standing assumption that asking the same question twice of an expert always yields the same answer.

This waiver, while admittedly artificial, has the merit of simplifying the analysis of the scheme, and, as we will see later, of making it regular.

The (unmodified) loyalty scheme is analyzed in the Appendix.

Theorem 5. *Let n experts be indexed by their reputation order (at round t) $r_1(t) > r_2(t) > \ldots > r_n(t)$ and have respective expertise $\epsilon_1, \ldots, \epsilon_n$ all of which are < 1. Then the order is quasi-stable under the **dual loyalty scheme** if and only if for each $i \in [n-1]$:*

$$\frac{\epsilon_i + \beta - 1}{(1-\epsilon_i)^2} \geq \frac{\epsilon_{i+1} + \beta - 1}{1 - \epsilon_{i+1}} \tag{3.13}$$

Proof. Observe that at each round a user is in one of $n+1$ states: State $i \in [n]$ when the user is "loyal" to expert i, and state 0 when the user is loyal to no one. The probability to be in each of the states is the result of a random walk among states given the transition probabilities between the states. We calculate these probabilities from the scheme rules:

Let the probability to be in state $i \in [0, n]$ be given by p_i, and the probability of a transition from state i to state j be given by $v_{i,j}$. Then by the rules of the scheme:

$$v_{i,j} = \begin{cases} \epsilon_j \prod_{k=1}^{j-1} (1-\epsilon_k) & i = 0, j \neq 0 \\ \prod_{k=1}^{n} (1-\epsilon_k) & i = 0, j = 0 \\ (1-\epsilon_i) v_{0,j} & i \neq 0, j \neq i \\ \epsilon_i + (1-\epsilon_i) v_{0,j} & i \neq 0, j = i \end{cases} \tag{3.14}$$

Given that $p_i = \sum_{k=0}^{n} p_k v_{k,i}$, we can calculate p_0:

$$\begin{aligned} p_0 &= p_0 v_{0,0} + p_1(1-\epsilon_1) v_{0,0} + \ldots + p_n (1-\epsilon_n) v_{0,0} = \\ &= v_{0,0} \Big[p_0 + \sum_{k=1}^{n} p_k (1-\epsilon_k) \Big] = \\ &= v_{0,0} \Big[1 - \sum_{k=1}^{n} p_k \epsilon_k \Big] \end{aligned}$$

As well as p_i, $\forall i > 0$:

$$\begin{aligned} p_i &= p_i\epsilon_i + p_0 v_{0,i} + p_1(1-\epsilon_1)v_{0,i} + \ldots + p_n(1-\epsilon_n)v_{0,i} = \\ &= p_i\epsilon_i + v_{0,i}\Big[p_0 + \sum_{k=1}^{n} p_k(1-\epsilon_k)\Big] = \\ &= p_i\epsilon_i + v_{0,i}\Big[1 - \sum_{k=1}^{n} p_k\epsilon_k\Big] = \\ &= \frac{1}{1-\epsilon_i} v_{0,i}\Big[1 - \sum_{k=1}^{n} p_k\epsilon_k\Big] \end{aligned}$$

Let us denote $1 - \sum\limits_{k=1}^{n} p_k\epsilon_k$ by M and conclude:

$$p_i = \begin{cases} M \prod\limits_{k=1}^{n} (1-\epsilon_k) & i = 0 \\ M \frac{\epsilon_i}{1-\epsilon_i} \prod\limits_{k=1}^{i-1} (1-\epsilon_k) & i \neq 0 \end{cases} \tag{3.15}$$

Solving for M gives $M = \left(1 + \sum\limits_{i=1}^{n} \left[\frac{\epsilon_i^2}{1-\epsilon_i} \prod\limits_{k=1}^{i-1} (1-\epsilon_k)\right]\right)^{-1}$. However, we do not need this value, since it is enough to know the ratios between the probabilities, where we find, for each $i \in [n-1]$:

$$\frac{\frac{\epsilon_i}{(1-\epsilon_i)^2}}{p_i} = \frac{\frac{\epsilon_{i+1}}{1-\epsilon_{i+1}}}{p_{i+1}} \tag{3.16}$$

We observe that for a user to arrive at state i ($i > 0$), she should be a satisfied customer of expert i. Therefore $\mathbb{E}[c_i(t)] = \frac{Np_i}{\epsilon_i}$. We are therefore ready to formulate the criterion for quasi-stability of the dual loyalty scheme:

Since the scheme is order-based, (3.2) defines the criterion for quasi-stability. Combining (2.4) and the above, the i'th pair is quasi-stable iff:

$$\frac{\epsilon_i + \beta - 1}{\epsilon_i} p_i \geq \frac{\epsilon_{i+1} + \beta - 1}{\epsilon_{i+1}} p_{i+1} \tag{3.17}$$

Finally, substituting (3.16):

$$\frac{\epsilon_i + \beta - 1}{(1-\epsilon_i)^2} \geq \frac{\epsilon_{i+1} + \beta - 1}{1-\epsilon_{i+1}} \tag{3.18}$$

Note that the requirement that all expertise levels be smaller than 1 is necessary, because if any experts have expertise of 1, the only possible steady-state is where **all** users are loyal to one of these experts.

Corollary 6. *The dual loyalty scheme is regular. This is directly observed from* (3.13).

Corollary 7. *Under the terms of Theorem 5 with a reward-only scheme (*$\beta = 1$*), quasi-stability requires the following equivalent inequalities for each* $i \in [n-1]$*:*

1. $\frac{1}{(1-\epsilon_i)^2}\epsilon_i \geq \frac{1}{1-\epsilon_{i+1}}\epsilon_{i+1}$
2. LEADER'S ADVANTAGE: $\frac{1}{1-\epsilon_i+\epsilon_i^2}\epsilon_i \geq \epsilon_{i+1}$
3. RECIPROCAL DIFFERENCE: $\frac{1}{\epsilon_i} - \frac{1}{\epsilon_{i+1}} \leq 1 - \epsilon_i$

Comparing these results with the corresponding results for the reputation-ordered scheme (Corollary 2), we observe that the loyalty property diminishes the value of a lead in reputation, although it does not nullify it: The leader's advantage is reduced from a factor of $1/(1-\epsilon)$ to $1/(1-\epsilon+\epsilon^2)$ and the reciprocal difference from 1 to $1-\epsilon$.

However, unlike in the reputation-ordered scheme where a leader's position with expertise of above $\frac{1}{2}$ is unassailable, loyalty does not allow for unassailable positions: A leader of expertise ϵ will be overtaken by a follower of expertise $\frac{\epsilon}{1-\epsilon+\epsilon^2}$ which is not greater than 1 and so feasible.

Two-sided quasi-stability is still feasible, but the window for it is narrower. The analogue of (3.9) is:

$$-(1-\epsilon_{i+1}) \leq \frac{1}{\epsilon_i} - \frac{1}{\epsilon_{i+1}} \leq 1 - \epsilon_i \tag{3.19}$$

In addition, we observe that the penalty-only scheme ($\beta = 0$) is, like in the reputation-ordered scheme, chaotic in any setting.

Example 2. *Let the initial setting in a dual loyalty scheme (*$n = 8, \beta = 1$*) be:*

$$(\epsilon_1 = 0.1, \epsilon_2 = 0.2, \epsilon_3 = 0.3, \epsilon_4 = 0.4,$$
$$\epsilon_5 = 0.5, \epsilon_6 = 0.6, \epsilon_7 = 0.7, \epsilon_8 = 0.8)$$

This setting will eventually settle on the quasi-stable

$$(\epsilon_4 = 0.4, \epsilon_5 = 0.5, \epsilon_6 = 0.6, \epsilon_7 = 0.7,$$
$$\epsilon_8 = 0.8, \epsilon_3 = 0.3, \epsilon_2 = 0.2, \epsilon_1 = 0.1)$$

4. The Rationality of Using Reputation as an Indicator of Expertise

As part of the model, we ascribed to each user a selection order of experts in which she queries experts in descending order of their reputation (barring private information about experts' history, as in the loyalty scheme).

The justification for employing such a selection scheme is that reputation is a positive indicator of expertise, i.e. that between any two experts, chances are 50% or better that the expert with the higher reputation has the higher expertise. But is this indeed the case? In other words, can a user, who knows that reputation is tallied through the feedback of a community of other users, each using some known or unknown selection rule, make a reasoned deduction that reputation signals expertise?

The question may appear puzzling in light of the results we have obtained, showing that often it is the less accomplished expert that is able to hold an indefinite lead in reputation over his betters, but to be aware of a possibility is not the same as knowing that it indeed occurred.

The question of rationality may be posed as follows: Suppose a particular user is aware of the expert selection methods employed by all other users. If in aggregate a community of users is more likely to reward the "losers" than the "winners" with a high reputation, the user would do well *not* to follow their collective advice, and eschew the high-reputation experts. However, if the community's behavior is sufficiently "well-behaved" to exclude such possibilities, using reputation as a signal for expertise is rationally justified.

That user communities may conceivably *not* be "well-behaved" is shown by the following example: Let all users have total recall of all their previous expert interactions, and let them each base their selection exclusively on expert success percentage, but using a *reverse* order: They give precedence to the expert that *failed* them most. Clearly such a scheme would reward the worst experts with the most customers, an advantage that may easily outweigh their lower success percentage and so provide the worst experts with the higher reputation feedback.

The anomaly in the above example is the irrational behavior of the users with their private information, i.e. their own experience: It makes no sense for them to prefer the experts that failed them most. We shall show that excluding irrational user behavior with their own experience is enough to make reputation a reliable signal of expertise:

In doing so, we want to allow the most general schemes through which users consider their experience with an expert: A user may choose to remember all previous encounters, or only the most recent m encounters, and she may attach significance to the order of experiences, e.g. A user who remembers the past two encounters with an expert, only one of which was successful, may value the experience higher if the success was on the most recent encounter, rather than in the penultimate one. However, for the valuation to be rational, a user must not value failure higher than success in any *particular* encounter, i.e. if a user remembers the most recent m encounters, and encounter $k \in [m]$ is the k'th most recent, then, changing that experience from success to failure, all other parameters held constant (i.e. all encounters except k, all experiences with other experts, and all expert reputations), may *not* advance the expert in the user's selection order. Briefly, in a rational user's selection order, an expert does not gain by failing in any particular trial.

As advancing in the selection order means a greater probability of the user becoming the expert's customer, we define this rationality requirement as *experience-monotonicity* of that probability:

Definition 5. *Let a user remember her past m encounters with an expert. Let her experience be $Z \subset [m]$, such that $i \in Z$ iff the i'th most recent encounter was successful. Let $C(Z)$ be the probability for the user, with experience Z before some round to become a customer of the expert in that round. A selection order, and a user that employs it, are called* ***experience-monotone*** *if:*

- *The order depends on nothing other than the current expert reputations and on a (possibly partial) recall of experts' success-failure ratio.*

- *for each pair of expert experiences* $Z1$ *and* $Z2$, $Z1 \subset Z2 \Rightarrow C(Z1) \leq C(Z2)$.

All selection orders previously considered are experience monotone: Clearly the loyalty scheme is experience monotone, since a recall of a previous round's success moves an expert to first position. So are all selection orders that have no recall. Note that the monotonicity requirement is on experience alone, and does not rule out taking a non-monotone, and apparently illogical view of reputation.

The (rather weak) restriction of experience monotonicity allows us to prove the following general result:

Theorem 6. *Observing the reputation of two experts (at some round* t*), the expert with the higher reputation is likely (with probability* $\geq 50\%$*) to have the higher expertise, provided:*

- *All users are experience-monotone.*

- *There is no prior information on expertise.*

- *There is no additional information showing that the reputation difference was previously (smaller* t*) larger or that it is smaller in the future (bigger* t*).*

Proof. **Lemma 1.** *If all users are experience-monotone:*

1. *For each expert, the expected number of customers in a round is monotonically non decreasing in his expertise.*
2. *For each expert, the expected reputation feedback in a round is monotonically increasing in his expertise.*

Proof. The second part of the lemma follows from the first part, recalling the definition of reputation feedback: $\mathbb{E}[w_i(t)] = (\epsilon_i + \beta - 1)\,\mathbb{E}[c_i(t)]$ and noting that

$\mathbb{E}[c_i(t)]$ is non-negative. We therefore prove the first part, which may be written formally as:

$$\frac{\partial}{\partial \epsilon_i}\,\mathbb{E}[c_i(t)] \geq 0 \qquad (0 \leq \epsilon_i \leq 1) \tag{4.1}$$

Note now that the *a priori* probability for a user to have a particular experience Z depends on the expertise: The probability that a user who had m encounters with an expert of expertise ϵ, had a particular experience Z equals $\epsilon^{|Z|}(1-\epsilon)^{m-|Z|}$.

We first show that, for each user, the probability of being a customer of the expert with expertise ϵ is non-decreasing in ϵ, when this probability is summed over all possible user experiences, with each experience given its *a priori* probability:

Let $R \equiv \{r_1, ..., r_n\}$ denote the vector of expert reputations, and let m_X be the number of encounters in user X's experience, and let $Z \subset [m_X]$ be X's experience with expert A, such that $i \in Z$ iff X's i'th most recent encounter with A was successful. Denote by $C_{X,A}(R, t, Z, m_X)$ the probability of user X being expert A's customer (at round t) when expert reputations are given by R and x's experience with A is Z. In the following we hold X, A, R, t and m_X constant, therefore we can and will write $C(Z)$ as shorthand for $C_{X,A}(R, t, Z, m_X)$.

Summing over all experiences $Z \subset [m]$ weighed by their *a priori* probability, we aim to show that, for each user, and provided $C(Z)$ satisfies monotonicity:

$$\frac{\partial}{\partial \epsilon} \sum_{Z \subset [m]} C(Z)\epsilon^{|Z|}(1-\epsilon)^{m-|Z|} \geq 0 \qquad (0 \leq \epsilon \leq 1) \Rightarrow \tag{4.2}$$

$$\sum_{Z \subset [m]} \big[|Z| - m\epsilon\big] C(Z)\epsilon^{|Z|-1}(1-\epsilon)^{m-|Z|-1} \geq 0 \qquad (0 \leq \epsilon \leq 1) \tag{4.3}$$

We will prove this by induction over m. For $m = 0$ the sum is empty so (4.3) is trivially true. Assume (4.3) for $m-1$ as the induction hypothesis, and expand (4.3) into two sums, the first over subsets containing m, the second over subsets not containing m:

$$\sum_{Z\subset[m]} \big[|Z| - m\epsilon\big] C(Z)\epsilon^{|Z|-1}(1-\epsilon)^{m-|Z|-1} =$$

$$\sum_{Z\subset[m-1]} \big[|Z| + 1 - m\epsilon\big] C(Z\cup\{m\})\epsilon^{|Z|}(1-\epsilon)^{m-|Z|-2} +$$

$$+ \sum_{Z\subset[m-1]} \big[|Z| - m\epsilon\big] C(Z)\epsilon^{|Z|-1}(1-\epsilon)^{m-|Z|-1} =$$

$$\epsilon \sum_{Z\subset[m-1]} \big[|Z| - (m-1)\epsilon\big] C(Z\cup\{m\})\epsilon^{|Z|-1}(1-\epsilon)^{m-|Z|-2} +$$

$$+ \epsilon(1-\epsilon) \sum_{Z\subset[m-1]} C(Z\cup\{m\})\epsilon^{|Z|-1}(1-\epsilon)^{m-|Z|-2} +$$

$$+ (1-\epsilon) \sum_{Z\subset[m-1]} \big[|Z| - (m-1)\epsilon\big] C(Z\cup\{m\})\epsilon^{|Z|-1}(1-\epsilon)^{m-|Z|-2} +$$

$$- \epsilon(1-\epsilon) \sum_{Z\subset[m-1]} C(Z)\epsilon^{|Z|-1}(1-\epsilon)^{m-|Z|-2} =$$

$$\epsilon A + (1-\epsilon)B + \epsilon(1-\epsilon) \sum_{Z\subset[m-1]} \big[C(Z\cup\{m\}) - C(Z)\big]\epsilon^{|Z|-1}(1-\epsilon)^{m-|Z|-2}$$

Where:

$$A \equiv \sum_{Z\subset[m-1]} \big[|Z| - (m-1)\epsilon\big] C(Z\cup\{m\})\epsilon^{|Z|-1}(1-\epsilon)^{m-|Z|-2}$$

$$B \equiv \sum_{Z\subset[m-1]} \big[|Z| - (m-1)\epsilon\big] C(Z)\epsilon^{|Z|-1}(1-\epsilon)^{m-|Z|-2}$$

Noting that if $C(k)$ is monotone then so is $C(k \cup \{m\})$, we infer $A \geq 0$ and $B \geq 0$ from the induction hypothesis. In conjunction with $\epsilon \geq 0$ and $1-\epsilon \geq 0$, we conclude that the first and second terms are non-negative. As $C(Z\cup\{m\}) \geq C(Z)$ by the monotonicity of $C(Z)$, the third term is also non-negative, and therefore the entire expression, completing the induction step and the proof of (4.2).

This proves (4.2) for each and every user. Therefore (4.1) is also proven, as the expected number of customers is the sum of selection probabilities over all users, and the lemma follows.

The lemma leads to a proof of the Theorem:

In the absence of information showing that the lead in reputation is shrinking, the presumption is that a lead in reputation is indicative of a higher or equal rate of increase of reputation. By the lemma, the expected rate of increase of reputation (reputation feedback) is higher for the expert with the higher expertise. Given that *a priori* the two experts are equally likely to have the higher reputation, by Bayes' rule *a posteriori* the expert whose lead was observed is likely to have the higher expertise.

Immediate corollaries are that observing all current expert reputations, the expert with the highest reputation is most likely to have the highest expertise, and that the expertise ranking order is most likely (out of all possible rankings) to be the reputation ranking order.

5. Discussion

5.1. From Reputation Systems to Reputation in General

Our analysis was framed in the context of online reputation systems, which are orderly and well-defined mechanisms which lend themselves to formal analysis. It would be natural to extend the analysis to the broader phenomenon of reputation, defined as the general estimation in which the expertise of an entity is held by the public. Mechanisms by which such a general estimation is formed in economic and social environments vary: It may be formed by word-of-mouth, by the influence of mass media, by the opinion of generally acknowledged authorities, by advertising, or by a combination of all of the above. We believe that our results are suggestive of the dynamics of reputation in this broad sense.

The next section exhibits one way a reputation model might work and be governed by our model without being managed by a centralized reputation system.

5.2. Market Share as Reputation: An Alternative Embodiment of the Model

An alternative embodiment of the model, retaining all of its key ingredients, is one where *market share* takes the place of reputation: Referring to the model as described in Section 2, we note that if we set the discounting factor α to 0, then reputation $r(t)$ and reputation feedback $w(t)$ become identical. Additionally setting $\beta = 1$ for positive feedback only, the value of $r(t)$ becomes the number of customers who have used an expert in a round, and were satisfied with his service: In effect, his market share. By further noting that the value of α was immaterial to most of the results we derived for the model (because the trend, set by $w(t)$, dominates the absolute values of $r(t)$ in the long run), we conclude that the bulk of our results holds when market share (thus defined) is substituted for reputation.

In particular, our result regarding the rationality of using reputation as a signal for expertise can certainly be carried over to this alternate embodiment, showing that market share is a positive signal for expertise, and therefore that using market share as a guide in selecting experts is rational user behavior.

5.3. Relation to Information Cascades

Information cascades [2] apparently offer a competing paradigm to the reputation-expertise dilemma analyzed by our model. Perhaps not coincidentally, a favorite example in information cascades concerns, like our opening question, restaurants: It shows the logic by which a user may join the longer queue to two restaurants, possibly in disagreement with her own private information. However we point out a fundamental difference: The dynamics of information cascades are independent of actual customer satisfaction through the objective success or failure of the service provider. On the other hand, feedback is a central aspect of our reputation model. The difference in a nutshell is that in the information cascades model a user would go to a restaurant based on the number of people queuing, while in our model a user would do the same based on the number of people recommending it on their way out.

While it is possible to model the public's rush to adopt a movie, a soft drink or a recording artist as information cascades, doing so would ignore the role of experience: Actually liking the movie, soft drink or artist, having experienced them, is largely a decision based on objective excellence or private taste, on which the influence of public opinion is limited. We therefore offer our model as appropriate for these situations. The information cascades model is appropriate for situations in which the truth value of beliefs is revealed in the distant future, or never, as for example regarding the reality of anthropogenic global warming, or of most political views. Our model is better suited for situations in which the truth value of beliefs is (noisily) revealed in time to influence the behavior of participants.

5.4. Further Study

In economics, a natural extension of the model is a pricing model, i.e., a model wherein experts put a price tag on their services, and users differ by their budget, or sensitivity to price. Price competition between several experts may be investigated, and price equilibrium sought. It stands to reason that a lead in reputation can be leveraged to a more profitable pricing strategy.

Pricing logically leads to a modeling of the **firm** as a special and important case of the expert in our model. In particular, it may be interesting to augment the model with the possibility of "buying" reputation (which we have eschewed in the current paper) by expenditure on advertising and other forms of marketing to raise a firm's brand/reputation. A further natural option for a firm is the possibility of "buying" expertise, modeling expenditure on R&D as a means of increasing the quality of a firm's products and services, represented as expertise in our model.

In this context an investigation of a generalized competition between firms, based on competition using prices, investment in reputation and investment in expertise, and governed by the reputation-expertise model, is interesting with potentially useful results.

A different direction for study arises from treating reputation as a local rather than a global attribute of an expert. "Local" here should be understood in terms of a social or geographical network in which feedback from user experience with an expert, rather than contributing to a global, common-knowledge reputation, influences only the immediate neighbors of the user in the social or geographical graph. Such a framework would then be suitable for studying social learning, a subject already extensively studied in the literature with various learning models, e.g. Bala and Goyal [1] and Rosenberg, Solan and Vieille [10]. In contrast to imitation, Bayesian deduction, and other learning mechanisms suggested by the literature, the reputation-expertise model suggests a learning mechanism of recommendation and trial: Individuals get suggestions from their neighbors but "learn" based on their own positive experiences. One question of importance in social learning, of whether the mechanism can predict the stability of diversity in a social network, is answered in some models in the negative, by, e.g. Bala and Goyal [1]. Our model, showing the self-perpetuating property of a high reputation, may lend itself well to explaining the formation of diversity barriers in which an entrenched custom or belief would successfully resist the invasion of a superior rival.

Appendix A. Relation to Google's Page Rank

One the main reasons we started to investigate the present subject was our desire to understand the dynamics underlying Google's search engine. We now exhibit a close relationship between our concept of reputation and Google's "page rank":

Google's page rank [3] is widely known as the main criterion by which pages are ranked in response to Google search. The page rank algorithm essentially computes an eigenvector for a matrix whose rows and columns represent all pages (or domains) on the web, and the coordinates of the resulting eigenvector are the page rank of each page (or domain). A simplified explanation of page rank is that it is the sum of "endorsements" by links to that page. Similarly, in our model, reputation of an expert is a sum of endorsements by satisfied customers. We now show that this qualitative similarity is in fact deeper, and that by likening our model's expert-user system to a part of the web, Google's page rank calculation for such a mini-web would be nearly identical to the numerical value of expert reputations in our model.

Consider a model with no discounting ($\alpha = 1$) and reward only ($\beta = 1$). Associate a node (i.e. a web-page or domain on the web) with each expert and as well as with each user, and assume that these nodes represent the entire web. Set the initial reputations $r_1(0), \dots, r_n(0)$ as follows: Let $r_i(0)$ be the initial number of hyperlinks from the user nodes to expert node i. Identify the reputation increment awarded to an expert on successful service to a user with the creation of a new hyperlink from the user's node to the expert's node. Finally, neglect the possibility that a user query has met failure by all experts, i.e. assume that every query is successfully treated by *some* expert.

According to this description at each round t a mini-web of expert and user nodes is reached, which we call $\Omega(t)$: each round, N new hyperlinks are added between user nodes and expert nodes, and at the end of round t, tN hyperlinks are added between user nodes and expert nodes, each user node with t outgoing links. Clearly the number of hyperlinks incoming to expert i is his current reputation $r_i(t)$.

According to the page-rank algorithm, the page-rank vector of all nodes in the web is given by the eigenvector (for the first eigenvalue $\lambda_1 = 1$) of the row-stochastic matrix $\mathbf{G} = \mu\mathbf{S} + (1-\mu)\frac{1}{M}\mathbf{J}$, where:

- M is the number of nodes in the mini-web.
- **G,S and J** are $M \times M$ matrices.
- $\mu \in [0, 1]$ is a parameter of the page-rank algorithm
- **S** is the adjacency matrix, defined thus:
  - Let $w_{i,j}$ be the number of links from node i to node j, and let $W_i = \sum_{j=1}^{M} w_{i,j}$. Then

$$\mathbf{S}_{i,j} = \begin{cases} \frac{w_{i,j}}{W_i} & W_i > 0 \\ \frac{1}{M} & W_i = 0 \text{ (i.e. node } i \text{ is a ``dangling'')} \end{cases}$$

- **J** is the all-1's matrix.

Both $\mathbf{G}$ and $\mathbf{S}$ are row-stochastic matrices. Therefore $\mathbf{G}$'s largest eigenvalue is 1 with corresponding eigenvector V i.e., $V = V\mathbf{G}$.

Claim 1. *The page rank calculated by the page-rank algorithm for each expert node in $\Omega(t)$ and the reputation of that expert at round t differ by a constant independent of the expert.*

Proof. Our expert-user mini-web has $M = n + N$ nodes where expert nodes are labeled $1, \ldots, n$ and user nodes $n+1, \ldots, n+N$.

The graph underlying $\Omega(t)$ is bipartite where every edge goes from a user node to an expert node. Clearly, $W_i = 0$ for $n \geq i \geq 1$, and $W_i = t$ for $N + n \geq i \geq n+1$. We can therefore calculate $\mathbf{G}$:

$$\mathbf{G}_{i,j} = \begin{cases} \frac{1}{n+N} & i \in [n] \\ \frac{\mu w_{i,j}}{t} + \frac{1-\mu}{n+N} & i \in [n+1, n+N], j \in [n] \\ \frac{1-\mu}{n+N} & i, j \in [n+1, n+N] \end{cases}$$

Let us find the eigenvector values for $\Omega(t)$, i.e. solve $V = V\mathbf{G}$:
Let $X \equiv \frac{1}{n+N}\left[\sum_{i=1}^{n} V_i + (1-\mu)\sum_{i=n+1}^{n+N} V_i\right]$. Expanding $V = V\mathbf{G}$ we get:

$$V_j = X \qquad \forall j \in [n+1, n+N] \tag{A.1}$$

$$V_j = X + \frac{\mu}{t}\sum_{i=n+1}^{n+N} w_{i,j} V_i \qquad \forall j \in [n] \tag{A.2}$$

Substituting (A.1) in (A.2):

$$V_j = X + \frac{X\mu}{t}\sum_{i=n+1}^{n+N} w_{i,j} \qquad \forall j \in [n] \tag{A.3}$$

But the sum in the above is, by the definition of $\Omega(t)$, equal to $r_j(t)$. Therefore:

$$V_j = X + \frac{X\mu}{t} r_j(t) \qquad \forall j \in [n] \tag{A.4}$$

Since the scale of the eigenvector is arbitrary, we may divide all eigenvector elements by $\frac{X\mu}{t}$, with the result:

$$V_j = \begin{cases} r_j(t) + \frac{t}{\mu} & j \in [n] \quad \text{(for expert nodes)} \\ \frac{t}{\mu} & j \in [n+1, n+N] \quad \text{(for user nodes)} \end{cases} \tag{A.5}$$

the top line of which is what we set out to show.

So reputation, as defined in our model, and Google's page rank, are closely related. In particular, their ranking is the same. In particular, for the sake of order-based selection schemes, such as the reputation-ordered scheme and the loyalty scheme, they are equivalent.

The conclusion lends support to our suggestion in the Introduction, that the dynamics of web search engines are instructive of the interaction of reputation and expertise.

Appendix B. More on the Behavior of the Model

Appendix B.1. The Loyalty Scheme

We turn our attention to the (unmodified) loyalty scheme:

Theorem 7. *Let n experts be indexed by their reputation order (at round t) $r_1(t) > r_2(t) > \ldots > r_n(t)$ and have respective expertise $\epsilon_1, \ldots, \epsilon_n$ all of which are < 1. Then the order is quasi-stable under the* ***loyalty scheme*** *if and only if for each $i \in [n-1]$:*

$$\frac{\epsilon_i + \beta - 1}{(1-\epsilon_i)^2 + \epsilon_i \prod_{k=1}^{i}(1-\epsilon_k)} \geq \frac{\epsilon_{i+1} + \beta - 1}{1 - \epsilon_{i+1} + \epsilon_{i+1} \prod_{k=1}^{i+1}(1-\epsilon_k)} \tag{B.1}$$

Proof. Following the proof of Theorem 5, *mutatis mutandis*, we have:

$$v_{i,j} = \begin{cases} \epsilon_j \prod_{k=1}^{j-1}(1-\epsilon_k) & i = 0, j \neq 0 \\ \prod_{k=1}^{n}(1-\epsilon_k) & i = 0, j = 0 \\ (1-\epsilon_i) v_{0,j} & i \neq 0, 0 < j < i \\ \epsilon_i & i \neq 0, j = i \\ v_{0,j} & i \neq 0, j > i \ \textit{or}\ j = 0 \end{cases} \tag{B.2}$$

Given that $p_i = \sum\limits_{k=0}^{n} p_k v_{k,i}$, we calculate p_0:

$$p_0 = p_0 v_{0,0} + p_1 v_{0,0} + \ldots + p_n v_{0,0} = v_{0,0}$$

Calculating p_i, $\forall i > 0$:

$$\begin{aligned}
p_i &= p_0 v_{0,i} + p_1 v_{0,i} + \ldots + p_{i-1} v_{0,i} + p_i \epsilon_i + \\
&+ p_{i+1}(1-\epsilon_{i+1}) v_{0,i} + \ldots + p_n (1-\epsilon_n) v_{0,i} = \\
&= p_i \epsilon_i + v_{0,i}\Big[p_0 + \sum_{k=1}^{i-1} p_k + \sum_{k=i+1}^{n} p_k (1-\epsilon_k)\Big] = \\
&= p_i \epsilon_i + v_{0,i}\Big[1 - p_i - \sum_{k=i+1}^{n} p_k \epsilon_k\Big] = \\
&= \frac{1}{1-\epsilon_i} v_{0,i}\Big[1 - p_i - \sum_{k=i+1}^{n} p_k \epsilon_k\Big]
\end{aligned}$$

In summary:

$$p_i = \begin{cases} \prod\limits_{k=1}^{n} (1-\epsilon_k) & i = 0 \\ \Big[\frac{\epsilon_i}{1-\epsilon_i} \prod\limits_{k=1}^{i-1} (1-\epsilon_k)\Big]\Big[1 - p_i - \sum\limits_{k=i+1}^{n} p_k \epsilon_k\Big] & i \neq 0 \end{cases} \tag{B.3}$$

As before we investigate the ratio of probabilities between two consecutive states. For each $i \in [n-1]$:

$$\frac{p_i \frac{1-\epsilon_i}{\epsilon_i}}{\prod\limits_{k=1}^{i-1} (1-\epsilon_k)} + p_i = \frac{p_{i+1} \frac{1-\epsilon_{i+i}}{\epsilon_{i+1}}}{\prod\limits_{k=1}^{i} (1-\epsilon_k)} + p_{i+1}(1-\epsilon_{i+1}) \tag{B.4}$$

Multiplying both sides by $\prod\limits_{k=1}^{i} (1-\epsilon_k)$, we derive the analogue of (3.16):

$$p_i \Big[\frac{(1-\epsilon_i)^2}{\epsilon_i} + \prod_{k=1}^{i} (1-\epsilon_k)\Big] = p_{i+1} \Big[\frac{1-\epsilon_{i+1}}{\epsilon_{i+1}} + \prod_{k=1}^{i+1} (1-\epsilon_k)\Big] \tag{B.5}$$

From which, in conjunction with (3.17) we get the criterion for quasi-stability:

$$\frac{\epsilon_i + \beta - 1}{(1-\epsilon_i)^2 + \epsilon_i \prod\limits_{k=1}^{i} (1-\epsilon_k)} \geq \frac{\epsilon_{i+1} + \beta - 1}{1 - \epsilon_{i+1} + \epsilon_{i+1} \prod\limits_{k=1}^{i+1} (1-\epsilon_k)} \tag{B.6}$$

As with the dual-loyalty scheme, the requirement that all expertise levels be smaller than 1 is necessary, because if any experts have expertise of 1, the only possible steady-state is where **all** users are loyal to one of these experts.

Corollary 8. *The loyalty scheme is not regular. However quasi-stability of a pair depends only on the expertise levels of the pair or higher-ranked experts, and is independent of lower-ranked experts.*

The non-regularity of the loyalty scheme implies that Theorem 3 cannot be applied to it. It is natural to ask whether (as in the regular and non-chaotic case) every initial expert order has a unique limit order. It turns out that the answer is negative, as the following example shows.

Example 3. *Consider a loyalty scheme with $n = 3, \beta = 1$ and the following initial setting:*

$$(\epsilon_1 = 0.45, \epsilon_2 = 0.55, \epsilon_3 = 0.66)$$

Both neighboring pairs in this setting are unstable. Swapping first experts 2 & 3 and then swapping experts 1 & 3 reaches the quasi-stable

$$(\epsilon_3 = 0.66, \epsilon_1 = 0.45, \epsilon_2 = 0.55)$$

Starting instead by swapping experts 1 & 2, followed by swapping 1 & 3 and finally 2 & 3 reaches a different quasi-stable order:

$$(\epsilon_3 = 0.66, \epsilon_2 = 0.55, \epsilon_2 = 0.45)$$

Corollary 9. *Under the terms of Theorem 7 with a reward-only scheme ($\beta = 1$), quasi-stability requires the following equivalent inequalities for each $i \in [n-1]$:*

1. $\frac{(1-\epsilon_i)^2}{\epsilon_i} + \prod_{k=1}^{i}(1-\epsilon_k) \leq \frac{1-\epsilon_{i+1}}{\epsilon_{i+1}} + \prod_{k=1}^{i+1}(1-\epsilon_k)$
2. RECIPROCAL DIFFERENCE: $\frac{1}{\epsilon_i} - \frac{1}{\epsilon_{i+1}} \leq (1-\epsilon_i)\Big[1-\epsilon_{i+1}\prod_{k=1}^{i-1}(1-\epsilon_k)\Big]$

Comparing the corresponding reciprocal difference results for the reputation-ordered scheme Corollary 2) and the dual loyalty scheme (Corollary 7), we see that the leader's advantage is smaller than in each of the two, and that the dual loyalty scheme is a half-way station in this respect between the reputation-ordered scheme and the loyalty scheme. Still the advantage persists. Indeed, the reciprocal difference is smaller than $1 - \epsilon_i$, the dual loyalty advantage, but larger than $(1 - \epsilon_i)(1 - \epsilon_{i+1})$.

The leader's advantage is smallest between the two leading experts, between which the reciprocal advantage is $(1 - \epsilon_1)(1 - \epsilon_2)$.

Two-sided quasi-stability is still feasible, but the window for it is even smaller. The analogue of (3.9) is:

$$
\begin{aligned}
-(1-\epsilon_{i+1})\Big[1-\epsilon_i\prod_{k=1}^{i-1}(1-\epsilon_k)\Big] &\leq \\
\frac{1}{\epsilon_i}-\frac{1}{\epsilon_{i+1}} &\leq \\
(1-\epsilon_i)\Big[1-\epsilon_{i+1}\prod_{k=1}^{i-1}(1-\epsilon_k)\Big]
\end{aligned}
$$

As in the other schemes, quasi-stability is impossible with penalty only ($\beta = 0$), and every setting is chaotic, because:

$$
\begin{aligned}
\frac{\epsilon_i - 1}{(1-\epsilon_i)^2+\epsilon_i\prod_{k=1}^{i}(1-\epsilon_k)} &\geq \frac{\epsilon_{i+1}-1}{1-\epsilon_{i+1}+\epsilon_{i+1}\prod_{k=1}^{i+1}(1-\epsilon_k)} && \Rightarrow \\
\frac{-1}{1-\epsilon_i+\epsilon_i\prod_{k=1}^{i-1}(1-\epsilon_k)} &\geq \frac{-1}{1+\epsilon_{i+1}\prod_{k=1}^{i}(1-\epsilon_k)} && \Rightarrow \\
-\epsilon_i(1-\prod_{k=1}^{i-1}(1-\epsilon_k)) &\geq \epsilon_{i+1}\prod_{k=1}^{i}(1-\epsilon_k)
\end{aligned}
$$

Which is impossible as the left-hand side is negative while the right-hand side is positive.

Appendix B.2. A Pathological Scheme

Selection schemes considered so far were grounded in logical user behavior. The following scheme has no logic to it, and in fact may be called illogical.

The purpose of this brief intellectual exercise is to get an idea on the range of possible behaviors of the model under various selection orders.

We propose a scheme called the **reverse ordered scheme** in which users query experts in ascending order of reputation, starting from the expert with the least reputation. As always, there are n experts with reputations $r_1(t) > \ldots > r_n(t)$, corresponding expertise $\epsilon_1, \ldots, \epsilon_n$, and reward/penalty factor β.

The expected number of customers of each expert i is:

$$\mathbb{E}[c_i(t)] = N \prod_{k=i+1}^{n} (1 - \epsilon_k) \tag{B.7}$$

And the condition for quasi-stability is:

$$\begin{aligned} \mathbb{E}[w_i(t)] &\geq \mathbb{E}[w_{i+1}(t)] & \Rightarrow \\ (\epsilon_i + \beta - 1)\, \mathbb{E}[c_i(t)] &\geq (\epsilon_{i+1} + \beta - 1)\, \mathbb{E}[c_{i+1}(t)] & \Rightarrow \\ (\epsilon_i + \beta - 1)(1 - \epsilon_{i+1}) &\geq \epsilon_{i+1} + \beta - 1 & \Rightarrow \end{aligned}$$

which yields the condition.

$$\frac{\beta}{\epsilon_i} - \frac{1}{\epsilon_{i+1}} \leq -1 \tag{B.8}$$

Compare this condition to the reciprocal difference condition of Theorem 1.

Examining the quasi-stability condition, we take note of interesting facts about this scheme:

- There exist quasi-stable reward-only settings, e.g.

  $$(\epsilon_1 = 1, \epsilon_2 = \frac{1}{2}, \epsilon_3 = \frac{1}{3}, \epsilon_4 = \frac{1}{4})$$

  For this to happen expertise levels must be in descending order and “not too close”. On the other hand the following setting is chaotic.

  $$(\epsilon_1 = \frac{2}{3}, \epsilon_2 = \frac{1}{2})$$

- **All** penalty-only settings are quasi-stable! This is the mirror-image of Corollary 3. Remarkably, while in this scheme reputation is purely result-based, reputation orders are virtually guaranteed to have no connection to skill.

- If a setting is chaotic then any expert in a chaotic pair has expertise $\geq 1 - \beta$.
- The scheme is order-based and regular.

Appendix B.3. Other Observations regarding Steady-States

I. In addition to whether a steady-state exists, we may take interest in the value of reputation at steady-state, or rather, in its expectation. Referring to (2.2), we note that in an order-based scheme, $w_i(t)$ is constant so long as the reputation order is stable, which is the definition of a steady-state. Therefore, from (2.2), letting $w_i = \mathbb{E}[w_i(t)]$ and assuming an order-based scheme, it is observed that $\lim_{t\to\infty} \frac{r_i(t)}{t} = w_i(t)$ in the absence of discounting ($\alpha = 1$). On the other hand, with discounting ($\alpha < 1$), applying expectation to both sides of (2.2) , i'th (constant) expected feedback at steady-state, $r_i = \mathbb{E}[r_i(t)]$ has a fixed point satisfying $r_i = \alpha r_i + w_i$, and therefore:

$$r_i = \frac{w_i}{1 - \alpha} \tag{B.9}$$

II. In the chaotic case, where no steady-state exists, and two-sided instability exists between an expert pair, we may ask what part of the time, on average, each of the experts has the lead. For order-based schemes, we provide the following answer:

Label a chaotic expert pair as 1 and 2. Let $w_i^j, i = 1, 2, j = 1, 2$ be i's reputation feedback when j has the lead. Then, define:

$$\Delta_1 \equiv w_2^1 - w_1^1$$
$$\Delta_2 \equiv w_1^2 - w_2^2$$

By the definition of two-sided instability, $\Delta_1 > 0$ and $\Delta_2 > 0$. Clearly, Δ_1 is the expected per-round change to the reputation difference $r_2 - r_1$ when expert 1 leads, while $-\Delta_2$ is the same when expert 2 leads. Let p_1 and p_2 be the probabilities that expert 1 and 2, respectively, holds the lead ($p_1 + p_2 = 1$).

The per-round expected change to the reputation difference is therefore $\Delta = p_1\Delta_1 - p_2\Delta_2$. As the reputation lead is expected to change an infinite number of times, necessarily $\Delta = 0$, therefore:

$$p_1 = \frac{\Delta_2}{\Delta_1 + \Delta_2}$$
$$p_2 = \frac{\Delta_1}{\Delta_1 + \Delta_2}$$

For example, using the reputation-ordered scheme (which is order-based), and with penalties only ($\beta = 0$):

Referring to (3.6) and (3.7), and noting that $\mathbb{E}[c_i(t)]$ depends only on experts ranked higher than i, which we may therefore mark as a constant C:

$$w_1^1 = C(\epsilon_1 - 1)$$
$$w_2^1 = C(\epsilon_2 - 1)(1 - \epsilon_1)$$
$$w_1^2 = C(\epsilon_1 - 1)(1 - \epsilon_2)$$
$$w_2^2 = C(\epsilon_2 - 1)$$

Therefore:

$$\Delta_1 = w_2^1 - w_1^1 = C\epsilon_2(1 - \epsilon_1)$$
$$\Delta_2 = w_1^2 - w_2^2 = C\epsilon_1(1 - \epsilon_2)$$

Expert 1's and 2's time-shares of the lead are in proportion $\frac{1}{\Delta_1} : \frac{1}{\Delta_2}$, and therefore in proportion $\frac{\epsilon_1}{1-\epsilon_1} : \frac{\epsilon_2}{1-\epsilon_2}$.

[1] Bala, V. and S. Goyal: Learning from Neighbors. *Review of Economic Studies*, 65, 595-621, 1998.

[2] Bikhchandani, S., Hirshleifer, D. and Welch, I: A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, **100**, 992-1026, 1992.

[3] Brin, S. and L. Page: The Anatomy of a Large-Scale Hypertextual Web Search Engine, *Computer Networks and ISDN Systems*, **33**: 107-117, 1998.

[4] Cabral, L.: Stretching Firm and Brand Reputation, *RAND Journal of Economics*, **31**, 658-673, 2000.

[5] Friedman, E., P. Resnick and R. Sami : Manipulation-Resistant Reputation Systems in *Algorithmic Game Theory, Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani, editors Cambridge University Press*, 677-697, 2007.

[6] Gittins, J. C. : Multi-armed bandit allocation indices. Wiley-Interscience Series in Systems and Optimization.. Chichester: John Wiley & Sons, Ltd.. ISBN 0-471-92059-2, 1989.

[7] Kreps, D.: Corporate Culture and Economic Theory. in J Alt and K Shepsle (Eds), *Perspectives on Positive Political Economy*, Cambridge: Cambridge University Press, 90-143, 1990.

[8] Kreps, D.M. and R. Wilson: Reputation and imperfect information, *Journal of Economic Theory*, **27**, 253-279, 1982.

[9] Mailath, G.J. and L. Samuelson: Who Wants a Good Reputation?, *Review of Economic Studies*, **68**, 415-441, 2001.

[10] Rosenberg, D., E. Solan, and N. Vieille: Informational externalities and emergence of consensus. *Games and Economic Behavior* **69** (2009), 979-994.

[11] Selten, R.: The Chain-Store Paradox, *Theory and Decision*, **9** (1978), 127-159.

[12] Tadelis, S.: What's in a Name? Reputation as a Tradeable Asset, *The American Economic Review*, Vol. 89, No. 3, pp. 548-563, June 1999.