

# Beyond Accuracy. Reputation for Partner selection with Lies and Retaliation

Isaac Pinyol<sup>1</sup>, Mario Paolucci<sup>2</sup>, Jordi Sabater-Mir<sup>1</sup>, and Rosaria Conte<sup>2</sup>

<sup>1</sup> Artificial Intelligence Research Institute, Barcelona, SPAIN

<sup>2</sup> Institute for Cognitive Science and Technology, Rome, ITALY

**Abstract.** In an unpredictable, heterogeneous world, intelligent agents depend on accurate social information; reputation, among the preeminent artifacts to transmit social evaluations, has been receiving growing attention by social scientists. A realistic description of reputation must include inaccurate information; in this paper, based on the distinction between image (agents' believed evaluation of a target) and reputation (circulating evaluation, without reference to the evaluation source), we model the spreading of information in a simple market with the presence of liars and the possibility of retaliation. While fear of retaliation inhibits the spreading of image, the detached character of reputation can be a cause of inaccuracy; The two forces could balance in different settings. In a set of simulations, with agents using the RePage platform for management of image and reputation, we compare the usage of image alone with the usage of both image and reputation. Reputation is shown to be preferable over image to allow for faster discover of scarce good sellers.

## 1 Introduction

In an unpredictable world, intelligent agents are shown to depend on accurate information ([1], [2], [3], [4]) for acting adaptively. More specifically, they depend on accurate social information for interacting into a heterogeneous multiagent world. Memory of past experience is a precious source of information, but is usually acquired at own expenses. As obtaining experience may be fatal in a world of cheaters, agents depend on one another to indirectly acquire information for partner selection, before interacting with, and in order to avoid, the bad guys.

In the last ten years or so ([5]; [6]; [7]), the role of indirectly acquired social information has been appreciated by social scientists to a fairly realistic degree. Indeed, reputation has received a growing attention as network-based social evaluation [8]. However, by this means the rate of inaccurate information circulating in a multiagent system increases, and the question as to how put up with such inaccuracy starts to be posed. Stated otherwise, if reputation is a mechanism for finding out the material cheaters, i.e. nonreciprocators in exchange of material goods, how find out the informational cheaters, i.e. deceitful informants?

Solutions to this problem in the MAS field usually rely on learning strategies ([9]; [10]), by means of which agents more or less gradually discover trustable informers and gain accurate information. Applied to detection of informational

cheating, learning techniques are rather more efficient than is the case with material cheating, since facing deception in reputation is less costly than trading with a gouger. However, this process results in shrieked social networks where agents get stuck, to the expenses of less experienced agents which are "new to the environment" [11]. But if by learning we mean feedback from experience that provides more accurate information, there are some distinct and interacting problems that need clarification and solution.

- On the side of the input, personal experience does not spare the costs of its acquisition: people do often pay dearly the privilege of telling who is a cheater and who is not!
- On the side of the output, once agents have got this precious information, why should they come to share it with others? In this sense, the idea of agents resorting to local and total beliefs ([9]) is somewhat naive.

These questions cannot be addressed by means of "try and test" techniques, but need theory-driven experimentation and implementation. Learning is not the key to solve any problem of adaptation: suffice to have a look at natural societies to find out that human agents neither always nor often converge on trustworthy informers, nor on rational expectations.

In this paper, we deal with deception by means of a theory of reputation, and closer to what is found to happen in natural societies [12], [13]. We will tackle the issue by means of a reputation system (REPAGE, [14]), that can metabolise inaccurate information rather than discard it thanks to a fundamental distinction between image (own opinion of a target) and reputation (others' opinion of a target). In substance, image is passed on as accurate evaluation, reputation is passed on as reported, hencefore not necessarily accurate evaluation (although in fact liars may choose either modality). The interplay between these two objects in the agents' minds and in social transmission allows both accurate and inaccurate information to spread over the population. We will simulate the effects of these assumptions over a simplified market, where buyers choose sellers of variable quality, endowed with a variable number of non-replenishing stock units.

Two assumptions about how REPAGE affects agents are made:

- In order to select information, agents are more conservative in image than reputation acceptance; in other words, they may accept as a reputation what they do not believe as an image.
- In order to avoid retaliation if found out to be inaccurate, agents transmit evaluations that they do not believe or are uncertain about, but only as reputation (others' opinion).

Thanks to image, incoming information affects agents' beliefs, and consequently partner selection, to a lower extent and in a more controlled way than it affects their message passing. Thanks to reputation, evaluations circulate into the system, providing continuous input to belief formation and revision. Reputation makes the system more dynamic and alert. Image makes it more selective and controlled. Reputation releases uncertain information, which might turn out

to be more or less accurate, and which agents will check before using it for their own purpose. A system based on image only is expected to hinder the information flow to the benefit of accuracy; a reputation-prone system is expected to collapse under uncertainty and inaccuracy.

## 2 Theory Introduction

### 2.1 Image And Reputation

Our proposal is based on a model of imAGE, REPutation and their interplay developed in [12]. Although both are social evaluations, image and reputation are distinct objects. Image is a simple evaluative belief [15]; it tells that the target is “good” or “bad” with respect to a norm, a standard, or a skill. Reputation is a belief about the existence of a communicated evaluation. Consequently, to assume that a target  $t$  is assigned a given reputation implies only to assume that  $t$  is reputed to be “good” or “bad”, i.e., that this evaluation circulates, but it does not imply to share the evaluation.

To select good partners, agents need to form and update own social evaluations; hence, they must exchange evaluations with one another. If agents should transmit only believed image, the circulation of social knowledge would be bound to stop soon. On the other side, agents that believe all the informations that they receive would be no more autonomous; in order to preserve their autonomy, agents need to *decide* independently whether to share or not and whether to believe or not others’ evaluations of a given target. Hence, they must

- form both evaluations (image) and meta-evaluations (reputation), keeping distinct the representation of own and others’ evaluations, before
- deciding whether or not to integrate reputation with their own image of a target.

Unlike other current systems, in REPAGE reputation does not coincide with image. Indeed, agents can either transmit their own image of a given target, which they hold to be true, or report on what they have “heard” about the target, i.e. its reputation, whether they believe this to be true or not. Of course, in the latter case, they will neither commit to the information truth value nor feel responsible for its consequences. Consequently, agents are expected to transmit uncertain information, and a given positive or negative reputation may circulate over a population of agents even if its content is not actually believed by the majority.

To remark the difference between the effects of REPutation and imAGE, we will examine all the simulation scenarios in the rest of the paper under the two main experimental conditions

- L1, where there is only exchange of image between agents
- L2, where agents can exchange both image and reputation.

Note that while L1 is comparable with a large body of similar literature (ex. [16]), the introduction (L2) of reputation as a separate object in a simulative experiment will be presented in this paper for the first time.

### 3 The Repage model and Architecture

Repage ([14]) is a computational system based on a cognitive theory of reputation [12] that proposes a fundamental distinction between image and reputation. The Repage architecture has three main elements, a memory, a set of detectors and the analyzer. The memory is composed by a set of references to the predicates hold in the main memory of the agent. Predicates are conceptually organized in levels and inter-connected. Each predicate that belongs to one of the main types (including image and reputation) contains a probabilistic evaluation that refers to a certain agent in a specific role. For instance, an agent may have an image of agent T (target) as a seller (role), and a different image of the same agent T as informant. The probabilistic evaluation consist of a probability distribution over the discrete sorted set of labels: {Very Bad, Bad, Normal, Good, Very Good}.

The network of dependences specifies which predicates contribute to the values of others. In this sense, each predicate has a set of precedents and a set of antecedents. The detectors, inference units specialized in each particular kind of predicate, receive notifications from predicates that changes or that appear in the system and uses the dependences to recalculate the new values or to populate the memory with new predicates.

Each predicate has associated a strength that is function of its antecedents and of the intrinsic properties of each kind of predicate. As a general rule, predicates that resume or aggregate a bigger number of predicates will hold a higher strength.

At the first level of the Repage memory we find a set of predicates not evaluated yet by the system. *Contracts* are agreements on the future interaction between two agents. Their result is represented by a *Fulfillment*. *Communications* is information that other agents may convey, and may be related to three different aspects: the image that the informer has about a target, the image that, according to the informer, a third party agent has on the target, and the reputation that the informer has about the target.

In level two we have two kind of predicates. *Valued communication* is the subjective evaluation of the communication received that takes into account, for instance, the image the agent may have of the informer as informant. Communications from agents whose credibility is low will not be considered as strong as the ones coming from well reputed informers. An *outcome* is the agent's subjective evaluation of a direct interaction, built up from a fulfillment and a contract.

At the third level we find two predicates that are only fed by valued communications. On one hand, a *shared voice* will hold the information received about the same target and same role coming from communicated reputations. On the other hand, *shared evaluation* is the equivalent for communicated images and third party images.

Shared voice predicates will finally generate *candidate reputation*; shared evaluation together with outcomes will generate *candidate image*. Newly generated candidate reputation and image aren't usually strong enough; new communications and new direct interactions will contribute to reinforce them until a

threshold, over which they become full-fledged image or reputation. We refer to [14] for a much more detailed presentation.

From the point of view of the agent structure, integration with the other parts of our deliberative agents is straightforward. Repage memory links to the main memory of the agent that is fed by its communication and decision making module, and at the same time, this last module, the one that contains all the reasoning procedures uses the predicates generated by Repage to make decisions.

## 4 Description of the Experiment

We have designed the simulation experiment as the simplest possible setting where accurate information is a *commodity*, meaning that information is both valuable and scarce. Since the system will be used as a proof of concept, we will not ground it with micro or macro data, but we will instead use a simplified generic economic metaphor of an agent-based market setting with instability. This simplified approach is largely used in the field [16], both on the side of the market design and of the agent design. We release the simplification on the agent design while keeping it on the side of the market design.

The experiment includes only two kind of agents, the buyers and the sellers. All agents perform actions in discrete time units (turns from now on). In a turn, a buyer performs one communication request and one purchase operation. In addition, the buyer answers all the information requests that it receives.

Goods are characterized by an utility factor that we interpret as quality (but, given the level of abstraction used, could as well represent other utility factors as quantity, discount, timeliness) with values between 1 and 100.

Sellers are characterized by a constant quality and a fixed stock, that is decreased at every purchase; they are essentially reactive, their functional role in the simulation being limited to providing an abstract good of variable quality to the buyers. Sellers exit the simulation when the stock is exhausted or when for certain number of turn they do not sell anything, and are substituted by a new seller with similar characteristics.

The disappearance of sellers makes information necessary; reliable communication allows for faster discover of the better sellers. This motivates the agents to participate in the information exchange. In a setting with permanent sellers (infinite stock), once all buyers have found a good seller, there is no reason to change and the experiment freezes. With finite stock, even after having found a good seller, buyers, should be prepared to start a new search when the good seller's stock ends.

At the same time, limited stock makes good sellers a scarce resource, and this constitutes a motivation for the agents not to distribute information. One of the interests of the model is in the balance between these two factors.

There are four parameters that describe an experiment: the number of buyers  $NB$ , the number of sellers  $NS$ , the stock for each seller  $S$ , and the distribution of quality among sellers. In 2 we defined the two main experimental situations, L1

where there is only exchange of image, and L2 where both image and reputation are used.

#### 4.1 Decision making module

In our experiments the decision making procedure is a key point that determines the performance of the whole system. From the seller side, this procedure is quite simple since they are limited to *sell* the products that buyers require and to disappear when the stock gets exhausted. From the point of view of the buyers, at each turn they have to ask one question to another buyer and buy some item from a seller. They may also answer a question from other buyers. Each of these actions requires the buyer to make some decisions:

**Buying action:** In this action the question is: which seller should I choose? The Repage system provides information about image and reputation of each one of the sellers. The easiest option would be to pick the seller with *better* image, or (in L2) better reputation if image is not available. We set a threshold for an evaluation (actually, for its center of mass, see [14] for definitions) to be considered *good enough* to be used to make a choice. In addition, we keep a limited chance to explore other sellers, controlled by the system parameter *risk*<sup>3</sup>. Figures 1 and 2 describe the reasoning procedure that agents use to pick the seller in the situations L1 and L2 respectively. Notice that image has always priority over reputation, since image imply an acknowledge of the evaluation itself while reputation only an acknowledge of what is said.

- |   |
|---|
| <ol style="list-style-type: none"><li>1. <i>Candidate_Seller</i> := Select randomly one image's seller</li><li>2. If <i>Candidate_Seller</i> is empty or decided to risk then <i>Candidate_Seller</i> := select randomly one seller without image</li><li>3. Buy from <i>Candidate_Seller</i></li></ol> |
|---|

**Fig. 1.** Buying action: Decision procedure for L1

- |   |
|---|
| <ol style="list-style-type: none"><li>1. <i>Candidate_Seller</i> := Select randomly one good enough seller image.</li><li>2. If <i>Candidate_Seller</i> is empty then <i>Candidate_Seller</i> := select randomly one good enough seller reputation</li><li>3. if <i>Candidate_Seller</i> is empty or decided to risk then <i>Candidate_Seller</i> := select randomly one seller without image</li><li>4. Buy from <i>Candidate_Seller</i></li></ol> |
|---|

**Fig. 2.** Buying action - Decision procedure for L2

---

<sup>3</sup> Risk is implemented as a probability (typically between 5% and 15%) for the buyer to try out unknown sellers

**Asking action:** As in the previous action, the first decision is the choice of the agent to be queried, and the decision making procedure is exactly the same than for choosing a seller, but dealing with images and reputation of the agents as informers (*informer image*) instead of as sellers. Figures 3 and 4 describe these procedures in situations L1 and L2 respectively.

- |   |
|---|
| <ol style="list-style-type: none"> <li>1. <i>Candidate_Informer</i> := Select randomly one <i>good enough</i> informer image</li> <li>2. If <i>Candidate_Informer</i> is empty or decided to risk then <i>Candidate_Informer</i> := select randomly one buyer without image as informer</li> <li>3. Ask to <i>Candidate_Informer</i></li> </ol> |
|---|

**Fig. 3.** Asking action - Decision procedure for L1

- |   |
|---|
| <ol style="list-style-type: none"> <li>1. <i>Candidate_Informer</i> := Select randomly one good enough informer image</li> <li>2. If <i>Candidate_Informer</i> is empty then <i>Candidate_Informer</i> := select randomly one good enough informer reputation</li> <li>3. if <i>Candidate_Informer</i> is empty or decided to risk then <i>Candidate_Informer</i> := select randomly one buyer without image as informant</li> <li>4. Ask to <i>Candidate_Informer</i></li> </ol> |
|---|

**Fig. 4.** Asking action - Decision procedure for L2

Once decided who to ask, the kind of question must be chosen. We consider only two possible queries: Q1 - Ask information about a buyer as informer (basically, how honest is buyer X as informer?), and Q2 - Ask for some good or bad seller (for instance, who is a good seller, or who is a bad seller?). Notice that this second possible question does not refer to one specific individual, but to the whole body of information that the queried agent may have. This is in order to allow for managing large numbers of seller, when the probability to choose a target seller that the queried agent have some information about would be very low. The agent will ask one of these two questions with a probability of 50%. If Q1 is chosen, buyer X as informer would be the less known one, that is, the one with less information to build up an image or reputation of it.

**Answering action:** Let agent *S* be the agent asking the question, *R* the agent being queried. Agents can lie, either because they are cheaters or because they are retaliating. When a buyer is a cheater whatever information being answered is changed to its opposite value. Retaliation is accomplished by sending inaccurate information from the point of view of the sender (for instance, sending "Idontknow" when really it has information, or simply giving the opposite value) when *R* has a bad image of *S* as informer. In L1 retaliation is done by sending an "Idontknow" message even when *R* has information. This avoids possible retaliation from *S* since an "Idontknow" message do not imply any commitment. If reputation is allowed, (L2) retaliation is accomplished in the same way as if the

agent were a liar, but converting all image to send into reputation, in order to avoid as well possible retaliation from  $S$ .

Because of the fear of retaliation, sending an image will take place only when an agent is very secure of that evaluation, in the sense of the REPAGE *strength* parameter included in every evaluation. This is yet another parameter *thStrength*, that allows to implement *fear of retaliation* in the agents. Notice that if *thStrength* is zero, there is no fear since whatever image formed will be a candidate to be sent, no matter its strength. As we increase *thStrength*, agents will become more conservative, less image and more reputation will circulate in the system.

Figures 5 and 6 describe the decision making process that agents use to answer the question Q1 in situations L1 and L2 respectively. In figures 7 and 8 is shown the processes agents use to answer Q2 in situations L1 and L2 respectively.

- |  |
|--|
| <ol style="list-style-type: none"> <li>1. <math>ImgX :=</math> Get image of agent <math>X</math> as informant</li> <li>2. if <math>ImgX</math> exists and <math>strength(ImgX) \geq thStrength</math> then send <math>ImgX</math> to agent <math>S</math>, END</li> <li>3. else send "Idontknow" to agent <math>S</math>, END</li> <li>4. if <math>ImgX</math> does not exist then send "Idontknow" to agent <math>S</math></li> </ol> |
|--|

**Fig. 5.** Answering Q1 - Decision procedure for agent  $R$ , L1

- |  |
|--|
| <ol style="list-style-type: none"> <li>1. <math>ImgX :=</math> Get image of agent <math>X</math> as informant</li> <li>2. if <math>ImgX</math> exists and <math>strength(ImgX) \geq thStrength</math> then send <math>ImgX</math> to agent <math>S</math>, END</li> <li>3. else convert <math>ImgX</math> to <math>RepX</math> and send <math>RepX</math> to <math>S</math>, END</li> <li>4. if <math>ImgX</math> does not exist then <math>RepX :=</math> Get reputation of agent <math>X</math> as informant</li> <li>5. if <math>RepX</math> exists then send <math>RepX</math> to <math>S</math>, END</li> <li>6. if <math>RepX</math> does not exist then send "Idontknow" to agent <math>S</math></li> </ol> |
|--|

**Fig. 6.** Answering Q1 - Decision procedure for agent  $R$ , L2

- |   |
|---|
| <ol style="list-style-type: none"> <li>1. <math>IG :=</math> Get good enough images of sellers; <math>IB :=</math> Get bad enough images of sellers</li> <li>2. if <math>IG</math> is not empty then <math>CandImage :=</math> Pick one randomly from <math>IG</math></li> <li>3. else if <math>IB</math> is not empty then <math>CandImage :=</math> Pick one randomly from <math>IB</math></li> <li>4. if <math>CandImage</math> is not empty and <math>strength(CandImage) \geq thStrength</math> then sent <math>CandImage</math> to <math>S</math>, END</li> <li>5. if <math>CandImage</math> is not empty and <math>strength(CandImage) &lt; thStrength</math> then send "Idontknow" to <math>S</math>, END</li> <li>6. if <math>CandImage</math> is empty then send "Idontknow" to agent <math>S</math></li> </ol> |
|---|

**Fig. 7.** Answering Q2 - Decision procedure for agent  $R$ , L1

<ol style="list-style-type: none"> <li>1. IG := Get good enough images of sellers; IB := Get bad enough images of sellers</li> <li>2. RG := Get good enough reputations of sellers;</li> <li>3. RB := Get bad enough reputations of sellers</li> <li>4. if IG is not empty then CandImage := Pick one randomly from IG</li> <li>5. else if IB is not empty then CandImage := Pick one randomly from IB</li> <li>6. if CandImage is not empty and strength(CandImage) <math>\geq</math> <i>thStrength</i> then send CandImage to <i>S</i>, END</li> <li>7. if CandImage is not empty and strength(CandImage) <math>&lt;</math> <i>thStrength</i> then convert CandImage to CandRep and send it to <i>S</i>, END</li> <li>8. if RG is not empty then CandRep := Pick one randomly from RG</li> <li>9. else if RB is not empty then CandRep := Pick one randomly from RB</li> <li>10. if CandRep is not empty send CandRep to <i>S</i>, END</li> <li>11. if CandRep is empty send "Idontknow" to <i>S</i>, END</li> </ol>
--

**Fig. 8.** Answering Q2 - Decision procedure for agent *R*, L2

## 5 Research Questions

We have two experimental conditions, with image only (L1) and with both Image and Reputation (L2). We will explore several values of the parameters in order to show how and where there is an advantage in using reputation. The hypotheses are:

- H1** Initial advantage: L2 shows an initial advantage over L1, that is, L2 grows faster.
- H2** Performance: L2 performs better as a whole, that is, the average quality at regime is higher than L1. Note that to obtain this result we are hardwiring a limitation in image communication, based on the theory that foresees large amounts of retaliation against mistaken image communications but not on the reputation side.
- H3** Cheaters effect: with a high number of cheaters, L2 tends to drop to L1 levels.

There are other hypotheses that we do not treat yet. They regard the relationship between efficiency, fairness, and the presence of cheaters. Actually, these are not yet really formulated as hypotheses but as questions.

- H3.B** Cheater effect: are cheaters always detrimental to the system? In particular, is the performance of the system always decreasing in the number of cheaters?
- H4** Fairness: what is the order relation between L1 and L2 in terms of fairness? For the calculation of fairness we can use simple measures of distribution in quality (averaged, accumulated).
- H5** Cheaters' advantage: do the cheaters effectively reach a significant advantage from their behavior?

## 6 Simulation Runs and Result Analysis

We have run simulations to examine the relationship between L1 and L2 with different levels in some parameters. The stock is fixed at 50, the number of

buyers to 25, and the number of sellers at 100. We included cheaters as well with percentages of 0%, 25% and 50%.

We run the simulations for 100 steps, and we explored the variation of good and bad sellers, from the extreme case of 1% of good sellers and 99% of bad sellers(A1), going through 5% good sellers and 95% bad sellers(A2), and 10% good sellers and 90% bad sellers(A3), and finally, to another extreme where we have 50% of good sellers and 50% of bad sellers(A4). For each one of these conditions and for every situation (L1 and L2) we run 10 simulations. In figures we present the accumulated average per turn of a concrete condition in both situations, L1 and L2.

### 6.1 Experiments without cheaters

In figure 9 we show results for the four conditions without cheater; both hypotheses H1 and H2 are verified. With the increase of good sellers the difference between L1 and L2 gets smaller, until in condition A4 there is no difference. Because of the good sellers increase, they can be reached by random search and the necessity of communicating social evaluations decreases. In the extreme condition A4, statistically every buyer would find a good seller at the second turn (there is a probability of 50% to get one in one turn). In condition A3 the probability to reach one good seller per turn is 0.1, then, in 10 turns approximately every one would reach a good one. In L1 the amount of useful communications (different from "Idontknow") is much lower than in L2, due to the fear of retaliation that governs this situation. In conditions where communication is not important, the difference between the levels disappears.

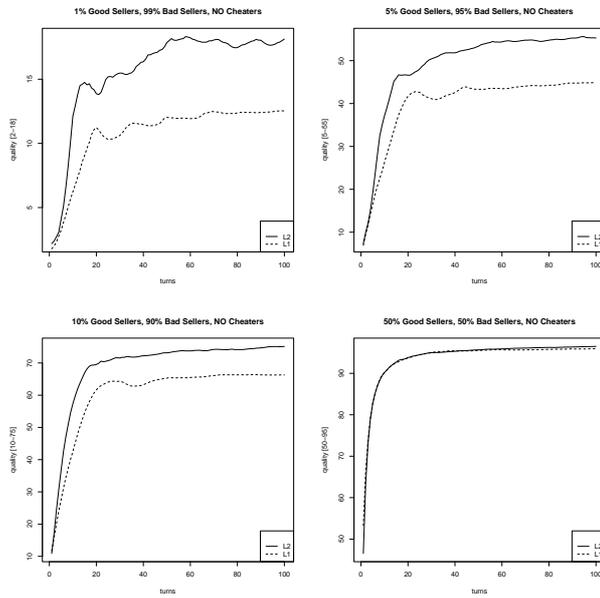
### 6.2 Experiments with cheaters

Figure 10 shows results for conditions A1, A2, A3 and A4 with 50% of cheaters. The increased amount of false information produces a bigger impact in situations and conditions where communication is more important. Quality reached in L1 shows almost no decrease with respect to the experiment without cheaters, while L2 quality tends to drop to L1 levels, supporting the hypotheses H3. This shows how the better performance of L2 over L1 is due to the larger amount of information that circulates in L2.

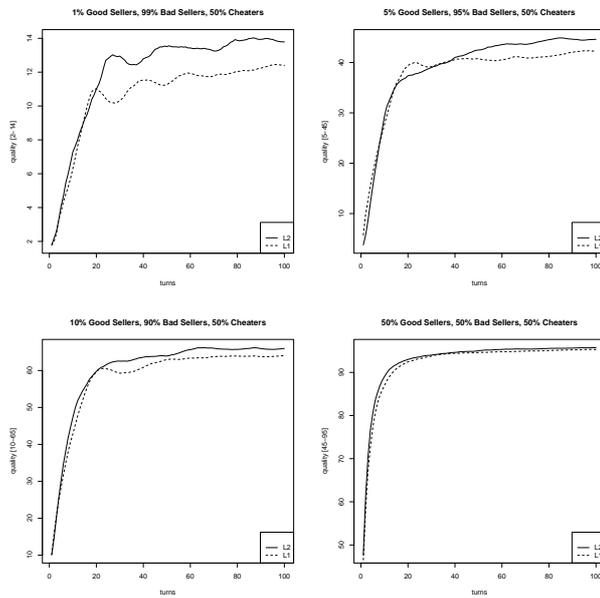
## 7 Conclusions and future work

The results we obtained indicate that using reputation and image instead of only image improve the average quality reached in the whole system. We consider these results as a proof of concept about the usefulness of the reputation model [12], under a set of assumptions that we discuss with a perspective on future works:

**Retaliation:** The presence of retaliation is crucial for the present results. We claim that the fact of communicating a social evaluation that is an image



**Fig. 9.** Accumulated average quality per turn without cheaters in condition A1, A2, A3 and A4 respectively



**Fig. 10.** Accumulated average quality per turn with cheaters in condition A1, A2, A3 and A4 respectively

implies a commitment from the source agent. From the theory, image is a believed evaluation (see section 2) and sharing it implies that the source agent is informing of what he/she *accepts* as true. This differs from reputation, since accepting a reputation do not imply to accept the nested belief. Because of that, sharing what an agent acknowledge as a reputation does not imply a personal commitment. Here we assume that the personal commitment associated to image transmission exposes the agent to a possible retaliation if inaccurate information was sent.

As a future work we will study in more depth the effect of cheaters over the whole system, considering the presence of a norm that prescribes agents to tell the truth, and the reputation mechanism as a social control artifact to exclude agents that do not follow the norm. This is where the hypothesis H3, H4 and H5 we described in section 5 take relevance.

**Communication and reputation:** There is no reputation without communication. Therefore, scenarios with lack of communication or few exchange of information cannot use reputation. However, in virtual societies with autonomous agents that have the freedom to communicate, that need to cooperate and have the right to choose partners, we consider that keeping a separation between image and reputation considerably increases the circulation of information and improves the performance of their activities. In our experiments, even when there is no panelization for direct interactions and considering at each turn only one possible question, the introduction of this difference already improves the average quality per turn. In scenarios where *quality* is scarce and agents are completely autonomous is where this social control mechanism make the difference.

**Decision making procedure:** The decision making schema we implemented (see section 4) determines the performance of the system. In fact, this is where the agent is taking advantage of the distinction between Image and Reputation ([12]). We will elaborate on this distinction, possibly reformulating it in terms of textitmeta decision making, a very promising future line of work to better ground and exploit the Image and Reputation artifacts.

## 8 Acknowledgments

This work was supported by the European Community under the FP6 programme (eRep project CIT5-028575, OpenKnowledge project FP6-027253 and SocialRep project MERG-CT-2005/029161) and the project Autonomic Electronic Institutions (TIN2006-15662-C02-01), and partially supported by the Generalitat de Catalunya under the grant 2005-SGR-00093. Jordi Sabater-Mir enjoys a RAMON Y CAJAL contract from the Spanish Government.

## References

1. Laffont, J.J.: The Economics of Uncertainty and Information. 4th ed. Cambridge, MA: MIT Press (1993)

2. Kudenko, D., Kazakov, D.: Adaptive Agents and Multi-Agent Systems. New York: Springer (2003)
3. Kluegl, F., Bazzan, A.: A case study on the role of information for implicit coordination. In: AAMAS 2004. (2004)
4. Molnr, K.: Learning with expert advice. JEL (2006)
5. Hirshleifer, J.: The private and social value of information and the reward to inventive activity. American Economic Review **61**, no. 4 (1971) 561–574 Reprinted in Uncertainty in Economics: Readings and Exercises. Edited by Peter A. Diamond and Michael Rothschild. Rev. ed. San Diego, CA: Academic Press, 1989. ISBN: 0122148517.
6. Carter, J., Bitting, E., Ghorbani, A.: Reputation formalization for an information-sharing multiagent system. Computational Intelligence **18**(4) (2002) 515–514
7. Yamagishi, T., Matsuda, M., Yoshikai, N., Takahashi, H., Usui, Y.: Solving the lemons problem with reputation. an experimental study of online trading. (2004)
8. Weesie, J., Buskens, V., Raub, W.: The management of trust relations via institutional and structural embeddedness. In Doreian, P., Fararo, T., eds.: The Problem of Solidarity: Theories and Models. Gordon and Breach 1998, Amsterdam (1990) 113–38
9. Yu, B., Singh, M.P.: Detecting deception in reputation management. In: AAMAS '03: Proceedings of the second international joint conference on Autonomous agents and multiagent systems, New York, NY, USA, ACM Press (2003) 73–80
10. Fullam, K., Barber, K.S.: Learning trust strategies in reputation exchange networks. In: AAMAS 2006. (2006)
11. Pradchayakool, P. and Daengdej, J., Supannika, K.: An approach for detecting deception in agents. In: Aerospace Conference IEEE. (2006)
12. Conte, R., Paolucci, M.: Reputation in artificial societies: Social beliefs for social order. Kluwer Academic Publishers (2002)
13. Conte, R., Paolucci, M.: Social cognitive factors of unfair ratings in reputation reporting systems. In: Proceedings of the IEEE/WIC International Conference on Web Intelligence - WI 2003. (2003) 316–322
14. Sabater, J., Paolucci, M., Conte, R.: Repage: Reputation and image among limited autonomous partners. Journal of Artificial Societies and Social Simulation **9**(2) (2006)
15. Miceli, M., Castelfranchi, C.: The Role of Evaluation in Cognition and Social Interaction. In: Human cognition and agent technology. Amsterdam:Benjamins (2000)
16. Sen, S., Sajja, N.: Robustness of reputation-based trust: boolean case. In: AAMAS '02: Proceedings of the first international joint conference on Autonomous agents and multiagent systems, New York, NY, USA, ACM Press (2002) 288–293