Arguing about reputation. The LRep language

Isaac Pinyol and Jordi Sabater-Mir

IIIA - Artificial Intelligence Research Institute
CSIC - Spanish Scientific Research Council
Bellaterra, Barcelona, SPAIN
{ipinyol,jsabater}@iiia.csic.es

Abstract. In the field of multiagent systems (MAS), the computational models of trust and reputation have attracted increasing interest since electronic and open environments became a reality. In virtual societies of human actors very well-known mechanisms are already used to control non normative agents, for instance, the eBay scoring system. In virtual societies of artificial and autonomous agents, the same necessity arises, and several computational trust and reputation models have appeared in literature to cover this necessity. Typically, these models provide evaluations of agents' performance in a specific context, taking into account direct experiences and third party information. This last source of information is the communication of agents' own opinions. When dealing with cognitive agents endowed with complex reasoning mechanisms, we would like that these opinions could be justified in a way such that the resulting information was more complete and reliable. In this paper we present LRep, a language based on an existing ontology of reputation that allows building justifications of communicated social evaluations.

1 Introduction

The field of multiagent systems has experienced an important growth and evolution in the past few years. These systems can be seen as virtual societies composed of autonomous agents where there is a need to interact with other members of the society to achieve their goals. As in human societies, these interactions usually involve an exchange of information. The problem of partners selection via the detection of good or bad potential partners, or how agents evaluate the credibility of received information, arises in a scenario like this. Human societies, along its history, have been using trust and reputation mechanisms for this purpose. These powerful social control artifacts have been studied from different perspectives, such as psychology (Bromley [1], Karlins et al. [2]), sociology (Buskens [3]), philosophy (Plato [4], Hume [5]) and economy (Marimon et al. [6], Celentani et al. [7]).

In multiagent systems the interest on these mechanisms has considerably increased and, as a consequence, numerous computational trust and reputation models have appeared in the literature. E-Commerce sites already use some of them (eBay [8], Amazon [9], OnSale [10]). These models consider reputation as a centralized global property. So, the reputation of each agent is public and all agents perceived the same reputation value. More sophisticated models ([11], [12], [13], [14], [15], [16], [17]) consider reputation as a subjective property. Therefore every agent has its own reputation system that provides evaluations of other agents calculated from external communication and direct experience, giving the agent its own vision of the society. Furthermore, other models (see [18], [19]) take into account social information when providing these evaluations.

One of these models is RepAge [17], a computational system based on a cognitive theory of reputation. This model is designed to be part of a cognitive agent,i.e., an agent endowed with beliefs, desires and intentions. Like other reputation models, RepAge uses social evaluations obtained from direct experiences and communicated social evaluations as source for calculations. However, this communication is quite simple and very limited, allowing only the exchange of single values associated reliability measure. In a real environment and for an agent that is able to make complex reasoning, an opinion without being justified can be very weak and not as useful as a fully justified opinion that points out where the information is coming from. With agents using a complex reputation model like RepAge, as important can be the followed procedure and the sources to calculate the final value as the final value itself.

In this paper we present LRep, a simple language that can be used with a model like RepAge to elaborate justifications of calculated values. These justifications can have different levels of detail. So, agents can decide the amount of extra information and the level of detail of them when there are communicating social evaluations.

In Section 2 we briefly introduce RepAge and its theory framework. Following this, in Section 3 we introduce an ontology of reputation and its specification using description logic. This ontology will be used to define the semantics of LRep. Afterwards, in Section 4 we define the syntax and semantics of LRep. In Section 5 we present several situations where the use of LRep and justification helps to improve the performance of cognitive agents. Finally, Section 6 presents the conclusions and future work.

2 The RepAge system

In order to present the RepAge system it is necessary to get in touch with the theoretical framework in which it is based. This framework is a cognitive theory of reputation developed by Conte and Paolucci in [20]. In this book they study the impact of the transmission of social evaluations in artificial societies, pointing out the important difference between information that is though to be true and information that is said.

RepAge is based on a model of imAGE, REPutation and their interplay developed in [20]. Although both are social evaluations, image and reputation are distinct objects. Image is a simple evaluative belief; 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.

Furthermore, agents using RepAge require to be cognitive to use all its potential. A cognitive agent has beliefs, goals and intentions that drive its deliberative process.

2.1 The RepAge architecture

As we stated, RepAge [17] is a computational system based on a cognitive theory of reputation [20] that proposes a fundamental distinction between image and reputation. The RepAge architecture 1 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 antecedents and a set of consequents. 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 informant has about a target, the image that, according to the informant, a third party agent has on the target, and the reputation that the informant 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 informant as informant. Communications from agents whose credibility is low will not be considered as strong as the ones coming from well reputed informants. 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 [17] 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 contain all the reasoning procedures uses the predicates generated by RepAge to make decisions.

3 The ontological dimension of reputation

As we have shown so far, reputation mechanisms play a crucial role in the way we conceive agents' societies. But social evaluations are more than simple ratio scores. In cognitive agents the fact of acknowledging certain reputation or image of other agents imply a mental state, a set of beliefs about the future performance of target agents, but at the same time, the formation of such a high

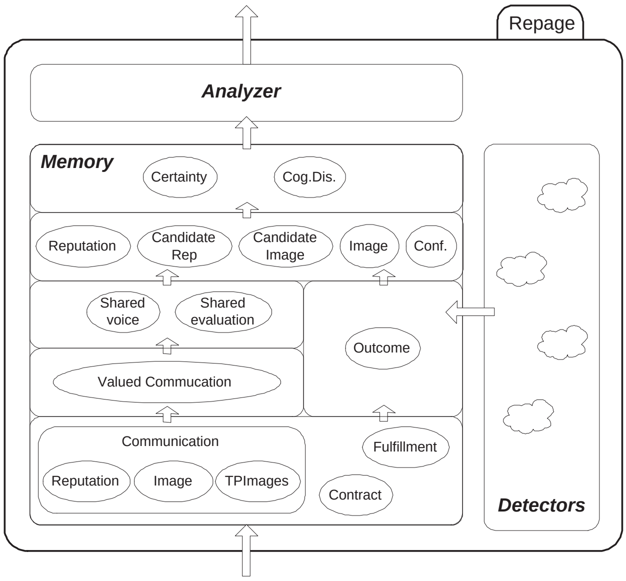

Fig. 1. The RepAge architecture

level predicates, require several intermediate cognitive steps, that generate a full taxonomy of interrelated predicates. From this point of view, it is easy to think about an ontology of reputation and image showing this structure. A possible ontology is defined in [21].

The concepts that appear in this ontology are very similar to the typology of predicates that RepAge (see previous section) defines. In [21] we define a mapping between RepAge predicates and the ontology (that is almost direct). Still though, we want to use as source of information this common ontology, since is not linked to any particular reputation model. A graphical representation of it is shown in Figure 2 and 3. However, a more formal approach is needed. Because LRep language is based on this ontology we need a formalism that allows us to refer instances of its concepts. For that, we decided to use description logic(DL). As we will explain, DL offers an elegant way to represent application domains, and its concepts have been used for the semantic web (in term of the language OWL DL) to describe ontologies. Furthermore, its syntax and semantics is very well known and accurately define(see [22]). In this section we first make a short introduction to what is a description logic system and why it is a good option to use as an ontology formalism. Afterwards, we give a description of the ontology using this formalism.

3.1 Description Logic

Description Logic (DL) is a knowledge representation formalism used to represent the application domain, the world. Its power relies on the formal logic-based semantics and the reasoning engine with which it is equipped. A DL system has two differentiate submodules, *TBox* and *ABox*.

On one side, the *TBox* contains a set of expressions in one of the languages of the $\mathcal{AL}$-languages family (see [22]), that define the terminology of the domain (the classes). These family of languages can be seen as fragments of first-order Logic(FOL) [22][1], but its powerful expressiveness simplifies the expressions and is specially suited for the definition of concepts. On the other side, the *ABox* contains assertions about named individuals in terms of the terminology defined in the *TBox*, the state of the world. In general, a knowledge representation system based on DL provides facilities to set and update knowledge bases, to manipulate it and to reason over it.

Because DL systems has a semantics that identifies its description language as segments of FOL, the set of predicates contain implicit knowledge, that can be made explicit using inference. Thus, the concept of satisfiability is defined in the classical way (see [22]). Having a DL system D, a concept C and an element a , we say that $D \models C(a)$ iff $C(a)$ can be inferred from D, that is, if it is deducible using some of the complete reason algorithms defined for DL systems (Structural Subsumption Algorithm or Tableau Algorithm, for instance)[22].

Nowadays, the interest of DL systems has considerably increased due to the popularity of ontologies for the semantic web and specifically because of the OWL language. The semantic web is using as standard the *OWL* language to structure knowledge contained in web sites, so to describe ontologies. This language(OWL) has three variants, one of them is *OWL DL*, a language that uses the concepts of description logic we have explained in this section.

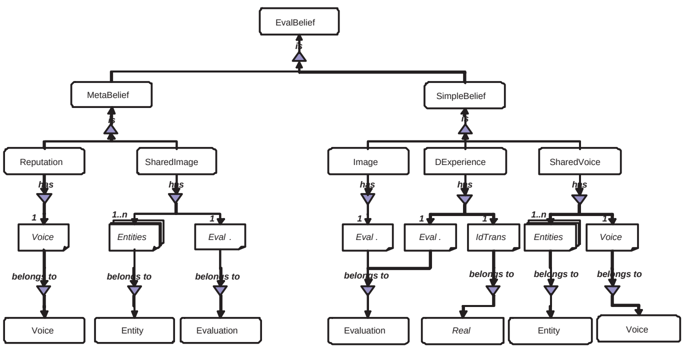

Fig. 2. The taxonomy, membership relations and components of evaluative beliefs

3.2 A DL version of the ontology

Having the ontology showed in Figure 2 and 3, few things have to be explained to get in touch with all necessary concepts.This ontology defines a taxonomy of

[1] So, all formulas of $\mathcal{AL}$-languages can be expressed as FOL formulas keeping the same semantic

evaluative beliefs, that represents beliefs that have some social evaluations. We divided them into *SimpleBelief* and *MetaBelief*. This division is conceptually important when talking about cognitive agents. An agent holding a simple belief acknowledge the evaluation that the belief contains, meanwhile a Metabelief is a belief about others' belief, an interpretation of what other agents think. Therefore, an agent holding a Metabelief do not need to believe the nested evaluation. For instance, I can belief that my friend thinks that his car is nice, but I don't necessary agree with this opinion. Then, we consider an Image, Direct Experiences and a SharedVoice as simple beliefs, and Reputation and Shared Images as Metabeliefs (see [21] for the details of this decision). The meaning is the same that have the predicates in RepAge. A direct experience should be understood as an outcome predicate in RepAge.

All these concepts are located in the bottom part of the taxonomy. A system using this ontology will have instances of these bottom concepts. All of them have at least as attribute an object *Evaluation* that contains information about the evaluation itself. Part of this information is the value of the evaluation, the representation of goodness and badness. In literature there are several possible representations, from simple boolean with bad/good, to probability distributions over some sorted set (like in the case of RepAge). In [21] four representation types are described, including transformation operations between them. For the sake of simplicity, in this first approach we will use a simple sorted labeled set, *VB, B, N, G, VG* meaning, *Very Bad*, *Bad*, *Neutral*, *Good* and *Very Good*.

At this point we have all the elements to understand a description of the ontology, that corresponds with the *TBox* of a DL system[2]:

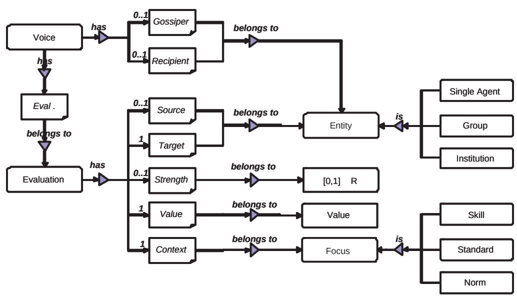

Fig. 3. The main classes and components of a social evaluation and voice

[2] The semantics of $(\leq\ nR)$ and $(=\ nR)$ is defined as $(\leq\ nR)^I = \{a \in \triangle^I where |\{b|(a,b) \in R^I\}| \leq n\}$ and $(= nR)^I = \{a \in \triangle^I where |\{b|(a,b) \in R^I\}| = n\}$, where I is an interpretation, $\triangle^I$ the domain of the interpretation, and R^I the interpretation of the relation R

$$
\begin{aligned}
Entity &\equiv SingleAgent \sqcup Group \sqcup Institution \\
Focus &\equiv Skill \sqcup Standard \sqcup Norm \\
Evaluation &\equiv \leq 1 hasSource.Entity \sqcap = 1 hasTarget.Entity \sqcap \\
&\equiv \sqcap = 1 hasContext.Focus \leq 1 hasStrength.\mathbb{R} \sqcap \\
&\equiv \sqcap \leq 1 hasValue.Value \\
Voice &\equiv \leq 1 hasGossiper.Entity \sqcap \leq 1 hasRecipient.Entity \sqcap \\
&\equiv = 1 hasEval.Evaluation \\
Image &\equiv SimpleBelief \sqcap = 1 hasEval.Evaluation \\
DExperience &\equiv SimpleBelief \sqcap = 1 hasEval.Evaluation \sqcap = 1 hasTrans.\mathbb{R} \\
ShVoice &\equiv SimpleBelief \sqcap = 1 hasVoice.Voice \sqcap \exists hasGossiper.Entity \\
ShImage &\equiv MetaBelief \sqcap = 1 hasEval.Evaluation \sqcap \exists hasSource.Entity \\
Reputation &\equiv MetaBelief \sqcap = 1 hasVoice.Voice
\end{aligned}
$$

In this case we consider as primitive concepts *SingleAgent*, *Group*, *Institution*, *Skill*, *Standard* and *Norm*. The concept *Value* is used to define the predicates *Value(VB)*, *Value(B)*, *Value(N)*, *Value(G)* and *Value(VG)*, as axioms of the system. All other concepts are defined using the $\mathcal{ALUN}$-language (see [22]).

4 The LRep: language framework for reputation and image justification

In this section we define both the syntax and semantics of the LRep language. The objective of this language is to provide a mechanism to represent not only the evaluation of an image or reputation but also a justification of that value. This justification should increase the richness of the exchanged information about image and reputation and therefore it increases the effectiveness of spreading reputation and image. That justification can be sometimes even more relevant than the evaluation itself (see section 5).

First, we will define the syntax of the language giving an informal semantics. Finally, we will give a formal semantics of the language.

4.1 Defining the basis of LRep

Let $A = \{a_1, \ldots, a_n\}$, $R = \{r_1, \ldots, r_m\}$ and $V = \{VB, B, N, G, VG\}$ be a set of agents, a set of roles, and a sorted set of evaluation labels respectively. We define the set *Eval* of all possible evaluations and evaluation as follows:

$$Eval = \{< a, r, v > | a \in A, r \in R, v \in V\} \tag{1}$$

We define a set of predicate letters P, and a set of quantifier letters N

$$P = \{I, R, ShI, ShV, DE, CI, CI_1, \ldots, CI_n, CR, CR_1, \ldots, CR_n\} \tag{2}$$

$$N = \{N_1, \ldots, N_n\} \tag{3}$$

Intuitively, the letters I, R, ShI and ShV refer to evaluations that are *Image*, *Reputation*, *Shared Image* and *Shared Voice*. The predicates CI, and CR refer to *Communicated Image* and *Communicated Reputation*. Concretely , CI_i and CR_i reefer to a *Communicated Image* and *Communicated Reputation* from an agent $a_i \in A$. DE refers to a *Direct Experience*. Notice that in the ontology, this predicate has an object evaluation and a real value. This second one refers to an identification number of the transaction that produced the direct experience.

4.2 Simple Predicate Formula (SPF) and Extended Predicate Formula (EPF)

Formulas in the LRep language are divided in SPF and EPF.

Let $e \in Eval$, $t \in \mathbb{R}$ and $1 \leq i \leq n$, then the following formulas are SPF:

- $I(e), R(e), ShI(e), ShV(e)$
- $DE(e,t), CI_i(e), CR_i(e)$

Let $1 \leq i \leq n$ and $e \in Eval$ then

- $\emptyset$ (empty formula) is an EPF
- If α is SPF then α is EPF
- If α is SPF then $N_i\alpha$ is EPF
- The formulas $N_iDE(e)$, $N_iCI(e)$ and $N_iCR(e)$ are EPF
- Inductively, if β and γ are EPF, then $\beta;\gamma$ is EPF

Intuitively, N_iX means that there have been received at least i communicated images or communicated reputations, or that the agent has had at least i direct experiences[3]. The formal semantics of the quantifier is defined in Section 4.4.

4.3 Justification

We define a justification in terms of LRep language as follows. Let α be a SPF and γ be an EPF, then a LRep formula is defined as:

$$\{\alpha : \gamma\} \tag{4}$$

The idea is that in the expression $\{\alpha : \gamma\}$ that is a *LRep* formula, the α predicate is the main element to communicate, and it is *justified* by the formula γ, that in fact it is a list of less generic predicates. For example we can have justifications like this:

[3] We decide N_i to be an lower bound instead of an exact number because this second one is too restrictive and leads to an honest-liars communication, forgetting the interesting option of telling a truth information but not exact. We have in mind to include in the future negative connective that will allow setting upper and lower bound

$$\{I(< a_1, r_1, VB >) : N_5CI(< a_1, r_1, B >); N_3DE(< a_1, r_1, VB >);$$
$$CI_{a_3}(< a_1, r_1, VB >)\}$$

Meaning that the Image of a_1 towards the role r_1 is very bad *because* we have received more than 5 communicated images saying that a_1 in r_1 is bad, we have experienced more that 3 times that the agent is very bad, and because a_3 communicated us that a_1 in the context r_1 is very bad. Of course, we are not talking about neither the truth of the explanation, nor the truth of the communication itself. Agents can lie, and of course can give partial information.

Given that, the syntax of LRep language can be defined using the following grammatic.

$LRep ::= \{SPF : EPF\}$
$SPF ::= I(E)|R(E)|ShI(E)|ShV(E)|DE(E, \mathbb{N})|Comm$
$Comm ::= CI_{agent}(E)|CR_{agent}(E)$
$E ::= < Target, Context, Value >$
$EPF ::= \emptyset|N_{\mathbb{N}}CI(E)|N_{\mathbb{N}}CR(E)|N_{\mathbb{N}}DE(E)|SPF|EPF; EPF$
$Context ::= norm|standard|skill$
$Target ::= agent|group|institution$
$Value ::= VB|B|N|G|VG$

4.4 Semantic of LRep

To define the formal semantics of the language we have to introduce the concept of *correctness* within a LRep expression. Saying that I had more than 10 direct experiences with a seller when I really had 2 is not correct taking into account my state of the world (where I only had 2 direct experiences). So, the semantics of LRep will be determined for the correctness of the expression towards certain state of the world. Of course, this model of the world will be represented as an instance of a DL system with the *TBox* defined in Section 3.2.

So, let $F =< T, A >$ be a DL system describing the state of the world of an agent, where T is the TBox of terminological terms composed by the concepts defined in Section 3.2, and A the ABox with the assertions describing the state of the world at certain moment of time. We say that a justification $J = \{\alpha : \emptyset\}$ or $J = \{\alpha : \gamma; \beta\}$ is correct towards the system F, written as $F \supset J$ iff each of the components of J is correct towards F. More formally:

$$F \supset \{\alpha : \emptyset\} \leftrightarrow F \supset \alpha$$

$$F \supset \{\alpha : \gamma; \beta\} \leftrightarrow F \supset \alpha \text{ and } F \supset \gamma \text{ and } \supset \beta$$

Then, the correctness of formulas SPF and EPF is defined in terms of the correctness of its atomic elements. For instance, considering the simple formula $DE(< y, r, v >, t)$ its correctness is defined as follows:

$$
\begin{aligned}
F \supset DE(< y, r, v >, t) \leftrightarrow \exists a, e \text{ such that } & \\
& F \models DExperience(a), hasEval(a, e), \\
& hasTrans(a, t) \text{ and } evalFine(e, y, r, v, F)
\end{aligned}
$$

where we define the predicate $evalFine$ as follows:

$$
\begin{aligned}
evalFine(e, y, r, v, F) = True \leftrightarrow\ & F \models hasTarget(e, y) \text{ and} \\
& F \models hasContext(e, r) \text{ and} \\
& F \models hasValue(e, v) \text{ and} \\
& F \models Value(v), Focus(r), Entity(y)
\end{aligned}
$$

Following the same idea, the correctness of all atomic elements of LRep is defined in the next table:

$$
\begin{aligned}
F \supset I(< y, r, v >) \leftrightarrow \exists a, e \text{ such that } & \\
& F \models Image(a), hasEval(a, e), \\
& \text{and } evalFine(e, y, r, v, F)
\end{aligned}
$$

$$
\begin{aligned}
F \supset CI_x(< y, r, v >) \leftrightarrow \exists a, e \text{ such that } & \\
& F \models ShImage(a), hasSource(a, x), hasEval(a, e) \\
& \text{and } evalFine(e, y, r, v, F)
\end{aligned}
$$

$$
\begin{aligned}
F \supset CR_x(< y, r, v >) \leftrightarrow \exists a, v, e \text{ such that } & \\
& F \models ShVoice(a), hasGossiper(a, x) \\
& hasVoice(a, v), hasEval(v, e) \\
& \text{and } evalFine(e, y, r, v, F)
\end{aligned}
$$

$$
\begin{aligned}
F \supset R(< y, r, v >) \leftrightarrow \exists a, v, e \text{ such that } & \\
& F \models Reputation(a), hasVoice(a, v) \\
& hasEval(v, e) \text{ and } evalFine(e, y, r, v, F)
\end{aligned}
$$

$$
\begin{aligned}
F \supset N_iDE(< y, r, v >) \leftrightarrow\ & |A| \geq i \text{ where } A = \{DExperience(a) | \exists e, t \\
& \text{such that } F \models hasEval(a, e), hasTrans(a, t) \\
& \text{and } evalFine(e, y, r, v, F)\}
\end{aligned}
$$

$$
\begin{aligned}
F \supset N_iCI(< y, r, v >) \leftrightarrow\ & |A| \geq i \text{ where } A = \{Entity(x) | \exists a, e \\
& \text{such that } F \models ShImage(a), hasSource(a, x) \\
& hasEval(a, e) \text{ and } evalFine(e, y, r, v, F)\}
\end{aligned}
$$

$$
\begin{aligned}
F \supset N_iCR(< y, r, v >) \leftrightarrow\ & |A| \geq i \text{ where } A = \{Entity(x) | \exists a, e \\
& \text{such that } F \models ShVoice(a), hasGossiper(a, x) \\
& hasVoice(a, v),\ hasEval(v, e) \\
& \text{and } evalFine(e, y, r, v, F)\}
\end{aligned}
$$

5 Using LRep

In this section we apply LRep in a concrete scenario. Let A be the set of agent names A ={*John*, *Debra*, *Laura*,...} and R a set of roles R ={*seller*, $informant$, *buyer*}. In this environment, everybody can play the three roles. In a typical transaction, an agent acting as a *buyer*, buys a specific product from another agent that acts as a *seller*. Also, there is the possibility to exchange information about other agents' performance, acting then as an *informant*. Agents are cognitive and use the RepAge model to deal with social evaluations. In this case they evaluate agents as *sellers* (whether they sell the products with the maximum quality, as they claim) and as *informants* (since they may not provide accurate information or even they may lie). Currently the exchange of social evaluations is done in terms of Image or Reputation. As shown in Section 2 there is an implicit commitment sending an Image (since it is the agent's own opinion) that does not exists when sending Reputation.

After introducing the scenario, we expose several cases where by using a justification, ambiguous situations become clearer and communications richer.

5.1 Case 1: Discrimination between weak and strong predicates

One of the main issues when exchanging social evaluations is the inherent subjectivity they have associated. Check for instance, the following communications:

C1:$\{I(< John, seller, VG >)\}$
C2:$\{I(< John, seller, VG >) : N_2DE(< John, seller, VG >)\}$
C3:$\{I(< John, seller, VG >) : N_{20}DE(< John, seller, VG >)\}$

The first communication, $C1$, indicates that the image the informant has of *John* as a *seller* is VG (very good). However, it does not tell us anything about the strength of it. Communications $C2$ and $C3$ show us some more details. Assuming that agents send correct information towards its vision of the world (in the sense we define in Section 4.4), we should agree that the justification in $C3$ gives more reasons to belief the communicated image than $C2$. And in this sense, communicated image in $C3$ is *stronger* than the one in $C2$ and of course than the one in $C1$. There are some models that represent this by using a value of reliability. However this reliability is calculated by the sender so again it is something subjective. For instance, in our example, all the informants have considered they had enough evidences to build up an Image, so, in the three communications the reliability value would be high. In terms of reputation we can have similar situations.

C1:$\{R(< John, seller, VG >)\}$
C2:$\{R(< John, seller, VG >) : N_2CR(< John, seller, VG >)\}$
C3:$\{R(< John, seller, VG >) : N_{20}CR(< John, seller, VG >)\}$

5.2 Case 2: Avoiding unreliable information

Another case where the use of LRep helps in the better understanding of the messages, is in the detection of information that should not be taken into account because the justification contradicts the state of the world that the recipient has. For instance, check the following justification:

$$
\begin{aligned}
\{I(< John, seller, B >) : \; & CI_{Laura}(< John, seller, B >); \\
& CI_{Debra}(< John, seller, VB >); \\
& I(< Laura, informant, VG >); \\
& I(< Debra, informant, VG >)\}
\end{aligned}
$$

In this case, the informant justifies its image of *John* as a seller pointing out that he has received two communicated images, one from *Laura* and another from *Debra* (that are considered very good informants), saying that *John* is mostly bad. However, if the recipient of the message has an image of *Laura* and *Debra* as informants that is very bad the image of *John* cannot be considered, at least without further knowledge that can solve the contradiction.

5.3 Case 3: Control of granularity

One interesting property that LRep has is the granularity of its predicates. In this sense, even in this first version it is already possible to give more and more detailed information to properly justify a predicate. For instance, consider the following communication:

$$
\begin{aligned}
\{R(< Laura, seller, VG >) : \; & ShV(< Laura, seller, G >); \\
& ShV(< Laura, seller, VG >)\}
\end{aligned}
$$

Here, this agent is justifying a reputation by means of two shared voices that at the same time are justified as follows:

$$
\begin{aligned}
&\{ShV(< Laura, seller, G >) : N_1CR(< Laura, seller, G >)\} \\
&\{ShV(< Laura, seller, VG >) : N_2CR(< Laura, seller, VG >)\}
\end{aligned}
$$

Another possible and more detailed justification of the two shared voices could be:

$$
\begin{aligned}
\{ShV(< Laura, seller, G >) : \; & CR_{Debra}(< Laura, seller, G >)\} \\
\{ShV(< Laura, seller, VG >) : \; & CR_{John}(< Laura, seller, VG >); \\
& CR_{John}(< Jorge, seller, VG >)\}
\end{aligned}
$$

Therefore, this justification could have included some information about the images of the informants, that supposedly are good. And these images, can be justified with the detail that the agent considers appropriate. The point of this discussion is to make the reader notice that using LRep, agents can reach the level of detail they want in the justifications.

5.4 Case 4: Putting everything together. Dialogs

Finally, extending LRep by allowing questions we can establish dialogs between two agents. In the following example we have agents A_1 and A_2 exchanging information. Initially, A_1 sends an image without any justification.

$$A_1 \rightarrow A_2 : \{I(< Laura, informant, VG >)\}$$

At this point, A_2 does not know A_1 very well, then it asks for more information:

$$\begin{aligned}
&A_2 \rightarrow A_1 : \{I(< Laura, informant, VG >)?\} \\
&A_1 \rightarrow A_2 : \{I(< Laura, informant, VG >) : \\
&\qquad CI_{Laura}(< Debra, seller, VB >); I(< Debra, seller, VB >) \\
&\qquad CI_{Laura}(< John, seller, VG >); I(< John, seller, VG >)\}
\end{aligned}$$

Again A_2 is not satisfied. It wants to know how the images about *Debra* and *John* where formed, so, it asks for it:

$$\begin{aligned}
&A_2 \rightarrow A_1 : \{I(< Debra, seller, VB >)?\} \\
&A_2 \rightarrow A_1 : \{I(< John, seller, VG >)?\} \\
&A_1 \rightarrow A_2 : \{I(< Debra, seller, VB >) : N_3DE(< Debra, seller, VB >)\} \\
&A_1 \rightarrow A_2 : \{I(< John, seller, VG >) : N_2DE(< John, seller, VG >)\}
\end{aligned}$$

Now, A_2 knows that the original information about *Laura* as *informant* is very good for A_1 because is based on that once, *Laura* gave information about *Debra* and *John* as very good and very bad sellers respectively, and that A_2 experienced with both of them observing that they behaved in the same way that *Laura* said.

Knowing this, the conclusions that A_2 may get depend on its own state of the world, its beliefs:

- **Ignore the information**: On one side it may have already had some direct experiences with *Debra* and *John* and they behaved the opposite of what A_1 claims. In this case, the information that *Laura* as *informant* is very good is not reliable for A_2.

- **Take the information as reliable**: On the other side, the evaluations of the direct experiences that A_2 and A_1 had with *Debra* and *John* may coincide, and then, A_2 may consider the original information reliable.

- **Need for more information**: Another case may come out when for instance, A_2 does not have any information about *John* or *Debra*. In this case, if for A_2 the original information is important enough and have the chance to do it, it may interact with both to acquire first hand experiences, or may be it may ask to another agent (with good image as informant) to contrast the information. The idea is that in justifications, every piece of information can be contrasted, either by direct experiences or by communications. So, in this example, the number of possible actions is quite high.

6 Conclusions and Future Work

As we stated from the beginning, we are dealing with cognitive agents. In our case it means agents that have beliefs, desires, intentions and goals to accomplish and that are able to reason about them. This is the context where a language like LRep has sense. By exchanging not only simple image/reputation values but justifications of these values, we are opening the possibility to reason about the process the informant followed to build those values and not only about the values themselves. Talking about image and reputation, and as we have shown with some examples in section 5, that extra information can be as useful as the value itself.

An important aspect of LRep is that the informant can decide how deep the justification has to be. So going from no justification at all to the exact details of the calculation, there is a wide range of possibilities from where the informant can choose according to its confidence with the receiver. Furthermore, the fact of using a common ontology of reputation for the LRep semantics, allows to apply LRep in other reputation models.

Future experiments are planed to be done using RepAge and LRep. In a scenario like the one described in [23], we have a set of buyers and sellers. Sellers sell items with certain quality (from a predefined minimum and maximum), and buyers want to buy always the maximum quality. Providing the agents with RepAge system, in [23] several experiments were run to observe the performance of the buyers per turn, varying several parameters (like number of sellers or buyers) and dealing with cheaters. The incorporation of the LRep language in this simulations will require two parallel phases. On one side, design more sophisticates decision making processes to take advantage of this new functionality, and on the other side, study the impact, the creation and motivation of sending false information in justifications.

Besides this, the LRep language is very simple, almost every atomic element in LRep coincide with an element of the ontology. Only the quantifiers define more sophisticate semantics. Extensions of LRep are expected, for instance, including universal or existential quantifiers. Also, more sophisticates protocols of communication should be taken into account.

7 Acknowledgments

This work was supported by the European Community under the FP6 programme (eRep project CIT5-028575 and OpenKnowledge project FP6-027253). Jordi Sabater-Mir enjoys a Ramon y Cajal contract from the Spanish Government.

References

1. Bromley, D.B.: Reputation, Image and Impression Management. John Wiley & Sons (1993)
2. Karlins, M., Abelson, H.I.: Persuasion, how opinion and attitudes are changed. Crosby Lockwood & Son (1970)
3. Buskens, V.: The social structure of trust. Social Networks (20) (1998) 265—298
4. Plato: The Republic (370BC). Viking Press (1955)
5. Hume, D.: A Treatise of Human Nature (1737). Oxford: Clarendon Press (1975)
6. Marimon, R., Nicolini, J.P., Teles, P.: Competition and reputation. In: Proceedings of the World Conference Econometric Society, Seattle. (2000)
7. Celentani, M., Fudenberg, D., Levine, D.K., Psendorfer, W.: Maintaining a reputation against a long-lived opponent. Econometrica **64**(3) (1966) 691—704
8. eBay: eBay, http://www.eBay.com. (2002)
9. Amazon: Amazon Auctions, http://auctions.amazon.com. (2002)
10. OnSale: OnSale, http://www.onsale.com. (2002)
11. Abdul-Rahman, A., Hailes, S.: Supporting trust in virtual communities. In: Proceedings of the Hawaii's International Conference on Systems Sciences, Maui, Hawaii. (2000)
12. Esfandiari, B., Chandrasekharan, S.: On how agents make friends: Mechanisms for trust acquisition. In: Proceedings of the Fourth Workshop on Deception, Fraud and Trust in Agent Societies, Montreal, Canada. (2001) 27—34
13. Schillo, M., Funk, P., Rovatsos, M.: Using trust for detecting deceitful agents in artificial societites. Applied Artificial Intelligence (Special Issue on Trust, Deception and Fraud in Agent Societies) (2000)
14. Yu, B., Singh, M.P.: Towards a probabilistic model of distributed reputation management. In: Proceedings of the Fourth Workshop on Deception, Fraud and Trust in Agent Societies, Montreal, Canada. (2001) 125—137
15. Carbo, J., Molina, J., Davila, J.: Trust management through fuzzy reputation. Int. Journal in Cooperative Information Systems (2002) in–press
16. Sen, S., Sajja, N.: Robustness of reputation-based trust: Boolean case. In: Proceedings of the first international joint conference on autonomous agents and multiagent systems (AAMAS-02), Bologna, Italy. (2002) 288—293
17. Sabater, J., Paolucci, M., Conte, R.: Repage: Reputation and image among limited autonomous partners. J. of Artificial Societies and Social Simulation **9**(2) (2006)
18. Carter, J., Bitting, E., Ghorbani, A.: Reputation formalization for an information-sharing multi-agent sytem. Computational Intelligence **18**(2) (2002) 515—534
19. Sabater, J., Sierra, C.: Regret: A reputation model for gregarious societies. In: Proceedings of the Fourth Workshop on Deception, Fraud and Trust in Agent Societies, Montreal, Canada. (2001) 61—69
20. Conte, R., Paolucci, M.: Reputation in artificial societies: Social beliefs for social order. Kluwer Academic Publishers (2002)
21. Pinyol, I., Sabater-Mir, J., Cuni, G.: How to talk about reputation using a common ontology: From definition to implementation. In: Proceedings of the Ninth Workshop on Trust in Agent Societies. Hawaii, USA. (2007) 90—101
22. Baader, F., McGuinness, D., Nardi, D., (eds.), P.P.S.: The description logic handbook. Cambridge university press (2003)
23. Pinyol, I., Paolucci, M., Sabater-Mir, J., Conte, R.: Beyond accuracy. reputation for partner selection with lies and retaliation. In: Proceedings of the Eighth International Workshop on Multi-Agent-Based Simulation. Hawaii, USA. (2007) 134—146