Towards a Functional Ontology of Reputation

Sara Casare
Intelligent Techniques Laboratory
University of São Paulo, Brazil
Av. Prof. Luciano Gualberto, 158 – trav. 3
05508-900 São Paulo SP BRAZIL
+55 11 30821968
sara.casare@poli.usp.br

Jaime Sichman
Intelligent Techniques Laboratory
University of São Paulo, Brazil
Av. Prof. Luciano Gualberto, 158 – trav. 3
05508-900 São Paulo SP BRAZIL
+55 11 30915397
jaime.sichman@poli.usp.br

ABSTRACT
This paper proposes a functional ontology of reputation for agents. The goal of this ontology is twofold. First, to put together the broad knowledge about reputation produced in some areas of interest such as psychology and artificial intelligence, mainly multi-agent systems. Second, to represent that knowledge in a structured form. The functional ontology of reputation employs the primitive categories of knowledge used in the Functional Ontology of Law proposed by Valente [16]. We claim that the concepts of the legal world can be used to model the social world, through the extension of the concept of legal rule to social norm and the internalization of social control mechanisms in the agent's mind, so far externalized in legal institutions.

Categories and Subject Descriptors
I.2.11 [Distributed Artificial Intelligence]: Multiagent systems

General Terms
Design, Experimentation, Standardization

Keywords
Reputation, Ontology, Interoperability

1. INTRODUCTION
Reputation is an indispensable condition for the social conviviality in human societies. The emergence of Internet based virtual societies has caused the migration of reputation related concepts from the world of human interactions to the world of virtual interactions.

The analysis of several computer based reputation models [5; 13; 17; 18 among others] indicates that the notion of reputation is used in an intuitive way. In most cases these models use neither a precise definition of reputation nor the theoretical or empiric bases from disciplines that have worked with reputation concepts much longer than Artificial Intelligence (AI), such as economy, sociology and psychology [1; 4; 9; 14; among other]. Several authors associate reputation to trust and reciprocity [1; 7; 11; 12]. Castelfranchi et al [7] consider reputation as a component of trust, while Ostrom [12] suggests that reciprocity, reputation and trust can help to overcome the agent temptations of short-run self-interest.

This work is organized as follows. Section 2 presents several reputation concepts based on the knowledge developed mainly in two research areas, Psychology and AI. Section 3 outlines a preliminary functional ontology of reputation. The goal of this ontology is to aggregate knowledge about reputation produced in different areas of interest, as well as to represent that knowledge in a structured form. In section 4, it is presented a possible usage of this ontology, illustrated by a hypothetical scenario. Finally, section 5 presents a discussion about the ontology produced so far, as well as some ideas for further work.

2. REPUTATION CONCEPTS
Reputation is a social product as well as a social process. It is a product, or property, in the sense that it consists of opinion agreement in some level; on the other hand, it may be seen as a process in the sense that there is a flow of information and influence in the social network. While reputation as a product may be seen as a cognitive representation (or a belief), reputation as a process consists of a set of beliefs’ transmission in the social network [4, p. 217; 8, p. 72].

This section defines the concepts used to identify the several aspects of a reputation, both as a product and as a process: reputation nature, roles involved in reputation formation and propagation, information sources for reputation, evaluation of reputation, reputation scope and reputation maintenance. Those aspects are based on Psychology [4; 9] and AI [8; 10; 11; 13; 17] research results, and constitute a set of ontological terms used in a functional ontology of reputation, presented in section 3.

2.1 Reputation Nature
The nature of the reputation distinguishes a reputation according to the kind of entity it is associated to. According to Bromley [4, p.4] there are different sorts of reputational entities, given that persons, groups of people and corporations can have a reputation, as well as products, services, geographical areas, and events in a general way, including activities. When associated to products...
and companies that manufacture them, a reputation is known as a brand image and corporate image, respectively. Activities, such as smoking and exercising, can have good or bad reputation, and so can events, such as carnival, and places like cities and countries.

Bromley [4, p. 6] points out that reputation, considered itself as a social phenomenon, can be a reputational entity, given the ambivalent feelings we have about the concept. Such ambivalence is illustrated by the way the word is used in daily life, both as a valuable and important social phenomenon, and as based on hypocrisy and reflecting accidental events.

2.2 Reputation Roles
Reputation formation and propagation involves several roles, played by the entities or agents that participate in those processes. According to Conte and Paolucci [8], reputation involves four distinct sets of agents: the evaluators, the targets, the beneficiaries, and the propagators.

Evaluators are those entities able to develop an evaluation or an evaluative belief about others as an effect of their social interactions and perceptions. Target entities are individuals, groups, or even artifacts that play the role of the evaluation object. Beneficiaries are individuals, groups, or organizations for which the evaluation of the target brings some benefit. Propagators, or third parties, are those entities able to transmit reputation information about a target to another entity. As the authors emphasize, there is a non-empty intersection between these four sets of agents. Therefore, an agent can be at the same time a member of the evaluators set as well as a member of the propagator and the beneficiaries’ sets.

2.3 Information Source of Reputation
Reputation, as part of collective systems of beliefs and opinions about people and things can be classified according to the origin of these beliefs and opinions. Beliefs can derive from several sources, such as direct experiences, received information, and social group prejudices. Those sources can be grouped in primary sources and secondary sources, according to whether the information is obtained by direct and indirect interaction [4].

Primary sources of information correspond to the agents that participate in social interactions, including those agents that take part in the action as well as those whose participation is limited to observe other agents’ actions. Secondary sources of information consist of the several opinions about agents that are not based on direct contact, such as information gathered indirectly, stereotypes and prejudices.

The terms used to denominate a reputation vary according to the author and the source of information. Bromley [4] designates primary reputation as one resulting from direct interactions or the observation of those interactions, and designates secondary reputation as one resulting from the opinions received from others. Conte and Paolucci [8] call image, or social image, the evaluation derived from primary sources of information, and they call reputation the evaluation derived from secondary sources of information. Yu and Singh [17] call local belief the evaluation resulting from direct interaction among the agents, and total belief one resulting from the combination of the local belief and the information received through the propagation of the local belief.

2.4 Reputation Evaluation and Measurement
A reputation can be summarized in a word or in an expression, such as "good" or "very good", or it can be detailed through examples and justifications. In general, a reputation can be evaluated through several types of measurements, different factors, and even different techniques of evaluation.

The measurement types that can be employed in the evaluation of a reputation allow different levels of detail, as in a level of esteem such as "bad", "good" and "great", or a more specific evaluation that considers the content of a reputation. That content evaluation allows identifying which are the aspects involved in a reputation as well as the relevance given to each one of them. The nature and the relevance of those aspects, or attributes, depend on the target entity to which they are related, and vary according to the interests and values of the members of a social group [4; 10]. Examples of attributes that can be used when we evaluate a person are: honesty, ability, and intelligence.

Besides the level of esteem and related attributes, we can measure a reputation through its extension in a specific social group. Examples of extension of an individual's reputation are: reputation known by most people; reputation known by some people; reputation known by few people [4, p. 250].

The factors involved in a reputation evaluation are those that somehow influence this evaluation. In a general way, a main factor is the behavior of the target entity [5; 17]. However, other factors can be considered, such as the reputation of the evaluator [18], the reputation of the propagator, as well as the social relationships between the target and the evaluator [13].

2.5 Reputation Maintenance
The reputation maintenance process deals with the modifications that occur in the content and structure of a reputation over time. Those alterations are consequence of changes that happened to agents as well as changes occurred in the environment in which they operate. This process reflects the consistency among the current behavior of the target agent and its expected behavior. Any discrepancy between current and expected behavior can provoke adjustments in reputation. For instance, if an agent achieves something socially important or suffers a significant failure, this information should be reconciled with its previous reputation [4, p. 216].

As the context is susceptible to alterations, an agent must be able to predict future changes in its environment in order to maintain consistent behavior. [9, p. 228].

Temporal aspects play an important role in the process of maintenance of a reputation. Some authors [5; 9; 10; 18] consider that the most recent behavior should weight more in a reputation maintenance process than the oldest ones, since a reputation should reflect the agent’s current behavior in spite of older ones. Those authors emphasize an interesting aspect related to the maintenance of a reputation and the control of its value. They consider that a reputation value cannot grow indefinitely; otherwise the most recent defections could not significantly affect a very high reputation.
2.6 Reputation Scope

The reputation scope distinguishes reputations according to the manner they are employed, in a local or a global way.

A reputation can be known as a global characteristic, shared by all members of a social group, or it can be known as a local property, maintained as an individual belief [11; 13]. We adopted the term global reputation to designate the first scope type and the term local reputation to designate the second one.

In a global scope, all members of the social group can contribute to form the reputation of a target entity, and this reputation is unique for the whole group. The target reputation is modified every time a group member makes a new evaluation of that target entity.

In a local scope, each member of the social group evaluates the target entity in an independent way. The reputation value is a characteristic associated to the relationship between the target entity and the evaluator entity.

2.7 Reputation Propagation

A reputation propagation process deals with various aspects involved in transmission of reputation. Information related to reputation is not just transmitted by the propagator entity, which, instead, selects, invents and manipulates the original information according to the restrictions and opportunities presented by the circumstances [4; 8; 13].

According to Conte and Paolucci [8], there are five aspects involved in an agent's decision on spreading or not a reputation: why to transmit, to whom to transmit, about whom to transmit, what to transmit, how to transmit. These aspects of reputation propagation can be associated to a primary reputation [17] as well as to a secondary reputation [8].

3. A PRELIMINARY FUNCTIONAL ONTOLOGY OF REPUTATION

This section outlines a work still in progress, a preliminary functional ontology of reputation. The goal of this ontology is to put together the broad knowledge about reputation produced in some areas of interest such as psychology and AI, and to represent that knowledge in a structured form.

The functional ontology of reputation employs the knowledge categories proposed by Valente [16] (see also [3]) in the work called the Functional Ontology of Law. We claim that the concepts of the legal world can be used to model the social world, through the extension of the concept of legal rule to social norm and the internalization of social control mechanisms in the agent's mind, so far externalized in legal institutions. While in the legal world a rule violation generates a legal punishment for the infringing agent, in the social world the penalty associated to an agent that violates a social norm is a bad reputation [6].

This ontology includes, as a kernel, the reputation concepts presented in section 2. That kernel is embedded in a broad framework, composed by the categories inspired by the Functional Ontology of Law. In that way we have outlined the whole mechanism related to reputation notion in the agent's mind.

As in the Functional Ontology of Law, the distinction among the categories of the reputation ontology will be accomplished according to a functional perspective, in which each component of the reputation system, embedded in the social system, exists to perform a specific function in the effort to achieve social objectives, such as trust, reciprocity and social cooperation.

The functional ontology of reputation contains four main categories: Reputative Knowledge, Responsibility Knowledge, Normative Knowledge and World Knowledge. We use class diagrams developed in Unified Model Language (UML) to represent the ontology taxonomy.

3.1 Reputative Knowledge

The Reputative Knowledge is the most characteristic category in the functional ontology of reputation. It was inspired by the Reactive Knowledge from the Functional Ontology of Law [16]. Its main function is to deal with the agent reward (good reputation) or the agent penalty (bad reputation), according to his behavior. This category models the products as well as the processes involved in the reputation notion, as described in section 2.

Reputative Knowledge main concepts are Reputation Property, Reputation Role and Reputation Process.

Figure 1: Reputative Knowledge representation

3.1.1 Reputation Property

The Reputation Property concept represents two reputation dimensions: reputation nature and reputation type.

Figure 2: Reputation Property representation

Reputation Nature concept distinguishes reputation according to the nature of a reputational entity: individual reputation, related to persons or individual agents; group reputation, related to the individuals or agent groups; product reputation, associated to products and services; location reputation, associated to geographical places; event reputation, associated to events in a general way; and activity reputation, related to the activities executed by agents.
The Reputation Nature concept distinguishes a reputation according to information source used in its formation. As we have seen in section 2.3, there is not a consensus among reputation researches about the terms used to denominate a reputation resulting from primary sources of information. We adopted the term proposed by Bromley [4], primary reputation, to designate this type of reputation. We called direct reputation the one derived from direct encounters among agents. We used observed reputation to refer to a reputation derived from the observation of those interactions.

We adopted the term secondary reputation to name the reputation resulting from the secondary sources of information. We adopted the term proposed by Mui et al [11], propagated reputation, to designate the one derived from received information. We called collective reputation the one agent inherit from their social group, and stereotyped reputation the one based on social prejudices.

The Reputation Roles concept represents those roles played by entities involved in reputative processes, such as reputation evaluation and reputation propagation. Putting together the target notion proposed by Conte and Paolucci [8] and the reputational entity notion proposed by Bromley [4], we defined the target role, played by all of the entities that act as reputation object, such as people (individuals or groups), things, products and service, places, events and activities.

Based on the evaluator, propagator and beneficiary sets of agents [8] we defined the evaluator role, the propagator role and the recipient role, respectively. The recipient role extends the notion of beneficiary set of agent in the sense that it makes explicit the action of receiving information about reputation related to the action of transmitting it. These roles can be played by persons or groups of people, as in daily life, or by software agents, as in online systems, such as e-commerce systems and electronic chat systems.

describes a reputation through the attributes that contribute in its composition as well as the relevance associated to them.

Figure 8: Reputation Measurement Type representation

The Reputation Evaluation Factor concept represents the four factors that somehow have influence a reputation evaluation. We adopted the following terms in order to identify these factors: Target Behavior, Evaluator Reputation, Propagator Reputation and Target Evaluator Social Relationship.

Figure 9: Reputation Evaluation Factor representation

Reputation Maintenance Process
The Reputation Maintenance concept represents the process that deal with reputation alterations over time. That concept involves two other processes, aggregation process and historical process.

We called aggregation process the one that deals with impact on reputation value caused by evaluation of agent’s current behavior: a positive impact improves reputation; a neutral impact maintains the same reputation; a negative impact decreases reputation. We called historical process that one in charge of historical dimension of reputation. That process specifies the historical extension of previous evaluations considered as well as the relevance associated to the most recent evaluations in spite of oldest ones.

Figure 10: Reputation Maintenance representation

Reputation Propagation Process
The Reputation Propagation process deals with reputation spread in social network. This process involves the following three concepts. Propagation Contents represents the information spread, in other words, what is transmitted. For example, an agent can decide to modify a reputation value before spreading it. Propagation Purpose designates reasons that an agent can consider before spreading a reputation. For example, an agent can decide to transmit a reputation because he agrees with its value. Propagation Form designates how a reputation is spread in social network. For example, during direct interaction among the agents, or in broadcast way.

Figure 11: Reputation Propagation representation

3.2 Responsibility Knowledge
The Responsibility Knowledge category was borrowed from the Functional Ontology of Law [3; 16]. Its main function is to associate a cause to a specific agent behavior, in order to define whether the agent must be considered responsible for this behavior or, instead, there are circumstances that attenuate its responsibility.

The Responsibility Knowledge represents four main notions, namely, attribution notion, actor-observer effect, responsibility assignment and responsibility restriction. Attribution is a term used in social cognition to refer to the causes for which the actions are interpreted in terms of personal characteristics (internal cause) or external circumstances (external cause), in order to provide a causal explanation for the associated behaviors.

Attribution processes are part of our inclination to impose a pattern of meaning to observed facts. Concerning reputation, one of major attribution effects is to give consistence and coherence to agent behaviors. A behavior has internal cause when interpreted as depending exclusively on the agent, for instance, agent’s motivation or agent's capacity. A behavior has external cause when considered as depending exclusively on external circumstances, such as opportunities or threats [4, p.37]. Attribution processes can be affected by the so-called actor-observer effect. That term is used to designate the tendency to attribute internal causes to other people’s behavior and to attribute external causes to one’s own behavior. Therefore, that effect can provoke a fundamental attribution mistake. Bromley points out that the actor-observer effect is most likely to affect our perception of people we know least well than of people we know better well [4, p. 38].

Responsibility assignment indicates that an agent should be held responsible for its behavior, given that it occurred under a condition interpreted as depending on internal causes. Responsibility restriction indicates that an agent cannot be held responsible for its behavior, given that it occurred under certain condition interpreted as depending exclusively on external cause instead of internal cause.
The following concepts represent the notions discussed above: Attribution, Internal Cause, External Cause, Actor-Observer Effect, Responsibility Assignment and Responsibility Restriction.

![Diagram](image)

**Figure 12: Responsibility Knowledge category**

### 3.3 Normative Knowledge and World Knowledge

Normative Knowledge category was borrowed from the Functional Ontology of Law [3; 16] and extended to deal with social norms instead of legal rules. Its main function is to classify the agent behavior, through the description of the social norms. The idea is to compare the agent actual behavior with the ideal one, prescribed by the norms, and then conclude whether a norm has been observed or violated. Normative Knowledge generates a normative status, adequate or inadequate, respectively, in order to classify the agent behavior.

World Knowledge category was borrowed from the Functional Ontology of Law [3; 16] and comprises the knowledge about the world to which the social norms are applied. Its main function is to provide a model of social behavior, which is used by the other categories.

These two categories will be discussed in a future work.

### 4. EXAMPLE OF THE USAGE OF THE FUNCTIONAL ONTOLOGY OF REPUTATION

A big picture of the relation among the four knowledge categories of the functional ontology of reputation is shown in figure 13 (inspired by Valente [16, p. 74]).

A cycle starts with the interpretation of the agent social behavior by the World Knowledge category. This category describes the agent behavior in terms of social norms. After this first step, Normative Knowledge category receives this behavior description and matches it with a social norm, generating the correspondent normative status (adequate or inadequate). Responsibility Knowledge category then defines whether the agent should be held responsible for its behavior (responsibility assignment or responsibility restriction). Finally, using the normative status and the responsibility information, Reputative Knowledge category is able to define the agent reputation, as a reward (good reputation) or a penalty (bad reputation).

As a scenario illustrating the usage of the functional ontology of reputation described above, imagine a society X, and agents A, B, C, belonging to this society. Agent A observes agent B behavior, who is smoking in a class room. Society X and agents’ behavior are represented by the inferior rectangle, while the superior rectangle represents the reputative system inside agent A’s mind.

![Diagram](image)

**Figure 13: Categories of a Functional Ontology of Reputation**

Through its World Knowledge category, agent A is able to describe agent B behavior in the following terms: “agent B is smoking in a closed place”.

Then Agent A, using its Normative Knowledge category, can look for a social norm related to that description, and indeed agent A finds the following one: “no smoking in closed place”. So, through this norm, agent A is able to classify agent B behavior as an inadequate one.

Thanks to its Responsibility Knowledge, which function is to associate an internal/external cause to a behavior, agent A concludes that agent B is smoking in a closed place of its own free will. As agent B behavior has an internal cause it provokes a responsibility assignment.

Since agent A has concluded “inadequate behavior with responsibility assignment” regarding agent B behavior, it can use its Reputative Knowledge in order to (i) verify reputation’s properties, (ii) identify reputation roles, and (iii) evaluate agent B reputation. First, agent A recognizes a primary reputation, seeing that it observes agent B behavior by itself. Second, it identifies that a target reputation role is played by agent B. Therefore, it concludes that it is dealing with an individual reputation.

Through a Reputation Evaluation Process, agent A is able to determine agent B reputation esteem level, as “bad guy”. However, until then agent B reputation in agent A’s mind was “very good guy”. Therefore, this new different evaluation, “agent B is a bad guy”, has to be aggregated to the current one. Then, agent A has to perform a Reputation Maintenance Process, which results in a new aggregated agent B reputation esteem level, such as “not so good guy”.

At the end of this process, agent A is able to provide agent B reputation for other agents in society X. For example, performing
5. DISCUSSION AND FURTHER WORK

In this paper we have presented some concepts used to identify reputation dimensions: reputation nature, roles involved in its formation and propagation, information sources for reputation, evaluation of reputation, reputation scope and reputation maintenance. Those concepts constitute a set of ontological terms that we have used to develop the kernel of a functional ontology of reputation. That ontology contains four main categories, distinguished by their function: Reputative Knowledge, Responsibility Knowledge, Normative Knowledge and World Knowledge. These categories have been borrowed from or inspired by the Functional Ontology of Law proposed by Valente [16].

The goal of this ontology is twofold. First, to put together, in a clear and coherent way, the broad knowledge about reputation disseminated in the literature. Second, to represent that knowledge in a common and structured form. We claim that the concepts of the legal world can be used to model the social world, through the extension of the concept of legal rule to social norm and the internalization of social control mechanisms in the agent's mind, so far externalized in legal institutions.

As we have seen in section 2, although there is a huge work on agent reputation, each reputation model defines its own basic terms. This ontology can be used as a meta-level ontological description of key concepts of reputation that could be mapped to different reputation models. One possible usage for this meta-model is to build special services on a middleware layer, enabling the interoperation of multi-agent systems that use different internal reputation models as criteria for social decision. A description logic language, Ontology Web Language – DL (OWL- DL) [2; 15], will be used to implement this Functional Ontology of Reputation.

6. ACKNOWLEDGMENTS

This work has been developed as part of the project “A proposal for a functional ontology of reputation for normative agents”, supported by FAPESP, Brazil, grant number 04/15055-6. Jaime Sichman was partially supported by CNPq, Brazil, grant number 304605/2004-2.

7. REFERENCES


