REPUTATION FORMALIZATION FOR AN INFORMATION SHARING MULTIAGENT SYSTEM

Jonathan Carter, Elijah Bitting and Ali A. Ghorbani
Faculty of Computer Science,
University of New Brunswick,
Fredericton, NB, E3B 5A3, Canada

We propose that through the formalization of concepts related to trust, a more accurate model of trust can be implemented. This paper presents a new model of trust that is based on the formalization of reputation. A multidisciplinary approach is taken to understanding the nature of trust and its relation to reputation. Through this approach, a practical definition of reputation is adopted from sociological contexts and a model of reputation is designed and presented.

Reputation is defined as role fulfillment. In order to formalize reputation, it is necessary to formalize the expectations placed upon an agent within a particular Multiagent System (MAS). In this case, the agents are part of an information sharing society. Five roles are defined along with the ways in which these roles are objectively fulfilled. Through the measurement of role fulfillment, a vector representing reputation can be developed. This vector embodies the magnitude of the reputation and describes the patterns of behavior associated with the direction of the vector.

Experiments are conducted to verify the sensibility of the proposed models for role fulfillment and overall reputation. The simulation results shows that the roles, defined for building reputation in an information-sharing MAS environment, react to different agent and user actions in a manner consistent with the formal definitions.

Key words: Agent, Information Sharing, Multiagent Systems, Reputation, Trust.

1. INTRODUCTION

Currently, many researchers are focusing their efforts on formalizing trust within multiagent systems based on an e-commerce framework. A formalization of trust will lead to models that accurately reflect the same characteristics and behaviors of trust that humans exhibit within human societies. Through this reflection, a foundation of trust between human users of such multiagent systems and their agent counterparts will be established. The lack of this establishment will naturally lead to an undermining of the goals of the multiagent system. Users will choose not to use a multiagent system that they do not have trust in. Hence, a dysfunctional multiagent system is established. Clearly, the trust formalization is needed to establish the necessary trust and confidence of the user in the multiagent system.

Trust cannot be properly formalized without a full conceptual understanding of its nature. In order to begin this journey, it is important to understand that trust is multidimensional and relies on many related concepts. Hence, examination of concepts related to trust is necessary.

In this paper, the concept of reputation is examined as it is closely related to the concept of trust (Blair, 2001). In general, there is a positive correlation between reputation and trust. If the concept of reputation is related to the concept of trust, the formalization of reputation should simplify further efforts to formalize trust. As such, the formalization of reputation is an excellent place to begin in formalizing trust.

Through the formalization, reputation can be quantified and used by agents in ways they see fit. Linking occurs between the concrete reputation model and the abstract trust to allow agents to make judgements about the trustworthiness of another agent in an objective way. It is important to remember that trustworthiness should not be one-dimensional as it is here. Formalization of other related concepts to trust will be required to fully formalize trust.

Defamation is a natural consequence of using the reputation construct. Defamation is defined as the act of destroying an agent's reputation through the dissemination of false

information about said agent through a third party (Duhaime, 2001). Defamation results in the loss of social power and status of the damaged party. As such, successful societies must maintain laws that protect the individual from defamation.

In order to prove that defamation has occurred, the damaged party must prove that the disseminated information is both false and has been filtered through a third party. In the case of an *information sharing society* (see Section 3), defamation could occur by falsely reporting the invalidity of an information agent. This leads to an unfair decrease in the victim's reputation, hence a loss of social power. This issue will be addressed in this paper.

This paper is organized as follows. Section 2 explores the relationship between trust and social order and describes the nature of reputation from a sociological and multiagent perspective. Sections 3 and 4 present the formal roles that result from the definition of reputation chosen. This discussion includes the discussion of concepts along with chosen models. Section 5 discusses the extensions of the formalization of reputation to trust and social values. In Section 6, the simulation results are given. Finally, conclusions of the present study are summarized.

2. TRUST, ORDER AND REPUTATION

Most generally, trust can be considered the expectation held by each member of a society that the existing natural and moral social orders will persist. This is to say that members of society have an inherent trust or faith that 'the sun will rise tomorrow' and society will exist more or less as it did today (Barber, 1959). In the following sections, trust in both natural and social orders will be examined.

2.1. Natural Order

Natural order relies upon the persistence of the actual environment. For example: every university student, upon arriving at school each morning, has an implicit expectation that the university buildings will still be there, standing as they had been the day before. In the case of a virtual environment such as ACORN (Agent-based Community Oriented Routing Network), this would be the expectation that the ACORN system and its resources will continue to be available for use. ACORN is a multi-agent architecture designed for information distribution and retrieval within an established network (See (Marsh, 1997, 2001) for more details)

2.2. Social Order

People trust in others, this is a basic fact; if this were not the case, people would simply not engage in any interaction with others for fear of harm or deceit (Luhmann, 1979). This trust results from expectation of the persistence of the moral social order, and is extremely important. People's expectation of the persistence of society (with it's existing moral and social order) allows them to consider consequences of their actions, positive or negative. The idea is that people's expectation that society will persist relates directly to their ability to consider the consequences of their actions to be faced in the future (note: consequences can only be realized if society continues to exist). This idea embodies the ancient law of cause and effect, which states that what people experience in their lives (effects), are direct results of past causes (cause). This idea of cause and effect (or karma), is necessary for continued moral actions (Barber, 1959). It is accepted almost universally by sociologists that in order for an individual to act rationally and deliberately, to effectively plan and make decisions, they must hold in trust many features of the social order. Garfinkel conducted many clever and

successful experiments that involved breaches of accepted social practices (Garfinkel, 1967). They showed how disturbed people tend to become when such a breach is encountered.

In the scope of the model presented in this paper, this idea of expectation of continued natural and moral social orders could be embodied as follows. Trust in the persistence of the natural environment can be assumed for simplicity sake. Trust in the persistence of the moral social order could be maintained by enforcement by the system in the form of a judicial entity who enforces the moral social order and punishes anyone who is guilty of breaching that order. The model could incorporate a judicial entity which would make use of a function $\mathcal{P}(u)$, where u is a user and $\mathcal{P}(u)$ returns the current penalty value for that user. The penalty values would be 0 for anyone not found guilty of an offence (innocent until proven guilty), otherwise the value returned would be the sum total of any outstanding (not yet expired) penalties. Penalties are assigned by the judicial entity. Their severity in terms of weight and duration would have to depend on the severity of the offence and the past record of the offender. These penalty values would be subtracted from user's trust values, effectively punishing those who breach the social order.

2.3. Order and Responsibility

Those in positions of power are expected to use that power in a just manner. This type of expectation results from the need to give power to those who have complex or specialized knowledge, which is out of reach to some members of the society. Upon receiving this additional power, the individuals wielding it are expected to act accordingly; they are entrusted with fiduciary responsibility. A relationship is established between the trustee and their society; this relationship mandates that the wielder of power use that power in a just manner.

Human society is always pushing to instill fiduciary responsibility in those in positions of power such as: parents, government officials, leaders of industry and teachers (Barber, 1959). The catch to this idea is that fiduciary responsibility depends heavily on the distribution of power within a society. In the case of equal distribution of power, each member of society has an equal amount of fiduciary responsibility toward all the others in the society. The idea of fiduciary responsibility only applies to situations where some are more capable than others in a society. In this situation, fiduciary responsibility falls on those who are more capable to not take advantage of the less capable (Barber, 1959).

In the scope of the model presented here, all the members of society have equal power or capability. The idea of fiduciary responsibility need not be implemented for the time being.

2.4. Sociological Considerations of Reputation

In order to understand the concept of reputation, it is useful to examine the definition provided by another well-known multiagent disciplinary study commonly referred to as Sociology. Sociologists have studied the concept of reputation as it relates to human societies.

Within Sociology, the concept of reputation lies within the realm of Sociology of Identity. In turn, Sociology of Identity is connected to the outside world through Symbolic Interactionism theory. This provides the logical connection between multiagent studies within the Artificial Intelligence (AI) community and the studies of identity and reputation within Sociology.

Sociologists see reputation as held by individuals, organizations, and objects. It is ascribed by society towards the individual and cannot be generated by the individual being judged. Reputation is viewed as both personal and commercial.

Personal reputation is the general estimation held by the public about the individual

in question. This estimation is based on the way in which identity is managed by the individual and presented towards society. Through presentation of the self, society constructs a reputation by looking at the objective behavior of the individual (Press, 1959). Commercial reputation is very similar to personal reputation in that it focusses on the estimation held by the public about a commercial entity. However, companies utilize public relation components to influence social perception. Hence, companies have more power than the individual in establishing and marketing their own reputation. This helps to explain the unequal distribution of social power between companies and individuals.

Erving Goffman (Press, 1959) makes the useful analogy of identity management to the dramatic world of theater. Individuals become actors on a stage negotiating their roles to society. On the front stage, individuals are polite with decorum. The back stage is where the individual exists as themselves. Front stage and back stage attain a certain fluidity and reputation is created through the management of both stages.

The audience has certain expectations or roles for the actors. Reputation is constructed based on the audience's belief that the actors have fully satisfied their roles. If an actor cannot act, they have failed as the role of actor and will develop a negative reputation amongst the audience members. In the same way, individuals and businesses have certain roles within society. If society judges that they have met their roles, they are rewarded with a positive reputation.

Positive reputation leads to confidence/trust in the individual along with a higher level of social status and power. Such individuals become sought out within society. Negative reputation leads to a loss of esteem held in society along with social status and power. Naturally, it is in an individual's best interest to maximize their positive reputation through identity management.

Within human societies, both reputations are so important that rules are established to protect reputation (Global Campaign, 2002). Defamation laws prevent an individual from destroying another individual's reputation through false accusations with third-parties (Duhaime, 2001). This suggest that within a multiagent system, the issue of defamation must be addressed in the formalization of reputation.

2.5. Reputation Within Distinct Societies

When this sociological examination of reputation is related to multiagent systems, we view these systems as societies. Throughout this paper, the terms society and MAS are used interchangeably. It becomes apparent that the reputation of an agent is based on the satisfaction of roles ascribed to the agent by the multiagent system. The roles of an agent must then be defined along with ways of objectively measuring them from a social perspective. Clearly, the defined roles of an agent must lead to the satisfaction of the society's goals. Otherwise, social status and power will be ascribed to agents that may not serve the society's best interests.

Using this definition of reputation as a base for formalization, each type of society will have it's own set of roles for it's agents because each type of society will have a different set of goals to achieve. The roles of an agent in an information sharing society will be entirely different than the roles of an agent in an e-commerce society because both societies are achieving different goals. Due to the dependency of reputation on role definition and measurement, the formalization of reputation is impossible to universalize as reputation is dependent on the goals of the multiagent society. Within this paper, formalization of reputation is restricted to that of an information sharing society.

An agent that makes transitions between societies will have to reestablish its reputation because the reputation value ascribed by the original society will be meaningless within the

different context of the new one. Each society has its own set of roles for satisfaction of differing goals. As such, the reputation ascribed as a result of these roles only makes sense in the context of that particular society where the roles were defined.

3. SOCIAL ROLES AND EXPECTATIONS

This section discusses the formal definitions of the agents' roles within an information sharing society along with proposed models describing the measurement of satisfaction of those roles over time. In order to define the roles, the purpose of the society and the goals of the agents must be outlined.

An information sharing society is a society of agents that attempt to exchange relevant information with each other in the hopes of satisfying a user's request. Each user can generate query or information agents. A query agent is an agent that traverses a network in search of finding 'useful' information for the agent's owner. Useful information has two definitions: (a) information that directly pertains to an immediate query issued by the user and (b) information that is relevant to the user's personal interests. An information agent is one that represents information relevant to the owner's personal interests and can freely navigate networks without any primary goals. Both types of agents contain the social network of known users connected to the agent's owner along with respective interests of those users.

These agents interact with each other in two ways. First, interaction occurs between a query agent and an information agent to relay information back to the query agent's owner. Secondly, interaction occurs to learn of a social network of people with related interests. This is provided so that agents can make recommendations to other agents about whom to see in the event that they cannot satisfy an information request. Ultimately, agents use each other to navigate towards the correct information agents relevant to their goals.

Through a brief examination of this society, it becomes clear that roles do exist for different agents. Together, the measurement of satisfaction of these distinct roles will objectively allow for a measurement of reputation and indirectly trust. Within an information sharing society, we define reputation, $\mathcal{R}$, as a 5-tuple

$$\mathcal{R} = (\Gamma, \Omega, \Upsilon, \Theta, \Psi)$$

where each tuple represents a role.

Definition 1. Social Information Provider (Γ): Users of the society should regularly contribute new knowledge of their friends to the society. The success of multiagent system is directly dependent on accurate data of the human society.

Definition 2. Interactivity Role (Ω): Users are expected to regularly use the system and maintain some form of interactivity. This helps to keep the society up-to-date.

Definition 3. Content Provider (Υ): Users should provide the society with knowledge objects that reflect their own areas of expertise. It is proposed that such objects will be of higher quality than objects that don't reflect the personal interests of their user.

Definition 4. Administrative Feedback Role (Θ): Users should provide feedback for the functionality of various aspects of the system. These functionalities include the quality of information objects.

Definition 5. Longevity Role (Ψ): Users should maintain an average reputation that is positive within society. Longevity of the society is achieved through longevity of the individual.

In Section 4, a model is presented that combines the measurement of these roles as a means of expressing reputation. Each role becomes normalized and weighted according to its importance. The process of weighting these roles is not a trivial task and will be elaborated upon in future work. Subsequent work will also analyze and compare the interdependency of these roles on each other.

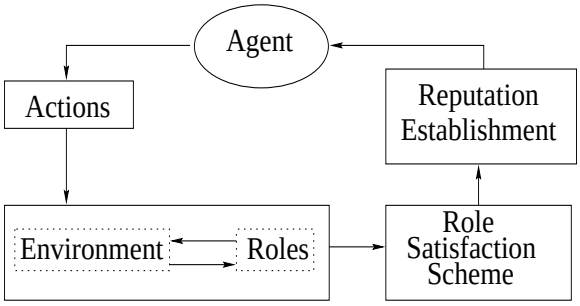

FIGURE 1. Reputation model framework

From a high-level perspective, the reputation model illustrated in Figure 1 shows how the reputation of any given agent in an information sharing MAS is established by determining the extent to which the agent fulfils the roles defined for that MAS. Roles are fulfilled by the performance of expected actions. Once the role fulfillment values are calculated (see Section 3 for formulae relating to the five roles defined in this paper), they are aggregated based on a weighting scheme such as the one described in Section 4. This weighted aggregate forms the reputation of the agent.

Notice that the roles themselves are dependant on both the environment and the actions of all agents within that environment. The definition of an environment can include the social goals. Due to the dependency of role definition on goals, the environment is said to have an influence on the roles. At the same time, the roles have an indirect influence on the environment through selective reinforcement of agent behaviors. These agent behaviors have direct impact on the environment. Hence, environment influences roles as roles influence the environment.

3.1. Social Information Provider

Within an information sharing society, knowledge of members of society forms the lifeblood of the functional society. Without the contribution of knowledge of the members, agents would never be able to learn and exchange 'useful' contact information. In order to facilitate knowledge exchange, it is critical that users provide honest and accurate descriptions of the people they know and the areas of their expertise. Users of the society should regularly contribute new knowledge of their friends to the society in order to maximize the likelihood of helping a fellow user in search of contact information. When users regularly fulfill this obligation, they are said to satisfy the social information provider role.

In the real world, this role is exemplified through the degree of connectivity of people to their communities. Some people have an unusually high degree of connection to those around them. Others use these people as "go-to" people for anything unfamiliar. They help facilitate contact between like-minded individuals.

As an example of the applicability of this role, consider the case of all agents within a small subset of the society containing no information about other users. In such a subset, query agents could only seek information from the immediately accessible information agents. If these agents could not satisfy the query, the query agent would have no means of contacting other agents outside the immediate environment and the query would go unfulfilled. Without knowledge of a social network, the society loses significant functionality and trust by the user.

Each recommendation, γ, made by a user represents a social connection and has a weight associated with it. The weight is a function of user's reputation and may decay over time. In other words, weight indicates the strength of a recommendation over time. Let $R_u(t)$ represent the reputation of the user u (recommendee) at time t. Let $\gamma_u(t) = \{\gamma_u^i(t) \mid i = 0..n-1\}$ represent a set of *distinct* recommendations stored by user u at time t. Note that $\gamma_u^i(t)$ denotes a distinct recommendation i stored by user u at time t. The weight of recommendation $\gamma_u^i(t)$ stored by user u at time t, $w_u^i(t)$, is defined as,

$$w_u^i(t) = e^{-\alpha(t-t_0)} R_u(t) \qquad 0 < \alpha < 1, \tag{1}$$

where α is a decay rate parameter.

Each particular recommendation made by a user u at time t_0 about a friend has a particular weight, $w_u^i(t)$, associated with it at the present moment t in measuring the satisfaction of this role. The weight is the product of a time decay factor and the recommendee's reputation. If the recommendee's reputation remains constant, the weight of the recommendation naturally decays because an old recommendation is no longer contributing towards the satisfaction of the social information provider role. Otherwise, it is possible that the weight could increase in time as the recommendee's reputation increases against the decay factor. Eventually, this contribution will decay to 0 even if the reputation reaches its maximum.

It is not allowed for negative reputations of recommendees to negatively impact the overall role satisfaction of the social information provider. Instead, the reputation is set to a minimal value for this calculation. In all cases, the recency factor α determines the rate of decay of the effects of this weight over time. Let n and W_u represent the number of recommendations made by a user and their collective weights, respectively. The total weight is given as,

$$W_u = \sum_{i=0}^{n} w_u^i(t) \tag{2}$$

It is possible to evaluate the degree to which the social information provider role has been satisfied by mapping the net value of W_u to a value in the interval $[0, 1]$.

Suppose ϕ is a real-valued function $\phi(.)$: $R \to [0, 1]$ mapping a value in R to the bounded interval $[0, 1]$ and satisfies the following

$$\lim_{x \to +\infty} \phi(x) = 1 \qquad \lim_{x \to -\infty} \phi(x) = 0$$

Assuming that Γ grows exponentially until an upper limit inherent in the system is approached, at which point the growth rate slows and eventually saturates. The sigmoid function is a suitable choice for ϕ to restrain the overall satisfaction value between $[0, 1]$. A model for the satisfaction of the role Γ is defined as,

$$\Gamma = \phi(W_u) = \frac{1}{1 + e^{\beta W_u}} \tag{3}$$

The exponential growth rate can be slowed down by controlling the 'sharpness' parameter β.

3.2. Interactivity Role

One of the primary reasons for formalizing trust is to lead to user confidence in the multiagent system. Such confidence will lead to regular use of the society by the user. Without such participation, the society becomes useless. The primary goal of an information sharing society is to provide information at the user's request. As such, it is expected that users should regularly use the system to satisfy the interactivity role. In order to model this role, it is important to define the concept of interactivity in quantitative terms. This role focusses on the quantity of interactions. As explained later, the administrative feedback role is responsible for ascribing a sense of quality to the interactions between users.

Interactivity is defined as a computational task that is demanded by the user. Tasks can include the following operations: logging in, checking e-mail, creating information agents, creating query agents, receiving updates of released agents, modifying profiles, and many more. Each task is equally weighted as one computational operation. Let op_u^d represent the number of operations performed by user u on day d. The total number of user operations, Top, performed by user u in the past D days is given as,

$$Top_u(D) = \sum_{d=1}^{D} op_u^d \tag{4}$$

The total number of operations, TOP, is a summation of operations conducted by all users, U, over the past D days.

$$TOP(U, D) = \sum_{u=1}^{U} Top_u(D) = \sum_{u=1}^{U} \sum_{d=1}^{D} op_u^d \tag{5}$$

In order to calculate the role participation, Ω, a ratio is defined between the total number of operations conducted by the user and the total number of operations in the system.

$$\Omega = \frac{Top_u(D)}{TOP(U, D)} \tag{6}$$

For this role, satisfaction is measured as a proportion of participation by the user in relation to everyone else using the society. As the user participates more often, the model dictates that the reward will approach a maximum value of 1. This follows under the assumption that the average number of user operations is the same. As the user leaves the system, the total contribution of operational units by the user towards the pool of units will approach 0 along with the reward.

There is the risk that individual users could flood the system with operations in order to artificially inflate their degree of satisfaction of this role. It is proposed that each interaction type takes on a weight. This weight is ascribed by proportion to the total number of operations that have occurred. Over time, these weights will stabilize as the patterns of use by users become regular. The operations which become most weighted will correspond to operations that promote use of the system. When calculating the total number of operations, TOP, this weight can be applied to each operation conducted over the considered days. We anticipate that 'useful' operations have a higher volume than other operations. Currently, we are working on the development of an adaptive scheme that can be used to prevent the misuse of a system.

Individual users cannot manipulate these weights of operations. As such, the act of flooding a system with useless interactions will result in a minimal increase in the degree of role satisfaction. It is only possible to manipulate these weights if everyone agrees to flood the system with trivial operations. This is a highly unlikely scenario.

3.3. Content Provider Role

Within an information sharing society, the quality of information represented through information agents is important in achieving the social goal of satisfying the users' demands for specialized knowledge. Users will lose confidence in a system that retrieves knowledge that is of low quality. Although the terms being used to describe the goal of high quality information are inherently vague, it is proposed that quality of information can be inferred by the source of the information. It is believed that users that create information agents related to their areas of expertise will produce higher quality content related to their interests than those who don't. As such, it can be inferred that an information agent is higher in quality if its subject corresponds to the user's interests.

With this belief in mind, a user can satisfy the content provider role by producing information agents that are directly related to their areas of interests. Within the real world, this role is exemplified through an expert. When people are searching for information on a topic, they want to find other people that are as knowledgeable as possible in this topic. Otherwise, risks can become unacceptably high for the individual.

Let $Uk(u)$ represent the set of all keys of personal interests expressed by the user u, N represent the total number of distinct keys belonging to user u, $Ak_u(a)$ defines the set of keys belonging to agent a that have been created by user u and M is the total number of distinct keys belonging to agent a.

$$Uk(u) = \{\, key_u^{(i)} \mid i = 1 \cdots N \,\}$$
$$Ak_u(a) = \{\, key_a^{(i)} \mid i = 1 \cdots M \,\}$$

Let $AK(u)$ denote the union of all the agent's keywords owned by user u.

$$AK(u) = \{\, Ak_u(1) \cup Ak_u(2) \cup Ak_u(3) \cdots \cup Ak_u(P) \,\},$$

where P represents the number of agents created by user u.

$$Mk(u, a) = \{\, Uk(u) \cap Ak_u(a) \,\} \tag{7}$$

Matching keywords, Mk, between the agents' keywords and the user's keywords is defined as the intersection of $Ak_u(a)$ and $Uk(u)$. Let m represent the number of matching keywords between user u and its agent a (i.e., $m = |Mk(u, a)|$) and $\Delta(u, a)$ represent the Hamming distance of user u and its agent a (i.e., $\Delta(u, a)$ denotes the number of keywords user u and agent a differ). We are interested in measuring the degree of similarity between user u and its agent a.

Definition 6. The degree of similarity between a user u and its agent a is generally defined as the number of keywords they have in common divided by the total number of keywords.

We define the similarity between user u and its agent a as,

$$\Lambda(u, a) = \frac{2m}{N + M}. \tag{8}$$

Alternatively, we can measure the similarity between user u and its agent a as follows:

$$\Lambda(u, a) = \frac{Uk(u) \cdot Ak_u(a)}{\mid Uk(u) \mid \cdot \mid Ak_u(a) \mid} = \frac{Uk(u) \cdot Ak_u(a)}{N \cdot M} \tag{9}$$

The quality measure of an agent, $Q(u, a)$, is a function of both Λ and Δ. Note that higher Λ implies higher similarity, whereas greater Δ shows less similarity. We shall define this function by

$$Q(u, a) = \frac{\Lambda(u, a)}{\Delta(u, a)} \qquad \Delta(u, a) \geq 1 \tag{10}$$

A higher value of this function implies greater quality. A user's quality is the function of the qualities of it's agents. We define the quality of user u as the average of the qualities of its agents.

$$Q(u) = \sum_{a=1}^{P} Q(u, a)$$

$$\overline{Q}(u) = \frac{1}{P} \sum_{a=1}^{P} Q(u, a) \tag{11}$$

Recall that P represents the number of agents created by user u. The degree of role satisfaction, Υ, is a real value between 0 and 1 and is calculated using the algebraic sigmoid function as follows,

$$\Upsilon = \frac{Q(u)}{\sqrt{1 + (Q(u))^2}} \tag{12}$$

When a user creates an agent, that agent always has implicit within it the user's sense of who he/she is. In a word, agent implies a view of the user with regard to information sharing. Through this final calculation, the degree to which a user is sharing knowledge related to their area of expertise is obtained. In order to make an accurate calculation of this role satisfaction, the user must accurately report their interests. Otherwise, users will be penalized for providing incomplete information through the reduction in role satisfaction.

3.4. Administrative Feedback Role

Ideally, an information sharing society would be free of problems and users would be satisfied with all aspects of the system. Such a dream is just that. Users should be encouraged to provide feedback on the peripheral qualities of the system. In this case, these qualities include: ease-of-use, speed, stability, and quality of information. Users are said to satisfy the administrative feedback role by providing such information.

Users should have an outlet to express content or discontent with another user's information objects or recommendations. Such expressions will lead to an increase or decrease in the other user's reputation while not affecting the reporting user's reputation provided that defamation is not true. If defamation is proven true, then the reporting agent's reputation is reduced while the victim's reputation is restored. Within this realm, users are said to satisfy the *content validity role.*

There is an inherent risk of damaging a user's reputation through the allowance of subjective expression of content validity by other users. Similarly, it might quite possibly create strong incentives to high reputation through disproportionately positive feedbacks. This introduces a subjective aspect of reputation based on the honesty of the other members of the society. Unfortunately, this risk cannot be controlled through any other means than through systems related to defamation. Clearly, it goes well beyond the scope of this work.

In reality, the content validity role is a subset of the administrative feedback role. As such, both can be modelled using the same equations.

$$\lambda_u(t) = \kappa(e^{-\alpha(t-t_0)})$$

The above equation expresses the value of a single feedback about user u at the current time t in relation to when the feedback was generated at time t_0. Each single feedback $\lambda_u(t)$ must be a function of time to allow both positive and negative feedback to decay over time. The expansion factor, α, is the same as defined in the social information provider rule in Section 3.1

The evaluation factor, κ, falls between the value of $[-1, 1]$ to allow subjective evaluation of content by a user towards another user's information agent. For other types of administrative feedback, the evaluation factor is ignored by setting its value to the default of 1. This factor unifies both the general administrative feedback and content validity roles into one model.

$$\lambda_u = \sum_{d=1}^{D} \sum_{i=0}^{F_d - 1} \lambda_u(d, i) \tag{13}$$

Together, a summation of individual feedbacks about a particular user u over the past D days (with F_d feedbacks each day) forms the basis of degree of satisfaction of the administrative feedback role. As users become more satisfied with a particular user's information agents, the degree of satisfaction of this role increases. Conversely, low quality information will receive a poor rating and the degree of satisfaction of this role will naturally decrease. Once again, we define the role satisfaction, Θ, as

$$\Theta = \frac{\lambda_u}{\sqrt{1 + (\lambda_u)^2}}. \tag{14}$$

It should be noted that because the users are given the chance to express a subjective measure of satisfaction of other users, the possibilities of positive and negative discrimination arise. To handle the negative discrimination and defamation, users and agents must be made anonymous inside the system. Once the possibility of negative discrimination and defamation have been removed, the idea of selective filtering to filter out any unusually high ratings can be applied to handle positive discrimination. The ideas of controlled anonymity and selective filtering used for this purpose are laid out in (Dellarocas, 2000).

3.5. Longevity Role

Lastly, the social longevity of the system is a direct result of the users' longevity within the system. Longevity is defined as the total time that a particular entity has a positive average reputation. Users should be encouraged to maintain a high reputation to promote the longevity of the system. Such users satisfy the longevity role.

In order to measure the degree of satisfaction of this role, a model must take into account the average reputation, $\overline{\mathcal{R}}_u$, of the user u up to the present time t along with the actual age of the user within the system (i.e., $\mathcal{T}_u = t - t_0 + 1$). We define the longevity role, ψ, as

$$\begin{aligned} \psi(t) &= \eta \overline{\mathcal{R}}_u \\ &= \left(\tfrac{-1}{\alpha \mathcal{T}_u} + 1 \right) \overline{\mathcal{R}}_u \end{aligned} \tag{15}$$

The coefficient η accounts for the age of the user within the system. As the user ages within the system, the average reputation in the long run (over the past D days) gains more significance as the user's history becomes more established. This is expressed through the approach of the multiplier to the value of 1. This rate of approach is dependent on the longevity establishment factor, α.

In the event that the average reputation is less than zero, the degree of satisfaction of this role is set to 0 as we believe it is unfair to further punish a user for having a negative reputation. Once again, the value of the degree of satisfaction can be mapped to a value between $[0, 1]$ through the use of an algebraic sigmoid function as,

$$\Psi(t) = \frac{\psi(t)}{\sqrt{1 + (\psi(t))^2}} \tag{16}$$

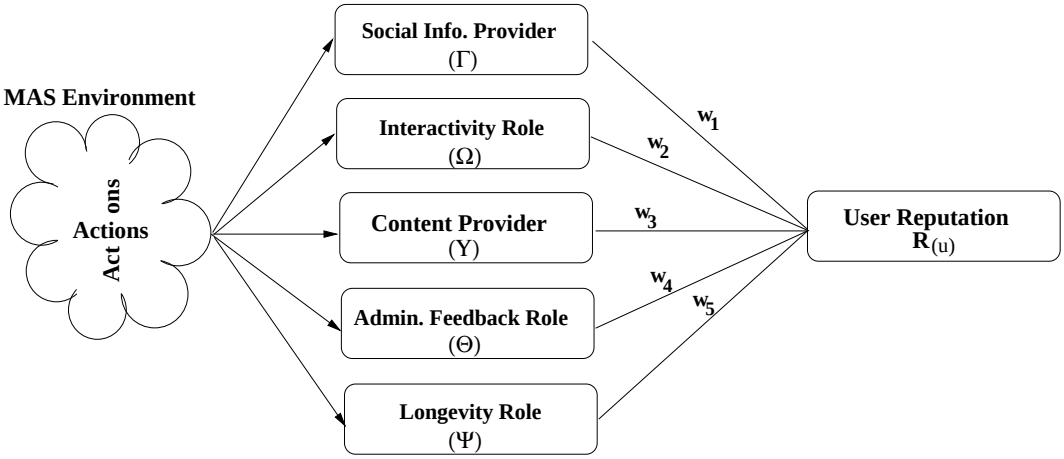

FIGURE 2. Determining user's reputation through the integration of roles.

4. INTEGRATION OF ROLES THROUGH REPUTATION

In the previous section, the expected roles of the users were formally expressed and models were established to express the degree of satisfaction of each of these roles. We should note that there are other ways to define, measure and quantify such roles, some of which could be probabilistic in nature. One should realize that the definitions and measurement schemes presented in this paper are the first attempt to define reputation as the degree of role fulfillment in the context of information sharing multiagent systems.

In order to determine a user's overall reputation, one must examine the degree of fulfilment of each of these roles and combine them to produce a scalar value (see Figure 2). Let

$$\mathbf{R}_u \;=\; \{r_1 = \Gamma,\; r_2 = \Omega,\; r_3 = \Upsilon,\; r_4 = \Theta,\; r_5 = \Psi\}$$

and

$$\mathbf{W}_u = \{w_i \mid i = 1..5\}$$

represent the corresponding weight vector. We define the user's reputation, $\mathcal{R}_u$, as,

$$\mathcal{R}_{(u)} = \sum_{i=1}^{5} w_{u,i} r_{u,i} = \mathbf{R}_u \cdot \mathbf{W}_u \tag{17}$$

The above role vector, $\mathbf{R}$, expresses the degree to which each of the previously mentioned roles has been satisfied independently of the others. The normalized weight vector, $\mathbf{W}$, expresses the degree to which each role's satisfiability must be weighted in calculating a final reputation. The distribution of weights amongst the roles is a reflection of social values within the society. A society may place more emphasis on a user being a good content provider than a user maintaining long term interactivity. Once again, these social values are entirely dependent on the goals of the society. As such, it is worth reemphasizing that it is impossible to universalize the calculation of reputation as the weighting function is dependent on social values.

Together, the dot product of these vectors will result in a scalar value that expresses the magnitude of the reputation. This overall reputation can be mapped to $[-1, 1]$ using a logistic function.

It is very important to capture the semantics of the roles and determine the significance of each role in relation with other roles. The weights can be determined in a number of ways. We can calculate the weight of a role by computing its statistical significance, $\mathcal{S}$, as follows,

$$\mathcal{S}_{r_i} = \frac{r_i - mean(\mathbf{R})}{\sigma^2(\mathbf{R})}$$

The weight distribution vector becomes the normalized representation of this new statistical significance vector. The roles which deviate most significantly from the measured mean become the biggest contributors to the calculation of reputation.

For the purposes of the experiments conducted in Section 6, the weight distribution function weighs all roles equally rather than being based on a significance distribution.

5. SOCIAL CONCEPT LINKS

5.1. Trust

In order to link reputation to trust, it is important to examine the implications of the proposed reputation model. In this model, the reputation vector $\mathbf{R}$ has both magnitude and direction.

The concept of reputation magnitude is quite clear. A person with a high reputation will have a large magnitude greater than 0 while a person with a negative reputation will have a large negative magnitude. In both cases, the person has had to perform many actions to acquire a large magnitude.

The implications of reputation direction are not as obvious. The direction primarily implies patterns of behavior on the user's part. Indirectly, the direction implies some subset of beliefs/philosophy of the user. The degree of similarity of two users' behavior is expressed through the Euclidean distance between their reputation vectors. This degree of similarity indirectly implies how different the two users really are with respect to each other. Trust is established amongst users with a positive reputation by examining how similar their values and beliefs are to one another. Users will trust other users that share similar beliefs and values to themselves (if both have positive reputations). The angle between two reputation vectors acts as an objective measure of similarity of actions, which are based on beliefs and values.

Desirability of one reputable agent to interact with another reputable agent is based on closeness or distance of the users that own the two distinct agents. The closeness, $\mathcal{C}$, of users u_1 and u_2 is defined as the square Euclidean distance between their reputation vectors.

$$\mathcal{C}_{u_1,u_2} = \|\mathbf{R}_{u_1} - \mathbf{R}_{u_2}\|^2,$$

where $\|\mathbf{x}\|$ is the L_2 norm (Euclidean norm) of the vector $\mathbf{x}$. An agent will trust another agent that has the highest reputation with closest patterns of behavior to itself amongst the agents available. Hence, a link between trust and reputation is established.

5.2. Culture/Values

Social values are expressed through the weight distribution function, w, when calculating the reputation of the user in question. Weights can be a function of many different things.

Commonly, weights could be a function of time. In this case, the calculation of reputation of a user becomes a function of the user's age within the system. This leads to the social stratification based on age. Certain segments of society will gain or lose reputation depending on their age bracket. Through this loss or gain comes a redistribution of power and status. Indirectly, this demonstrates the power of the weight functions in segmenting society.

6. EXPERIMENTS

In order to gain a better understanding of the nature of individual roles in a mathematical sense we have carried out the following experiments. Through these experiments we have also examined the effect of roles defined for building reputation. The last experiment is intended not only to show how the longevity role contributes to the reputation value, but also to illustrate the effect of aggregation of all the values obtained from the individual roles.

6.1. Social Information Provider Role

Simulations are carried out to demonstrate how this role builds reputation. The situation being simulated is: 100 recommendations are made by a single user u on each of 100 days. The reputations corresponding to each of the 100 recommendations are random floating point numbers in the range $(0, 1)$. The value of β in Equation 3 is set at 0.8 for all cases. Figure 3 shows that making recommendations about others with positive reputations builds reputation. It also shows that reputation decays over time in the absence of such recommendations (see the reputation curve after day 100).

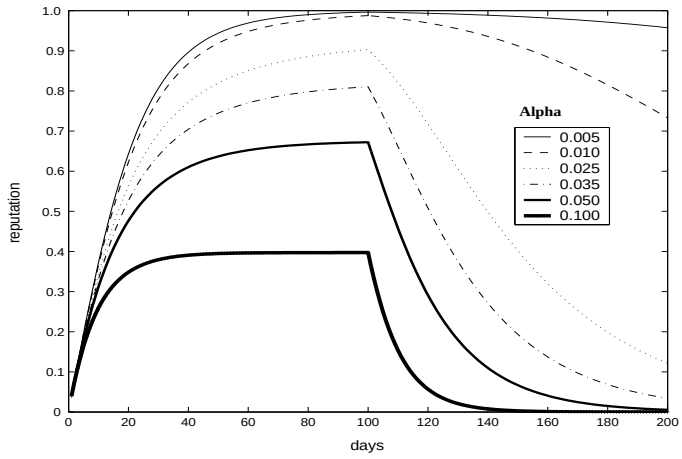

FIGURE 3. Continuous recommendations until day 100

6.2. Interactivity Role

In order to demonstrate how this role contributes in building reputation in different situations, two different cases are considered. The first case is when a user y performs many

operations on day 0, and then remains inactive as user x performs operations for each of the next 100 days. The results of this test are given in Figure 4. It is seen that the reputation from this role is directly proportional to the proportion of operations performed by a specific user with respect to the system.

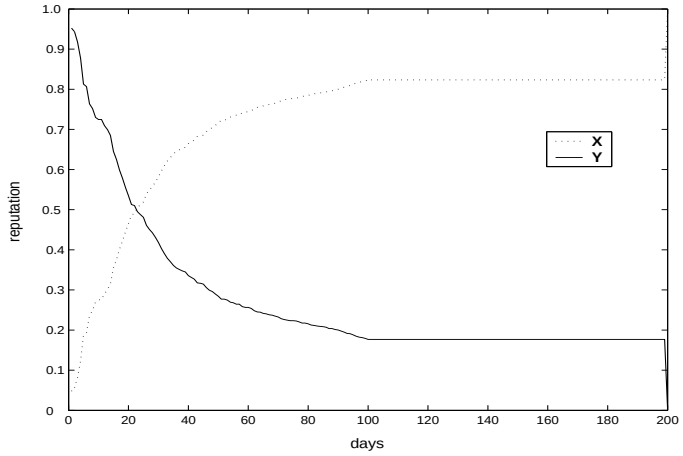

FIGURE 4. User x dominates interactions

The second test represents the results from the cases with one other user (x performs roughly $\frac{1}{2}$ of the total system operations) and with 10 other users (x performs roughly $\frac{1}{11}$ of the total system operations). Figure 5 shows that reputation from this role depends on the proportion of the total system operations.

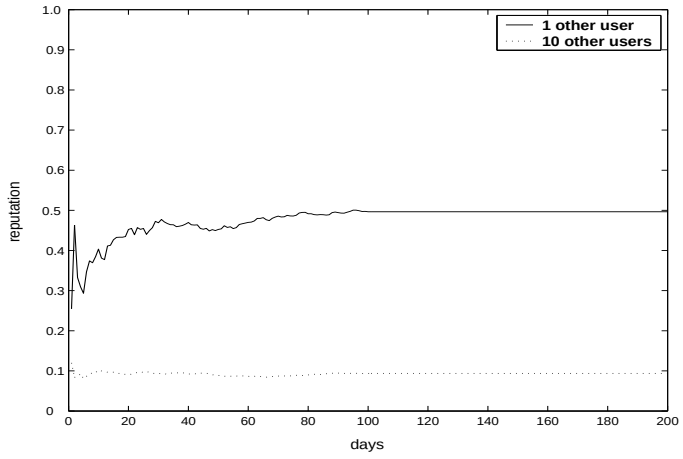

FIGURE 5. Random distribution of operations amongst all users

Notice that the reputation values from this role are much smaller when 11 users are performing operations compared to only 2. This demonstrates that the reputation values

from this role are divided among however many users are performing operations.

6.3. Content Provider Role

Two simulations are carried out to demonstrate the function of the content provider role. For simplicity, the 26 letters of the alphabet are used to represent 26 distinct keywords. Each of the two simulations were run twice with 1 and 3 agents. The first simulation was to model destroying identity. The user and agent(s) start with identical keyword lists. Additional keywords are added to the user list every day after day 50. The second simulation was to model building true identity. The agents start with full lists and the user list is empty. After day 50, the user adds one of the agent keywords to the user list every second day until the user has all the keywords. Figure 6 shows reputation values corresponding to building and destroying identity with both 1 and 3 agents. It demonstrates that the degree of role fulfillment does not decay over time like some of the other roles. As the agent and user keyword lists match, higher reputation is obtained from this role.

6.4. Administrative Feedback Role

The experimental simulations described here are intended to give the reader a better understanding of the behavior of the administrative feedback role. Two test cases are presented and several runs are executed for each case to demonstrate the effect of various values of α. For the first 50 days, user x receives random positive feedbacks. On day 51, the feedbacks become negative until day 100.

Figure 7 shows how reputation values for the administrative feedback role are directly affected by the value of the feedbacks received. In Figure 7, we demonstrate that the reputation value is proportionate to the value of the feedbacks received. When the negative feedbacks start on day 51, the reputation values immediately begin to drop and eventually end up below zero.

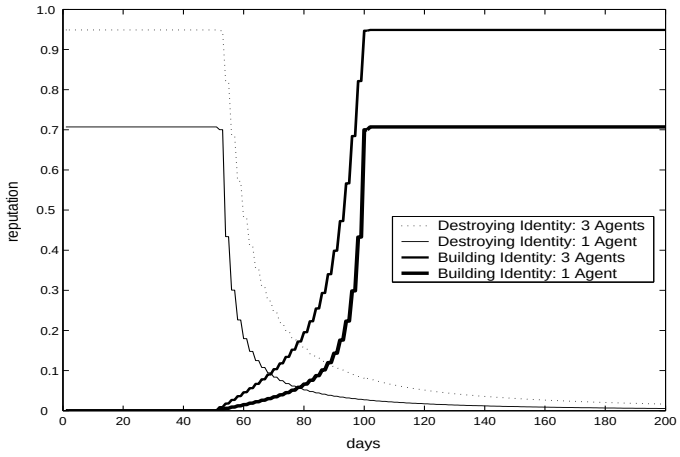

FIGURE 6. Random distribution of operations amongst all users

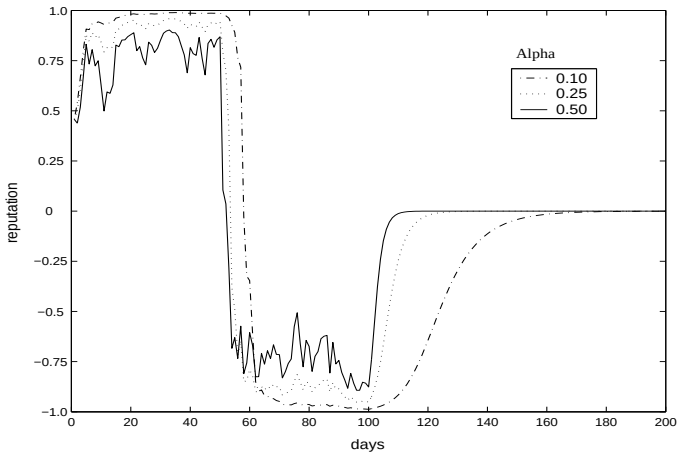

Figure 7. Reputation from positive to negative feedback

6.5. Longevity Role

The longevity role depends upon the aggregation of all the roles. It is necessary to demonstrate the impact of the longevity role on the overall reputation. To achieve an aggregation of the reputation values obtained from each role, the values were each scaled by 1/5 and the results summed to obtain the values plotted here. Two cases are considered. The first case models the possible behavior of a 'good' or reputable user. The second case demonstrates the results of 'bad' behavior. The good behavior includes the following: providing feedback about others with positive reputation, regularly performing actions on the system, providing valid content and receiving positive feedbacks. The bad behavior includes the following: providing no feedback, actions or content, and receiving negative feedbacks. Figure 8 shows that the longevity role has a small positive impact on positive reputations and none on negative reputations.

The formalization presented here takes an essentially distinct direction in comparison to existing models of reputation or trust. That is to say that no models like it exist; results from comparing these models, therefore, would be of little value.

7. CONCLUSION

Reputation is ascribed to the individual by society and is based on external observable behavior. The chosen behavior to observe is based on the criteria that it satisfies role fulfillment. As such, a set of roles for the multiagent system must be clearly defined and easily observable.

Through a summation of the degree of satisfiability of each of these roles, a reputation can be established. Reputation can then be extended to account for components of trust, social values, and social segregation in an objective way.

In this paper, the roles of the information sharing society are formalized and the final reputation value is consequently formalized. Through the formalization of reputation, partial formalization of trust is given. It is important to note that trust by an agent is based

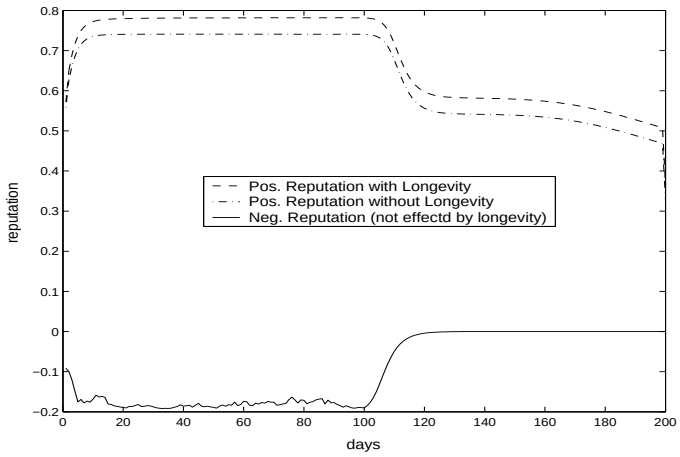

FIGURE 8. Longevity effect on reputation - Case 1

on many other components than reputation. Formalization of trust can only be achieved through formalization of the related concepts. These related concepts include the following: self-esteem, reputation, defamation, expectations and prestige.

Simulation studies are carried out to examine the effect of roles defined for building reputation in an information sharing multiagent system. The simulation results shows that the roles, defined for building reputation in an information-sharing MAS environment, react to different agent and user actions in a manner consistent with the formal definitions. One item that our studies bring to light is the fact that the reputation resulting from the interactivity role is split among all the users. This seems problematic and will require some additional work.

8. ACKNOWLEDGMENTS

This work was partially funded by the Natural Science and Engineering Research Council of Canada (NSERC) through grant RGPIN 227441-00 to Dr. Ali Ghorbani.

References

Alfarez Abdul-Rahman and Stephen Hailes. Relying on Trust to find Reliable Information. Dept. of Computer Science, University College London, London WC1E 6BT, UK, 1999.

Abdul-Rahman, Alfarez, and Hailes, Stephen. Supporting Trust in Virtual Comminities. Dept. of Computer Science, University College London, London WC1E 6BT, UK, 2000.

Barber, B. *The logic and limits of Trust*, pp. 9–17. Grammercy Press, New Brunswick, NJ, 1959.

Blair, M. The SocioWeb. http://www.socioweb.com/markbl/socioweb; accessed 12-15-01.

Dellarocas, C. Mechanisms for Coping with Unfair Ratings and Discriminatory Behavior in Online Reporting Systems. http://ccs.mit.edu/dell/icis2000.pdf

Duhaime, L. Canadian Defamation Law. http://www.duhaime.org/ca-defam.htm; accessed 12-20-01.

Garfinkel, H. *Studies in Ethnomethodology*, pp. 172–173. Prentice-Hall, Englewood Cliffs, NJ, 1967.

Global Campaign for Free Expression. Defining Defamation: Principles on Freedom of Expression and Protection of Reputation. http://www.article19.org/docimages/714.htm; accessed 1-15-02.

Luhmann, N. *Trust and Power*. Wiley, New York, NY, 1979.

Marsh, S. and Masrour, Y. Agent Augmented Community Information — The ACORN Architecture. *Proceedings CASCON 97*, pp. 72–81, Toronto, 1997.

Marsh, S., Ghorbani, A. A. and Bhavsar, V. C. The ACORN Multi-Agent System. Submitted to Autonomous Agents and Multi-Agent Systems, Kluwer Academic Publishers, August 2001.

UK, N., Press, H. and Erving, G. *The Presentation of Self in Everyday Life*. Horizon Scientific Press, Norfolk, UK, 1959.

Castelfranchi, C. and Falcone, R. Principles of Trust for MAS: Cognitive Anatomy, Social Importance, and Quantification. In *Proceedings of the 3rd International Conference on Multi-Agent Systems*, pp. 72-79, 1998.

Sabater, Jordi and Sierra, Carles. REGRET: A Reputation Model for Gregarious Societies. CSCIC-Spanish Scientific REsearch Coucil, Bellaterra, Catalonia, Spain, 2000.

Josang, Audun. The Right Type of Trust for Distributed Systems Department of Telematics, THe Norwegian University of Science and Technology N-7034. Trondheim, 1996.

Sander, Tomas and Tschudin Christian F. In *Protecting Mobile Agents Against Malicious Hosts*. International Computer Science Institute, Berkeley, CA 94704, USA, February 1998.

Braynov Sviatoslav. Trust Revelation in Multiagent Interaction. Department of Computer Science and Engineering State University of New York at Buffalo. Buffalo, NY USA.

Birk Andreas. Boosting Cooperation by Evolving Trust. Vrije Universiteit Brussel, Artificial Intelligence Laboratory, Brussels, Belgium.