Proceedings of the Symposium

# Social Networks and Multi-Agent Systems Symposium (SNAMAS-09)

A symposium at the AISB 2009 Convention (6-9 April 2009)
Heriot-Watt University, Edinburgh, Scotland

Symposium Chairs
Giulia Andrighetto
Guido Boella
Jaime Sichman
Harko Verhagen

Published by SSAISB:
The Society for the Study of Artificial Intelligence
and the Simulation of Behaviour
http://www.aisb.org.uk/

**ISBN - 1902956753**

# Social Networks and Multi-Agent Systems Symposium (SNAMAS-09)

A two-day symposium at AISB 2009 (6-9 April 2009).
http://snamas.di.unito.it/index.html

## PROGRAMME CHAIRS

Giulia Andrighetto, ISTC-CNR Rome
Guido Boella, University of Turin
Jaime Sichman, University Sao Paulo
Harko Verhagen, Stockholm University

## INTRODUCTION

One of the most interesting research topics in the field of multiagent systems is the definition of models with the aim of representing social structures such as organizations and coalitions, to control the emergent behavior of open systems. Organizations and coalitions are composed by individuals, related to each other by different possible kinds of relations such as dependencies on goals, conflicts on resources, similar beliefs and so on. One important issue is how to represent these relations. Moreover, like human organizations, these social structures are characterized also by an high degree of dynamism.

In dealing with societal issues, the multiagent systems field took inspiration mostly from organizational theory in economics and legal theory, while less attention is devoted to the research area describing the relations among the individuals inside human organizations and their dynamics: social network analysis.

Social network analysis has emerged as a key technique in modern sociology, anthropology, social psychology, communication studies, information science, organizational studies, economics as well as a popular topic of study. Social network analysis views social relationships in terms of nodes and ties. Nodes are the individual actors within the networks, and ties are the relationships between the actors. Research in a number of academic fields has shown that social networks play a critical role in determining the way problems are solved, organizations are run, and the degree to which individuals succeed in achieving their goals.

Despite the common object of study, multiagent systems and social network analysis use concepts like agents, relationships, dependencies, and so on which often have only superficial similarities. The aim of this symposium is to underline the differences and the similarity points between social network analysis and multiagent systems in the representation of the social structures and their dynamics, and to promote the interchange of knowledge and methodologies among the two research fields.

## TOPICS

Topics of interest include but are not limited to:

- Emergent behaviour in multiagent systems and social networks analysis
- Simulation of social systems
- Learning evolution and adaptation in multiagent systems and social networks analysis
- Artificial social systems
- Societal aspects
- Models of personality, emotions and social behaviour
- Organizations in Multiagent systems and Social Networks

## PROGRAMME COMMITTEE

Guido BOELLA, University of Turin, ITALY
Kathleen CARLEY, Carnegie Mellon University, USA
Cristiano CASTELFRANCHI, ISTC-CNR Rome, ITALY
Rosaria CONTE, ISTC-CNR Rome, ITALY
Gustavo A. Gimenez LUGO, Federal Technological University of Parana, BRASIL
Jaime SICHMAN, University of Sao Paolo, BRASIL
Carles SIERRA, IIIA-CSIC, SPAIN
Pietro TERNA, University of Turin, ITALY
Leendert VAN DER TORRE, University of Luxembourg, LUXEMBOURG
Harko VERHAGEN, Stockholm University, SWEDEN
Serena VILLATA, University of Turin, ITALY

## Table of Contents

# Trust and Relational Capital

**Cristiano Castelfranchi, Rino Falcone, and Francesca Marzo**

**Istituto di Scienze e Tecnologie della Cognizione**

**National Research Institute of Italy (ISTC-CNR)**

**Abstract.** Trust can be viewed at the same time as an instrument both for an agent selecting the right partners in order to achieve its own goals, and for an agent of being selected from other potential partners in order to establish with them a cooperation/collaboration and to take advantage from the accumulated trust. In this paper we will analyze trust as the agents' *relational capital*. Starting from the classical dependence network with potential partners, we introduce the analysis of what it means for an agent to be trusted and how this condition could be strategically used from it for achieving its own goals, that is, why it represents a form of power. The idea of taking another agent's point of view is especially important if we consider the amount of studies in social science that connect trust with *social capital* related issues. Although there is a big interest in literature about 'social capital' and its powerful effects on the wellbeing of both societies and individuals, often it is not clear enough what is it the object under analysis. Individual trust capital (relational capital) and collective trust capital not only should be disentangled, but their relations are quite complicated and even conflicting. To overcome this gap, we propose a study that first attempts to understand what trust is as *capital of individuals*. In which sense "trust" is a capital. How this capital is built, managed and saved. In particular, how this capital is the result of the others' beliefs and goals. Then we aim to analytically study the cognitive dynamics of this object.

## 1 INTRODUCTION

In multi-agent systems trust is a growing field of analysis and research and ways to calculate it have already been introduced to enhance studies on commercial partnership, strategic choice, and on coalition formation. In particular, in almost the present approaches the focus is on the trustier and on the ways for evaluating the trustworthiness of other possible trustees. In fact, there are no so many studies and analyses about the model of *being trusted*. Also our socio-cognitive model of trust (1, 2) was about the cognitive ingredients for trusting something or somebody, and how trust affects decision, which are the sources and the basis for trusting, and so on; we never modelled what does it mean to be trusted (with the exception of the work on trust dynamics (3) in which the focus was on the reciprocation and potential influences on the trustworthiness) and why it is important. In this paper we address this point, analyzing what it means that trust represents a strategic resource for agents that are trusted, proposing a model of 'trust as a capital' for individuals and suggesting the implication for strategic action that can be performed. Our thesis is that to be trusted: i) increases the chance to be requested or accepted as a partner for exchange or cooperation; ii) improves the 'price', the contract that the agent can obtain.

The need of this new point of view directly derives from the fact that in multi-agent systems it is strategically important not only to know who is trusted by whom and how much, but also to understand how being trusted can be used by the trustee. It has been already shown that using different levels of trust represents an advantage in performing some task such as allocating task or choosing between partners. Therefore, having "trust" as a cognitive parameter in agents' decision making can lead to better (more efficient, faster etc.) solutions than proceeding driven by other kind of calculation such as probabilistic or statistical one. This study already represented an innovation since usually trust has been studied as an effect rather than a factor that causes the developing of social network and their maintenance or structural changing.

In order to improve this approach and to better understand dynamics of social networks, now we propose a study of what happens on the other side of the two-way trust relationship, focusing on the trustee, in particular on a cognitive trustee. Our aim is an analytical study of what it means to be trusted. The idea of taking the other point of view is particularly important if we consider the judge amount of studies in social science that connect trust with social capital related issues.

Our claims are: (a) to be trusted usually is an advantage for the trustee (agent $Ag_i$); more precisely received trust is a capital that can be invested, and that requires decision and costs to be cumulated; (b) it is possible to measure this capital, which is relational, that is it depends on a position in a network of relationships; (c) trust has different sources: from personal experience that the other agents have with $Ag_i$; from circulating reputation of $Ag_i$; from $Ag_i$ belongingness to certain groups or categories; from the signs and the impressions that $Ag_i$ is able to produce; (d) the value of this capital is context dependent (and market dependent) and dynamic; (e) received trust strongly affects the 'negotiation power' of $Ag_i$ that cannot simply be derived from the "dependence bilateral relationships".

Although there is a big interest in literature about 'social capital' and its powerful effects on the wellbeing of both societies and individuals, often it is not clear enough what is it the object under analysis. To overcome this lack, we propose a study that first attempts to understand what trust is as capital of individuals. How is it possible to say that "trust" is a capital? How is this capital built, managed and saved? Then we aim to analytically study the cognitive dynamics of this object, with a particular focus on how they depend on beliefs and goals.

## 2 TRUST AND RELATIONAL CAPITAL

Social Capital (4, 5, 6, 7) can be seen as a multidimensional concept and can be studied in its relation both with social norms and shared values and with networks of interpersonal relations. While in the former case studies about conventions and collective attribution of meanings can be useful to study how social capital can be a capital for the society, in the latter, one of the basic issues that need to be studied is how it can happen that networks of relations can be built, which ways they develop, and how they can both influence individual behaviours and be considered as an individual capital.
We also would like underline that social capital is an ambiguous concept. By social a lot of scholars mean in fact 'collective', some richness, advantage of any for the collective; something that favors cooperation, and so on. On the contrary we assume here (as a first step) an individualistic perspective, considering the advantages of the trusted agent, not the advantages for the collectivity, and distinguishing between 'relational capital' (8) and the more ambiguous and extended notion of 'social capital'. The individual (or organization) $Ag_i$ could use its capital of trust, for anti-social purposes. In economic literature the term "capital" refers to a commodity itself used in the production of other goods and services: it is, then, seen as a human-made input created to permit increased production in the future. The adjective "social" is instead used to claim that a particular capital not only exists in social relationships but also consists in some kind of relationships between economical subjects. It is clear that for the capital goods metaphor to be useful, the transformative ability of social relationships to become a capital must be taken seriously. This means that *we need to find out what is the competitive advantage not simply of being part of a network, but more precisely of being trusted in that network.*

The additional value of trusting has been shown as a crucial argument in decision making and in particular in choice of rely on somebody else for achieving specific goals included in the plans of the agents. In these studies trust has been analysed as valuation of the other and expectations on it, and has being shown how these characteristics and mechanisms, being part of the decision process at the cognitive level, represent an advantage for the society in terms of realizing cooperation among its actors and for the trustier in terms of efficiency of choices of delegation and reliance (9).

Changing the point of view, we now want to focus on the trusted agent. What does imply to be trusted for the trustee? The intuitive answer could be that: i) the probability to be chosen for exchange or for partnership will grow; but also that: ii) the *negotiation power* of that agent will increase.

However, to account for this it is necessary to rethink the whole theory of negotiation power based on dependence (10,11,12,13). Try to build a theory of dependence including trust does not mean to base the theory of social capital on dependence, but to admit that the existent theory of dependence network and the consequent theory of social power is not enough without the consideration of trust. What we need, then, is a comprehensive theory of trust from the point of view of the trusted agent, in order to find out the elements that, once added to the theory of dependence, can explain the *individual social power in a network*, on one hand, and, on a second phase, the *social capital meant as a capital for the society.*

Once a quantitative notion of the value of a given agent is formulated calculating on *how much the agent is valued by other agents in a given market for a given task*, we can say that this trust-dependent value is a real capital. It consists of all the relationships that are possible for the agent in a given market and, together with the possible relationships in other markets, it is the so-called *relational capital* of that agent. It differs from simple relationships in given networks, which are a bigger set, since it only consists of relationships the agent has with those who not only need it but have a good attitude toward it and, therefore, who are willing to have it as a partner. How much it is appreciated and requested? How many potential partners depends on $Ag_i$ and would search for $Ag_i$ as partner? How many partners would be at disposal for $Ag_i$'s proposals of partnership, and what "negotiation power" would $Ag_i$ have with them?

These relationships form a capital because (as any other capital) it is the result of investments and is costly cumulated to be spent. In a certain sense it represents a strategic tool to be competitive, and, as well as it happens with other capitals such as the financial one, it is sometimes even more important than the good which is sold (being it either a service or a material good). For example when $Ag_i$ decides of non keeping a promise to $Ag_j$, it knows that $Ag_j$'s trust in $Ag_i$ will decrease: is this convenient for future relationships with $Ag_j$? Will $Ag_i$ need counting on $Ag_j$ in future? Or, is this move convenient for reputation and other relationships? For this reason it is very important to study how it is possible for the agent to cumulate this capital without deteriorating or waste it: since the relational capital can make the agent win the competition even when the good it offers is not the best compared with substitutive goods offered in the market, it should be shown quantitatively what this means and what kind of dynamical relationships exist between quality of offered good and relational capital.

## 3 COGNITIVE MODEL OF BEING TRUSTED

### 3.1 Objective and Subjective Dependence

The theory of trust and the theory of dependence are not independent from each other. Not only because – as we modelled (1, 2), before deciding to actively trust somebody, to rely on it ($Ag_i$), one ($Ag_j$) has to be dependent on $Ag_i$: $Ag_j$ needs an action or a resource of $Ag_i$ (at least $Ag_j$ has to believe so). But also because *objective* dependence relationships (10) that are the basis of adaptive social interactions, are not enough for predicting them. *Subjective* dependence is needed (that is, the dependence relationships that the agents know or at least believe), but is not sufficient; it is also necessary to add to (i) the belief of being dependent, of needing the other, (ii) the belief of the trustworthiness of the other, of the possibility of counting upon it. If I wouldn't not feel dependent on, I couldn't rely on the other.
The theory of dependence includes in fact two types of dependences: (1) the *objective dependence*, which says who needs whom for what in a given society (although perhaps ignoring this). This dependence has already the power of establishing certain asymmetric relationships in a potential market, and it determines the actual success or failure of the reliance and transaction; (2) the *subjective (believed) dependence*, which says who is believed to be needed by who. This dependence is what determines relationships in a real

market and settles on the negotiation power; but it might be illusory and wrong, and one might rely upon unable agents, while even being autonomously able to do as needed.

More Formally, let $Agt=\{Ag_1,..,Ag_n\}$ a set of *agents*; we can associate to each agent $Ag_i \in Agt$:

- a set of *goals* $G_i=\{g_{i_1},..,g_{i_q}\}$;
- a set of *actions* $Az_i=\{\alpha_{i_1},..,\ \alpha_{i_z}\}$; these are the elementary actions that $Ag_i$ is able to perform;
- a set of plans $\Pi=\{p_{i_1},..,p_{i_s}\}$; the $Ag_i$'s plan library: the set of rules/prescriptions for aggregating the actions; and
- a set of *resources* $R_i=\{r_{i_1},..,r_{i_m}\}$.

The achievement/maintenance of each goal needs of actions/plans/resources. Then, we can define the *dependence relationship* between two agents ($Ag_j$ and $Ag_i$) with respect to a goal $g_{jk}$, as: *Obj-Dependence* $(Ag_j, Ag_i, g_{jk})$ and say that:

*An agent $Ag_j$ has an Objective Dependence Relationship with agent $Ag_i$ with respect to a goal $g_{jk}$ if for achieving $g_{jk}$ are necessary actions, plans and/or resources that are owned by $Ag_i$ and not owned by $Ag_j$. More in general, $Ag_j$ has an Objective Dependence Relationship with $Ag_i$ if for achieving at least one of its goals $g_{jk} \in G_j$, are necessary actions, plans and/or resources that are owned by $Ag_i$ and not owned by $Ag_j$.*

As in (12) we can introduce the *unilateral*, *reciprocal*, *mutual* and *indirect* dependence (see Figure1). In very short and simplified terms, we can say that the difference between reciprocal and mutual is that the first is on different goals while the second is on the same goal.

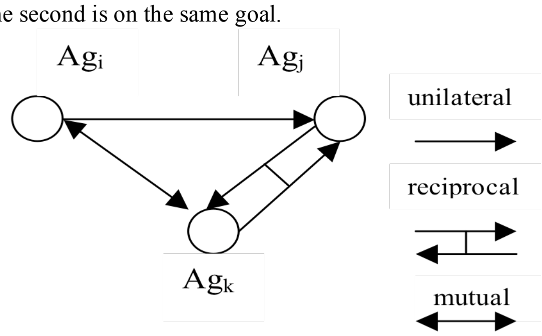


Figure1

If the world knowledge would be perfect for all the agents, the above described objective dependence would be a common belief about the real state of the world. In fact, the important relationship is the network of dependence *believed by each agent*. In other words, we cannot only *associate* to each agent a set of goals, actions, plans and resources, but we have to evaluate these sets as believed by each agent (the subjective point of view), also considering that they would be partial, different each of others, sometime wrong, and so on. In more practical terms, each agent will have a different (subjective) representation of the dependence network as exemplified in Figure1. So, we introduce the $Bel_kG_z$ that means the Goal set of $Ag_z$ believed by $Ag_k$. The same for $Bel_kAz_z$, $Bel_k\Pi_z$, and $Bel_kR_z$. In practice, the dependence relationships should be re-modulated on the basis of the agent subjective interpretation.

We introduce the *Subj-Dependence*$(Ag_j, Ag_i, g_{jk})$ that represents the $Ag_j$'s point of view with respect to its dependence relationships.

In a first approximation each agent should correctly believe the sets it has, while it could mismatch the sets of other agents.

We define *Dependence-Network(Agt,t)* the set of dependence relationships (both subjective and objective) among the agents included in *Agt* set at the time *t*. Each agent $Ag_j \in Agt$ must have at least one dependence relation with another agent in *Agt*.

## 3.2 Dependence and Negotiation Power

Given a *Dependence-Network(Agt,t)*, we define *Objective Potential for Negotiation* of $Ag_j \in Agt$ about one of its own goals $g_{jk}$ -and call it $OPN(Ag_j, g_{jk})$- the following function:

$$OPN(Ag_j, g_{jk})=f\left(\sum_{i=1}^{n}\frac{1}{1+p_{ki}}\right)$$

Where:

*f* is in general a function that preserves monotonicity (we will omit this kind of functions in the next formulas);

*n* represents the number of agents in *Agt* set that have a dependence relation with $Ag_j$ with respect to $g_{jk}$ (this dependence relation should be either reciprocal or mutual: in other words, there should also be an action, plan, or resource owned by $Ag_j$ that is necessary for $Ag_i$);

$p_{ki}$ is the number of agents in *Agt* that are competitors with the $Ag_i$ on the same actions/plans/resources (useful for $g_{jk}$) in a not compatible way ($Ag_i$ is not able to satisfy at the same time all the agents).

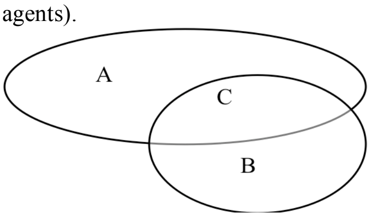


Figure 2

In Figure2 we show the objective dependence of $Ag_j$: A represents the set of agents who depend from $Ag_j$ for something (actions, plans, resources), B represents the set of agents from which $Ag_j$ depends for achieving an own specific goal $g_{jk}$. The intersection between A and B (part C) is the set of agents with whom $Ag_j$ could potentially negotiate for achieving $g_{jk}$. The greater the overlap the greater the *negotiation power* of $Ag_j$ in that context.

However, the negotiation power of $Ag_j$ also depends on the possible alternatives that its potential partners have: the few alternatives to $Ag_j$ they have, the greater its negotiation power (see below).

We can define the *Subjective Potential for Negotiation* of $Ag_j \in Agt$ about one of its own goals $g_{jk}$ -and call it $SPN(Ag_j, g_{jk})$- the following function:

$$SPN(Ag_j, g_{jk}) = \sum_{i=1}^{n}\frac{1}{1+p_{ki}}$$

Where we have the same meanings as for the previous formula but now we make reference to the believed (by $Ag_j$) dependence relations (not necessarily true in the world): in particular are believed both *n* (the number of direct dependences) and *p* (the indirect, competitive dependences).

Analogously, we can interpret Figure2 as the set of believed relationships (by $Ag_j$) among the agents. In this case we have the subjective point of view. It is also possible to introduce a modulation factor that takes into account the special kind of dependence: reciprocal (x=r), mutual (x=m):

$$SPN(Ag_j, g_{jk}) = \sum_{i=1}^{n}\frac{m_x}{1+p_{ki}} \quad \text{with } 0<m_x<1$$

Usually, we can say that $m_m \geq m_r$. More in general, we can say that the

*Subjective Potential for Negotiation* of $Ag_j \in Agt$ about the whole set of its own goals ($G_j$) in the *Dependence-Network(Agt,t)* is:

$$SPN(Ag_j, G_j) = \frac{1}{s}\sum_{k=1}^{s}\sum_{i=1}^{ns}\frac{m_x}{1+p_{ki}}$$

Where $s$ is the number of goals of $Ag_j$, and $ns$ is the number of other agents in the set $Agt$, that have a dependence relation with $Ag_j$ with respect to the goal $g_{jk}$.

$p_{ki}$ is the number of agents in $Agt$ that are competitors with the $Ag_i$ on the same actions/plans/resources (useful for $g_{jk}$) in a not compatible way.

In words, the global subjective potential for negotiation of an agent in a dependence network with respect to all its own goals is the sum of beliefs above showed[1].

## 3.3 The Trust Role in Dependence Networks

Before taking into account the trustee's point of view we would like to introduce into the dependence network also the trust relationships. In fact, *although it is important to consider dependence relationship between agents in a society, there will be no exchange in the market if there is not trust to enforce these connections*. Considering the analogy with the Figure2, we will have now a representation as given in Figure3 (where D includes the set of agents that $Ag_j$ considers trustworthy for achieving $g_{jk}$).

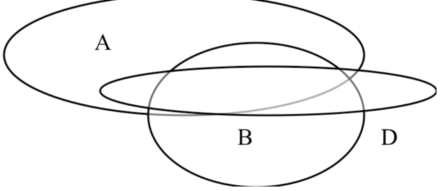


Figure3

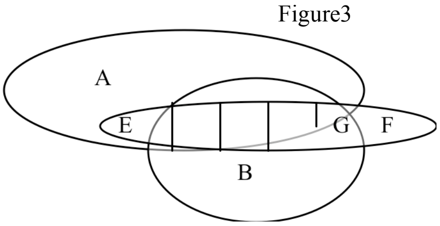


Figure4

We have now a new subset (showed outlined in Figure4) containing the potential agents for negotiation. The analysis of the part E, F and G will result in: part E includes agents who depend form $Ag_j$, who are trusted but on different tasks; part F includes agents not depending from $Ag_j$ and trusted on different tasks; part G includes agents trusted for achieving the goal $g_{jk}$ but not depending from $Ag_j$.

Not only the decision to trust presupposes a belief of being dependent, but notice that a dependence belief (*BelDep*) implies on the other side a piece of Trust (as modelled in (1,2)). In fact to believe to be dependent means:

- (*BelDep-1*) to believe not to be able to perform action $\alpha$ and to achieve goal $g$; and

- (*BelDep-2*) to believe that $Ag_i$ is *able* and in condition to achieve $g$, to perform $\alpha$.

Notice that (*BelDep-2*) is precisely one component of Trust in our analysis: the *positive evaluation* of $Ag_i$ as competent, able, skilled, and so on. However, the other fundamental component of trust as evaluation is lacking: reliability, trustworthiness: $Ag_i$ really intends to do, is persistent, is loyal, is benevolent, etc. Thus he will really do what $Ag_j$ needs.

Given the basic role played by "believed networks of dependence", established by a believed relationship of dependence based on a belief of dependence, and given that this latter is one of the basic ingredient of trust as a mental object, we can claim that this overlap between theories is the crucial issue and our aim is namely to study it deeply.

So introducing also in the *Subjective Potential for Negotiation* (of $Ag_j \in Agt$ about one of its own goals $g_{jk}$) the basic beliefs about trust (1,2) we have:

$$SPN(Ag_j, g_{jk}) = \sum_{i=1}^{n}\frac{Bel_j(DoA_i * DoW_i)}{1+p_{ki}}$$

*Where:*

$m_m = m_r = 1$;

$DoA_i$ is the degree of ability (with respect to the goal $g_k$) of the agent $Ag_i$ as believed by $Ag_j$ ($Bel_j(DoA_i)$);

$DoW_i$ is the degree of willingness (with respect to the goal $g_k$) of the agent $Ag_i$ as believed by $Ag_j$ ($Bel_j(DoW_i)$);

$DoA_i$ and $DoW_i$ respectively represent the $Ag_i$'s ability and willingness of using actions/plans/resources for the goal $g_k$. We do not consider here the possible relations between the values of $DoA_i$ and $DoW_i$ with the $p_{ki}$ variable. $1 \geq DoA_i, DoW_i \geq 0$.

Let us, now, explicitly recall what are the cognitive ingredients of trust and reformulate them from the point of view of the trusted agent. In order to do this, it is necessary to limit the set of trusted entities. It has in fact been argued that trust is a mental attitude, a decision and a behavior that only a cognitive agent endowed with both goals and beliefs can have, make and perform. But it has been underlined, also, that the entities that is trusted is not necessarily a cognitive agent.

When a cognitive agent trusts another cognitive agent, we talk about social trust. We consider that the set of actions, plans and resources owned/available by an agent can be useful for achieving a set of tasks ($\tau_1$, ..., $\tau_r$).

We take now the point of view of the trustee agent in the dependence network: so we present a cognitive theory of trust as a capital, which is, in our view, a good starting point to include this concept in the issue of negotiation power. That is to say that if somebody is potentially strongly needed by other agents, but it is not trusted, its negotiation power does not improve.

We call the *Subjective Trust Capital* of $Ag_i \in Agt$ about an its own task $\tau_k$ the function:

$$STC(Ag_i, \tau_k) = \sum_{j=1}^{n} Bel_i(Bel_j DoA_i * Bel_j DoW_i)$$

Where $n$ is the number of agents need the task $\tau_k$.

$Ag_j$, $Ag_i \in Agt$.

In words, the cumulated trust capital of an agent $Ag_i$ with respect to a specific task $\tau_k$, is the sum (on all the agents need that specific task in the network dependence) of the corresponding abilities and willingness believed by each dependent agent. The

[1] An interesting problem is that an agent could be a competitor towards itself for achieving its own goals; for example: 1) $Ag_j$ needs action $\alpha_r$ both for $g_s$ and $g_t$ and there is only an agent in $Agt$ that has $\alpha_r$ but is unable to provide two times the action $\alpha_r$. 2) $Ag_j$ needs action $\alpha_r$ for $g_r$ and $\alpha_s$ for $g_s$ and for both the actions $\alpha_r$ and $\alpha_s$ it depends only from $Ag_i$ that can provide only an action.

subjectivity consists in the fact that both the network dependence and the believed abilities and willingness are believed by (the point of view of) the agent $Ag_i$.
As showed in (2) we call Degree of Trust of the Agent $Ag_j$ on the agent $Ag_i$ about the task $\tau_k$ (DoT($Ag_j$ $Ag_i$ $\tau_k$)):
$$DoT(Ag_j\ Ag_i\ \tau_k) = Bel_j DoA_i * Bel_j DoW_i$$
At the same way we can also call the self-trust of the agent $Ag_i$ about the task $\tau_k$ we can write:
$$ST(Ag_i, \tau_k) = Bel_i(DoA_i * DoW_i)$$
From the comparison between $STC(Ag_i, \tau_k)$, $DoT(Ag_j\ Ag_i\ \tau_k)$ and $ST(Ag_i, \tau_k)$ a set of interesting actions and decision are taken from the agents (we will see in the next paragraph).
Starting from the Trust Capital we would like evaluate the usable part of this trust capital. In this sense, we introduce the *Subjective Usable Trust Capital* of $Ag_i \in Agt$ about an its own task $\tau_k$ as:
$$SUTC(Ag_i, \tau_k) = \sum_{j=1}^{n} \frac{Bel_i(Bel_j DoA_i * Bel_j DoW_i)}{1 + p_{kj}}$$
where $p_{kj}$ is (following the $Ag_i$'s belief about the beliefs of $Ag_j$) the number of other agents in the dependence network that can achieve the same task with a trust value comparable with the one of $Ag_i$. We have two *comparable trust values* when the difference between them is in a range under a given threshold that could be considered meaningless with respect to the achievement of the task.

## 4 DYNAMICS OF RELATIONAL CAPITAL

What has not been considered enough in organization theory is the fact that the *relational capital* is peculiar in its being crucially based on beliefs: again, what makes relationships become a capital is not simply the structure of the networks (who "sees" whom and how clearly) but the levels of trust which characterizes the links in the networks (who trusts whom and how much). Since trust is based on beliefs – including, as we said, also the believed dependence (who needs whom) – it should be clear that relational capital is a form of capital, which can be manipulated by manipulating beliefs.

### 4.1 Increasing, decreasing and transferring

For what concerns the dynamic aspects of this kind of capital, it is possible to make hypotheses on how it can increase or how it can be wasted, depending on how each of basic beliefs involved in trust are manipulated.
First, let us consider what kind of strategies can be performed to enforce the other's dependence beliefs and his beliefs about agent's competence.
i) $Ag_i$ can make the other agent dependent on him by making the other lacking some resource or skill (or at least inducing the other to *believe* so).
ii) $Ag_i$ can make the other agent dependent on him by activating or inducing in it a given goal (need, desire) on which the other is not autonomous (14) (or believes so).
iii) Since dependence beliefs is strictly related with the possibility of the others to see the agent in the network and to know her ability in performing useful tasks, the goal of the agent who wants to improve her own relational capital will be to *signaling* her presence and her skills (15,16,17). While to show her presence she might have to shift her position (either physically or figuratively like, for instance, changing her field), to communicate her skills she might have to hold and show something that can be used as a signal (such as certificate, social status etc.). This implies, in her plan of actions, several and necessary sub-goals to make a signal. This sub-goals are costly to be reached and the cost the agent has to pay to reach them can be taken has the evidence for the signals to be credible (of course without considering cheating in building signals). It is important to underline that using these signals often implies the participation of a third subject in the process of building trust as a capital: a third part which must be trusted (2). We would say the more the third part is trusted in the society, the more expensive will be for the agent to acquire signals to show, and the more this signals will work in increasing the agent's relational capital. We will see later how this is related with the process of transferring trust from an agent to another (building reputation).
Obviously also $Ag_i$'s *previous performances* are 'signals' of trustworthiness. And this information is also provided by the circulating *reputation* of $Ag_i$ (18, 19).
In formal terms, we can say that $Ag_i$ has to work for increasing: $Bel_j DoA_i$ and consequently $Bel_i Bel_j DoA_i$.
iv) Alternatively, $Ag_i$ could work for reducing the believed (by $Ag_j$) value of ability of each of the possible competitors of $Ag_i$ (in number of $p_{kj}$) on that specific task $\tau_k$.
Let us now consider how willingness beliefs can be manipulated. In order to do so, consider the particular strategy performed to gain the other's good attitude through gifts (20). It is true that the expected reaction will be of reciprocation, but this is not enough. While giving a gift the agent knows that the other will be more inclined to reciprocate, but she also knows that her action can be interpreted as a sign of the good willingness she has: since she has given something without being asked, the other is driven to believe that the agent will not cheat on him. Then, the real strategy can be played on trust, sometimes totally and sometimes only partially – this will basically depend on specific roles of agents involved.
Again in formal terms, we can say that $Ag_i$ has to work for increasing: $Bel_j DoW_i$ and as a consequence $Bel_i Bel_j DoW_i$. Alternatively, it could work for reducing the believed (by $Ag_j$) value of willingness of each of the possible competitors of $Ag_i$ (in number of $p_{kj}$) on that specific task $\tau_k$.
An important consideration we have to do is that a dependence network is mainly based on the set of actions, plans and resources owned by the agents and necessary for achieving the agents' goals (we considered a set of tasks each agent is able to achieve). The interesting thing is that the dependence network is modified by the dynamics of the agents' goals, from their variations, from the emergency of new ones, from the disappearance of old ones, from the increasing request of a subset of them, and so on (21). On this basis changes the role of each agent in the dependence network, changes in fact the trust capital of the agents.
*Relational capital* can be also circulated inside a given society. If somebody has a good reputation and is trusted by somebody else, she can be sure this reputation will pass and transfer to other actors – and this is always considered in marketing strategies of making voice circulate. What is not clear yet is how these phenomena work. But when trust on an agent circulates, it is strategically important for the agent to know very well how this happens and which ways (not only figurate) trust takes to expand. In fact, not all the ways are the same: it is possible that

being trusted by a particular agent can mean that she just has one more agent in her relational capital, but gaining the trust of another agent can be very useful to her and exponentially increase her capital thanks to the strategic role or position of this other agent. That said, it should be clear the importance of understanding if and how much an agent is able to manage this potentiality of her capital.

Basically, here also, the role of agents involved play a crucial part: for this reason it is necessary for agent to know the multiplicative factor represented by the recognized and trusted evaluator in the society. It is not necessarily true, in fact, that when somebody trusts somebody else and this trusts a third one, the first one will trust the third one: the crucial question is "which role the first recognize to the second". If the second one is trusted as an evaluator by the first one, than she can trust the third one for specific goals.

Usually how well these transitive process works depends on what kind of broadcasting and how many links the valuator has and how much she is trusted in each of those links, so, basically, it recursively depends on the valuator's relational capital.

## 4.2 Strategic behavior of the trustee

Until now we did not talk about subjective difference in the way trust is perceived by the two parts of the relationship. We must take into account the fact that there is often a difference between how the others actually trust an agent and what the agent believes about; but also between this and the level of trustworthiness that agent perceive in herself. Since being able is not necessarily the cause of trust: it can be the case of a diffuse atmosphere that makes the others trust the agent although the agent has not all the characteristics to be trusted. These subjective aspects of trust are fundamental for managing this capital, since it can be possible that the capital is there but the agent does not know to have it. Can it be possible to use the relational capital even if who uses it is not aware of having it?

At the base of the possible discrepancy in subjective valuation of trustworthiness there is the perception of how much an agent feels trustworthy in a given task and the valuation that agent does of how much the others trust her for that task.

In addition, this perception can change and become closer to the objective level while the task is performed. These factors must be taken into account and studied together with the different components of trust, in order to build hypotheses on strategic actions the agent will perform to cope with her relational capital. Then, we must consider what can be implied by these discrepancies in terms of strategic actions: how they can be individuated and valued? How the trusted agent will react when aware of them? She can either try to acquire competences in order to reduce the gap between others' valuation and her own one, or exploiting the existence of this discrepancy, taking advantage economically of the reputation over her capability and counting on the others' scarce ability of monitoring and testing her real skills.

## 5 CONCLUSIONS

As we said, individual trust capital (relational capital) and collective trust capital not only should be disentangled, but their relations are quite complicated and even conflicting. In fact, since the individual is in competition with the other individuals, he has a better position when trust is not uniformly distributed (everybody trusts everybody), but when he enjoys some form of concentration of trust (an oligopoly position in the trust network); while the collective social capital could do better with a generalized trust among the members of the collectivity.

## REFERENCES


[1] Castelfranchi C., Falcone R., Principles of trust for MAS: cognitive anatomy, social importance, and quantification, *Proceedings of the International Conference of Multi-Agent Systems (ICMAS'98)*, pp. 72-79, Paris, July, 1998.

[2] Falcone R., Castelfranchi C., (2001). Social Trust: A Cognitive Approach, in *Trust and Deception in Virtual Societies* by Castelfranchi C. and Yao-Hua Tan (eds), Kluwer Academic Publishers, pp. 55-90.

[3] Falcone R., Castelfranchi C. (2001), The socio-cognitive dynamics of trust: does trust create trust? In *Trust in Cyber-societies: Integrating the Human and Artificial Perspectives* R. Falcone, M. Singh, and Y. Tan (Eds.), LNAI 2246 Springer. pp. 55-72.

[4] Bourdieu, P. 1983: Forms of capital. In: Richards, J. C. ed. Handbook of theory and research for the sociology of education, New York, Greenwood Press.

[5] Coleman, J. C. 1988: Social capital in the creation of human capital. American Journal of Sociology 94: S95-S120.

[6] Putnam, R. D. 1993: Making democracy work. Civic traditions in modern Italy. Princeton NJ, Princeton University Press.

[7] Putnam, R. D. 2000: Bowling alone. The collapse and revival of American community. New York, Simon and Schuster.

[8] Granovetter, M. (1973). The strength of weak ties. American Journal of Sociology, 78, 1360-1380.

[9] Castelfranchi, C., Falcone, R., (1998) Towards a Theory of Delegation for Agent-based Systems, *Robotics and Autonomous Systems*, Special issue on Multi-Agent Rationality, Elsevier Editor, Vol 24, Nos 3-4, , pp.141-157.

[10] Castelfranchi C., and Conte R., The Dynamics of Dependence Networks and Power Relations in Open Multi-Agent Systems. In Proc. COOP'96 – Second International Conference on the Design of Cooperative Systems, Juan-les-Pins, France, June, 12-14. INRIA Sophia-Antipolis, 1996. P.125-137).

[11] Sichman, J, R. Conte, C. Castelfranchi, Y. Demazeau. A social reasoning mechanism based on dependence networks. In *Proceedings of the 11th ECAI*, 1994.

[12] Castelfranchi, C., Miceli, M. e Cesta, A., Dependence relations among autonomous agents. In E. Werner, Y. Demazeau (Eds), *Decentralized A. I.* - 3 , pp. 215-227, North Holland, Amsterdam, 1992.

[13] Conte, R. e Castelfranchi, C. (1996) Simulating multi-agent interdependencies. A two-way approach to the micro-macro link. In U. Mueller & K. Troitzsch (eds) *Microsimulation and the social science*. Berlin, Springer Verlag, Lecture Notes in Economics.

[14] Castelfranchi, C. Falcone R. (2003), From Automaticity to Autonomy: The Frontier of Artificial Agents, in Hexmoor H, Castelfranchi, C., and Falcone R. (Eds), Agent Autonomy, Kluwer Publisher, pp.103-136.

[15] Schelling, T., *The Strategy of Conflict*. Cambridge, Harvard University Press, 1960.

[16] Spece, M. 1973 Job market signaling. *Quarterly Journal of Economics, 87*, 296-332.

[17] R. Bliege Bird & E. Alden Smith "Signaling Theory, Strategic Interaction, and Symbolic Capital", *Current Antropology*, vol. 46, n.2. April 2005.

[18] R. Conte and M. Paolucci, Reputation in Artificial Societies. Social Beliefs for Social Order. Kluwer 2002.

[19] A. Jøsang and R. Ismail. *The Beta Reputation System*. In the proceedings of the 15th Bled Conference on Electronic Commerce, Bled, Slovenia, 17-19 June 2002.

[20] Cialdini, R. B. 1990: Influence et manipulation, Paris, First.

[21] Pollack, M., Plans as complex mental attitudes in Cohen, P.R., Morgan, J. and Pollack, M.E. (eds), *Intentions in Communication*, MIT press, USA, pp. 77-103, 1990.

# *The Paradox of Elegance –* A Very Short Introduction to the Topology of Complex Social Systems and the "Small World"-Paradigm in the Realm of Law

**Ugo Pagallo[1] and Giancarlo Ruffo[2]**

**Abstract.** This paper aims to follow up the proposal of the AISB'09 Convention organizers in order to promote the interchange of knowledge and methodologies between the "multi-agent systems"-analyses and the "social network" perspectives. The idea is, first of all, to summarise ten years of research on the "small world"-paradigm and to present the most recent work on the topology of complex legal systems. In light of the extraordinary discoveries achieved in this field, we then stress limits and open problems, both at theoretical and practical levels, of the new paradigm. Law seems, indeed, a good field in which scientists can deepen that fundamental interchange of knowledge and methodologies we need in the study of contemporary social systems.

## 1 INTRODUCTION

In order to promote the interchange of knowledge and methodologies between the "multi-agent systems"-analyses and the "social network" perspectives, this paper comes in four sections.

First, we summarise ten years of research on the "small world"-paradigm and, hence, we recall some work on the topology of complex legal systems.

Secondly, we present our own research and results over the last three years on some possible applications of the paradigm to extremely hot legal issues of today's debate as it happens with privacy and copyright.

Thirdly, we mention some other possible applications of the paradigm to the "multi-agent systems"-approach as in the case of the development of legal ontologies.

Finally, we stress both limits and problems of the "small world"-perspective. Notwithstanding its merits in current legal research, Gregory Chaitin's paradox on the "elegance" of computer programs recalls why we cannot demonstrate to have find out Leibniz's (and even Gödel's) characteristica universalis. Law is indeed a good field in which to test that interchange we need among different theoretical outlooks that deal with complex social networks.

## 2 THE LAST DECADE OF THE "SMALL WORLD" PARADIGM

Since the idea of small world-networks first appeared in the pioneering work of Stanley Milgram and later with the sociological research of Mark Granovetter, in few years it has become one of the key words of contemporary scientific research by fostering a large set of empirical studies on the topology of complex systems. Of course, significant effort has been made in order to structure analytical models able to capture the nature of small world-networks. Here it suffices to mention only three of these. The first one was proposed by Watts and Strogatz [1998]: they suggested to rewire randomly a small fraction of the edges belonging to a low-dimensional regular lattice in order to prove that the degrees of separation in the network would exponentially decrease. In fact, the peculiarity of the small world-model depends on the apparent deviation from the properties of both random and regular networks since, like random networks, small world-networks present a short characteristic path length but, like regular networks, they also have high clustering coefficients. This model considers social networks as a structured and ordered world with nodes in the graph tending to be linked each other when they share one or more neighbors. Moreover, random links can determine some shortcuts to distant nodes, thus limiting the diameter of the network.

However, this first small world-model is not easily applicable from an "algorithmic" perspective: Milgram's experiments, as well as everyone's experience in ordinary life, prove not only that shortcuts exist, but that people in a social network are able to find them, with some approximation. This means that if we want to define the behavior of a rational agent living in a small world, we can also assume that he/she can reach a distant point in few steps using local information and without a global "bird's eye view." Hence, distant nodes are not simply linked in a random way, like the small-world model suggests, because connections are biased toward closeness or similarity criteria. Kleinberg [2000] extended the model in order to capture this effect observed in real life networks.

On the other hand, the third work we need to mention is Barabási's research [2002]: Indeed, he noted that most real world networks grow by continuous addition of new nodes whereas the likelihood of connecting to a node would depend upon the node degrees. In other words, this sort of special attachment in a growing system explains what Watts and Strogatz apparently missed, i.e., the power-law distribution of the network in a topological scale-free perspective. This means that most real small world-networks are characterized by few nodes with very high values and by most nodes with small degree. So that the presence of hubs, i.e., a small fraction of nodes with a much higher degree than the average, is the key to comprehend why small world-networks in the real world can be both highly clustered and scale-free when small, tightly interlinked clusters of nodes are connected into larger, less cohesive groups.

To shed further light on the new paradigm let us introduce its legal applications by considering two studies from 2005. The first work is by Seth Chandler [2005] who built an electronic map of 26000 decisions issued by the U.S. Supreme Court from early 19th Century till now. He assumed each case as a node of

[1] Department of Law, University of Torino, Via Sant'Ottavio, 54, Torino, Italy.
Email: ugo.pagallo@unito.it.
[2] Department of Computer Science, University of Torino, Corso Svizzera, 185, Torino, Italy.
Email: giancarlo.ruffo@unito.it.

the network and each citation as a link while links between nodes are intended as directional arrows rather than simple lines. The result is a network with very low density that nevertheless also has a main core. In fact, only 258047 out of 365 million possible citations really exist but there also are decisions among the large group of weakly connected cases that are both well cited and interdependent. These decisions are the hubs of the network and, not surprisingly, Chandler claims that this main core substantially concerns "rights of free speech and association under the American constitution." [2005, 20; and Pagallo 2007, 205] More particularly, we have got 122 nodes each with 28 or more links to the other cases of the main core so that the density is more than 500 times greater than the density of the network as a whole. By grasping these cases as hubs of a small world-network, it becomes then easy to understand why any First Amendment decision would reverberate more readily through law than a decision made in any other field. In a nutshell, hubs offer the common connections mediating the short path lengths between other nodes in the network.

These results were (partially) confirmed in June 2005 by Thomas Fowler and Sangick Jeon who presented the network of 30,288 U.S. Supreme Court majority opinions from 1754 to 2002 (actually, they also accounted for the decisions of the Supreme Court of Pennsylvania contained in the first volume of the U.S. Supreme Court Reporter). Again, each case is considered as a vertex or a node of the network and each citation as an arc or a link. The total number of links to and from the node represents its degree (in and out). The overall result is a list of all of the cases that are connected together by 220,500 citations according to the power-law distribution of a small world-network. From this viewpoint, it is not only possible to highlight the "good hubs" (well cited) as well as the "good authorities" (most interdependent nodes) of the network. We are also able to follow the rise and fall of a precedent's importance in a continuously evolving legal system such as the U.S. common law. While the most authoritative cases before the American civil war involved freedom of contract, namely the contract clause, after the war and until the end of the 1930s with the New Deal, the main core became balance of power in order to regulate commercial issues in a federal system. Whereas this perspective confirms Chandler's conclusions – in that the contemporary main core of the U.S. Supreme Court jurisprudence is given by rights of free speech as Justices shifted their focus towards civil liberties – Fowler and Jeon also back up our network perspective. "In particular, the power-law tail in the degree distribution of inward and outward citations in the precedent network suggests that there is something systematic about the evolution of law that mimics the evolution of other network phenomena." [2005, 33]

## 3 LEGAL TOPOLOGIES

There are two main applications of the "small world"-topological approach to complex legal systems.

First, it is possible to deepen the analysis of some fields of artificial intelligence and the law concerned with the study of legal reasoning and argumentation through computational methods as in the example of case-based legal reasoning, of knowledge discovery in legal databases, or of legal information retrieval [see Pagallo, Ruffo 2007a]. For instance, we can determine which are the most connected cases, i.e., the hubs of the network in the double perspective of good hubs that cite many good authorities and good authorities cited by many good hubs. By quantifying the grade by which some decisions accelerate the transmission of ideas in the legal network, we obtain two kinds of precedents, i.e., well founded in law and influential. From this further standpoint, electronic maps permit not only to determine and quantify the authority cases of the legal network while comparing them with expert rankings, in that the small world paradigm also fits several other realms of law. We have electronic maps built on the articles of the U.S. federal commercial code, on the network of legal decision-making both at Capitol Hill and in Stockholm, but even on scholarly publications [as shown by Th. Smith 2005].

However, there is a second possible realm of applications due to our network approach: Indeed, after Barabási's work on the Internet and the Web, several scientific papers have demonstrated the existence of small world-patterns and power laws-distribution of information that characterize any peer-to-peer system at different levels. For example, there is a significant evidence of spontaneous clustering of users by content distribution in, say, both Gnutella and Kazaa systems as shown by Iamnitchi et al. [2004] and Ruffo et al. [2008]. Whereas different models can be used in order to detect this phenomenon like "data-sharing graphs" or "affinity networks," what is striking is the fact that the topology remains the same of such complex networks as the U.S. Supreme Court jurisprudence or the Web studied by Barabási: notwithstanding the nature of the system or its peculiar constituents, the probability that a vertex in a complex network is connected to other vertices decays according to a power law. So, it is possible to exploit these topological properties of the network in order to get some interesting applications on recommender systems, digital privacy, and copyright.

On one hand, our work shows how hubs can be considered as vectors for exploiting all the very opportunities of this technology in order to protect personal data. As we illustrated in previous work [i.e., Pagallo, Ruffo 2007b], it is still Gnutella's small world-features that suggest to adopt a decentralized recommendation scheme based on spontaneous affinities. By exploiting partnership degree and users relationships it is not necessary to get user profiles and users are not required to give feedbacks to a data collector entity. In other words, we do not need to trade off personal data for digital personalization and data protection on the Web, because there is a brilliant way to update Amitai Etzioni's [2004, 45] dichotomy between "liberalizing technologies" and public protective ones. On the other hand, it is not a mystery that hubs may be intended as simple targets in order to break these spontaneous networks and, hence, these new emerging digital communities. As a matter of fact, legal troubles of P2P systems prove, among other things, that the topology of these systems can be (and has been) exploited to develop new methods for attacking copyright infringements [as discussed in Pagallo 2008a-b].

Yet, the panoply of all possible applications, pro or contra privacy, pro or contra copyright, does not imply that technology should be considered as something "neutral," i.e., a simply means to obtain whatsoever end. Rather, it is important to remark the mutual interaction through which technology reshapes both legal concepts and their own environmental framework, while political decisions influence or attempt to determine possible developments of technology. After all, legal troubles of P2P systems with both copyright and privacy interests illustrate some peculiarities of the U.S. legal system as well as some key differences between U.S.- and EU-law determined by the same technology.

## 4 LEGAL ONTOLOGIES

One of the most fruitful ways to produce formal descriptions of knowledge is currently the development of "structured dictionaries" and research efforts on formal ontologies. For example, Gianmaria Ajani and his research group have shown many interesting applications, including a tool based on an ontology designed to recover legal information and build conceptual dictionaries. Avoiding as much as possible the polysemy of legal terms as well as terminological and conceptual faux-amis, an experiment has been conducted on taxonomy, where "terms are supported by a clearly identified concept inserted in a knowledge structure of a text and related to the other concepts that belong to such a structure." [see Ajani, Ebers, 2005] From a technical viewpoint, it is quite interesting to stress the "bottom up fashion" starting with legal terms defined by scholars. Actually, a traditional "top-down approach works well for the topmost level, where the basic conceptual primitives are precisely defined (concept, relation, role, qualia, processes, etc.), and the representation instruments are put at the disposal of those who build the ontology." [Lesmo et al. 2007] But, alas, a lot of problems arise when core ontology level is involved.

From our perspective, it would be very interesting to shed light on the topological aspects of these experiments by considering the semantic hubs of the network since it is not unlikely that further electronic reconstructions of legal networks as, say, in the case of ECJ jurisprudence or, in Italy, of the Constitutional Court would present some known features: high clustering coefficients, short diameter, main cores. So, we have three different applications.

First, at a national level, it would be easy to compare hubs within a same legal field of the network, in order to highlight both affinities and discrepancies of different national legal systems. While a semantic approach to legal ontologies has to select terms, a topological approach indicates which terms are fundamental for comparison, confronting, for example, authority scores with expert rankings.

Secondly, a linguistic field as legal science evolves, so that a topological approach would make it possible to follow the semantic evolution of the network. For instance, by taking into account the jurisprudence of the European Court of Justice, it might be very interesting to investigate how the main core has changed throughout the decades, measuring the authority score of, say, its first fundamental sentences in institutional cases (as in the paradigmatic example of the Kompetenz-Kompetenz issues).

Finally, it is still the topology of our system that obliges us to concentrate on the hubs of every taxonomy. Indeed, even in the case of linguistic systems it is highly probable to find determinate clustering coefficients that go along with a diameter shorter than that of regular networks. The set of nodes with highest degree could represent the main core of the taxonomy in so far as our empirical research would conduct us, with fair probability, to the shortest average distance-concepts with their specific grades of "betweenness." This means that these special hubs will probably be the nodes mostly tested by scholars in order to check the quality of the taxonomy. In scientific research, this outlook really offers a good way to think about our own "trials and errors."

## 5 THE PARADOX OF ELEGANCE: A CONCLUSION

We gave you a short account of the research on the small world paradigm and the topology of complex legal networks over the last ten years. Its outcomes are remarkable: we do know that this is the case of various complex social networks as the Internet and the Web, the jurisprudence of the U.S. Supreme Court and the network of legal decision making both at Capitol Hill and in Stockholm, the articles of the U.S. federal commercial code and the information shared via P2P systems on the Internet. These topological properties have been exploited in several different ways. Think of it as electronic maps in order to describe the semantic structure and evolution of legal networks, or as recommender systems to avoid overloading information and to guarantee privacy on the Web. In any case, it is always as if we were witnessing "something systematic about the evolution of law that mimics the evolution of other network phenomena." [see again Fowler, Jeon 2005, 33.] So, could we have found Leibniz's (and Gödel's) characteristica universalis via the small world-paradigm?

We do believe this is not the case for a couple of reasons.

First, there is a practical explanation that concerns Friedrich Hayek's classical distinction between cosmos and taxis, i.e., evolution vs. constructivism, spontaneous orders vs. human (political) planning. In a nutshell, Hayek's idea was that the informational complexity of cosmos should not be reduced by any taxis and, furthermore, orders spontaneously evolve from such informational complexity. This means that any constructivist approach as some multi-agent systems-perspectives suggest cannot cope with all the complexity of contemporary legal systems; moreover, even if you deal with cosmos complexity through a small world-topological viewpoint, it is certainly possible to describe and follow how complex social systems evolve but, alas, you cannot foresee or determine their evolution!

The second reason is theoretical and formalizes the very connection of information and complexity in contemporary legal systems. It may be summed up with Gregory Chaitin's thesis [2005, 125] on why you cannot prove that a program is "elegant," i.e., the smallest possible program that produces that particular set of theorems you have to deal with when analysing the informational complexity of the legal cosmos. In fact, if there were such an algorithm, we could always use it to find all the elegant programs!

This paradox of elegance, of course, recalls us the best epistemological output of the debate in the 1900s, i.e., "falsificationism."

In this context, however, we would like to stress another provisional morality: Both for practical and theoretical reasons, even (good) scientific results do not authorise to assume a close and self-referential perspective. This is why we believe work on social topology and small world distribution of information in (social and legal) complex systems should remain open and ready to promote the interchange of knowledge and methodologies among some of the most interesting and promising fields of contemporary scientific research.

## REFERENCES


[1] Ajani, Gianmaria and Ebers, Martin (eds) [2005]. Uniform Terminology for European Contract Law. Baden-Baden: Nomos Verlag.

[2] Ajani, Gianmaria et al. [2007]. The Multilanguage Complexity of European Law: Methodologies in Comparison, edited by G. Ajani, G. Peruginelli, G. Sartor, and D. Tiscornia, Florence: European Press Academic Publishing.

[3] Barabási, Albert-Lázslo [2002]. Linked: The New Science of Networks, New York: Perseus.

[4] Chaitin, Gregory [2005]. Meta Math! The Quest for Omega, New York: Pantheon.

[5] Etzioni, Amitai (2004). How Patriotic Is the Patriot Act? Freedom versus Security in the Age of Terrorism, New York-London, Routledge.

[6] Fowler, Thomas and Jeon, Sangick [2005]. The Authority of Supreme Court Precedent: A Network Analysis, on line at http://jhfowler.ucdavis.edu accessed 20.12.2008.

[7] Iamnitchi A., Ripeanu M., Foster I. [2004]. Small-world file-sharing communities. In Proc. of the 23rd Conference of the IEEE Communications Society (InfoCom 2004). Hong Kong.

[8] Kleinberg, J. [2000] The Small World phenomenon: An algorithmic perspective. In Proceedings of the 32nd ACM Symposium on Theory of Computing (STOC 2000), pp. 163-170.

[9] Lesmo, Leonardo – Boella, Guido – Mazzei, Alessandro – Rossi, Piercarlo [2007]. Multilingual conceptual dictionaries based on ontologies: analytical tools and case studies, in The Multilanguage Complexity of European Law, op. cit., pp. 179-194.

[10] Pagallo, Ugo [2007]. “Small world” Paradigm and Empirical Research in Legal Ontologies: a Topological Approach, in The Multilanguage Complexity of European Law, op. cit., pp. 195-210.

[11] Pagallo, Ugo [2008a]. Let Them Be Peers: The Future of P2P Systems and Their Impact on Contemporary Legal Networks, paper for the Infosoc meeting on “The Future of… Conference on Law and Technology” organized by the European University Institute, Fiesole, Florence, on Oct. 28-29, 2008, at http://www.one-lex.eu/futureof/ and to be published in a special edition of the European Journal of Legal Studies (EJLS).

[12] Pagallo, Ugo [2008b]. Ethics Among Peers: From Napster to Peppermint, and Beyond, paper for the 5th itAIS Conference on "Challenges and Changes: People, Organizations, Institutions and IT" organized by the Italian Association for Information Systems in Paris, France, on Dec. 13-14, 2008, at http://eventseer.net/e/7947/ and to be published by Springer, 2009.

[13] Pagallo, Ugo and Ruffo, Giancarlo [2007a]. P2P Systems in Legal Networks: Another “Small World” Case, in “Eleventh International Conference on Artificial Intelligence and Law”, Acm: Stanford, CA., pp. 287-288.

[14] Pagallo, Ugo and Ruffo, Giancarlo [2007b]. On the Growth of Collaborative and Competitive Networks: Opportunities and New Challenges, in Ethicomp Working Conference 2007, edited by S. Rogerson e H. Yang, Yunnan University, pp. 92-97.

[15] Ruffo, Giancarlo and Schifanella, Rossano [2008]. A Peer-to-Peer Recommender System Based on Spontaneous Affinities. in ACM Transactions on Internet Technology (TOIT), ACM Press, New York, NY, USA. (forthcoming).

[16] Smith, Thomas A. [2005]. The Web of Law, San Diego Legal Studies Research Paper No. 06-11. Available at SSRN: http://ssrn.com/abstract=642863 or DOI: 10.2139/ssrn.642863.

[17] Watts, Duncan and Strogatz, Steven [1998]. Collective Dynamics of “Small-World” Networks. Nature, 393, pp. 440-442.

# Modelling Network-Enabled C2 using Multiple Agents and Social Networks

**Tim Grant** [1][2]

**Abstract.** This paper describes work in progress in developing an architecture for Command & Control systems based on an empirical model of the military commander's thinking processes. It shows how social network constructs could be added to a multi-agent system to model network-enabled capabilities in a complex, real-world domain, using the events of September 11, 2001, as a case study. Areas for further research are identified.

## 1 INTRODUCTION

Military operations have changed dramatically in nature since the end of the Cold War in 1989. Nowadays, coalitions are formed from the forces of many nations ("combined operations") and multiple services ("joint operations"), allied with civilian organizations such as commercial suppliers, government departments, international and non-governmental organizations, and the media ("civil-military cooperation"). Instead of all-out warfare, operations may take the form of defence, diplomacy, and development (the "three D's"), often simultaneously.

**Figure 1. RNLA's Battlefield Management System.**

In military organizations, monitoring and control of operations is known as Command & Control (C2), defined as: "the exercise of authority and direction by a properly designated commander over assigned and attached forces in the accomplishment of the mission" [1]. C2 functions include "planning, directing, coordinating, and controlling forces and operations", and a C2 system consists of "an arrangement of personnel, equipment, communications, facilities, and procedures".

The dramatic changes in the nature of military operations, coupled with developments in Information and Communication Technologies (ICT), have required equally dramatic changes in military C2. Twenty years ago, communications in-theatre were by voice or morse code over VHF/UHF radio, with the positions of own and enemy forces being plotted on paper maps using pencils. Long distance communications were by telex or by voice over HF radio. Today, military vehicles are fitted with GPS receivers, and broadcast their position periodically by data messages over Combat Net Radio to the superior commander and to other vehicles from the same unit. A in-vehicle display (see Figure 1) shows the position of all vehicles in the unit, automatically plotted in real time on a map or satellite photograph of the area. The same display can be copied by satellite links to headquarters, possibly thousands of kilometres away, or by data links to friendly ships and aircraft. The same system can be used to disseminate operational orders from the commander to all vehicles in the unit, as well as providing email and chat functionality. The key benefit is increased C2 tempo and agility.

These changes have less visible organizational consequences. Linking all units, vehicles, ships, aircraft, and individual soldiers, sailors, and airmen to headquarters means that the entire military organization will soon be networked together. Instead of emphasizing the flow of formal communications (situation reports and operations orders) up and down the organizational hierarchy, peer-to-peer communications and collaborative processes are becoming increasingly important. Trials and small-scale operational applications have shown that there are several pitfalls, including information overload and the temptation for high-ranking officers to micro-manage operations. Military doctrine must change accordingly. The NATO term for the desired future state is Network-Enabled Capabilities (NEC)[3], defined as "the Alliance's cognitive and technical ability to federate the various components of the operational environment, from the strategic level (including NATO headquarters) down to the tactical levels, through a networking and information infrastructure (NII)" [2].

Modern C2 system architectures must reflect these changing needs. As the NATO definition shows, a NEC-era C2 system is seen as a federation of units networked to one another, lending itself to agent-based architectures. In 2005, the Netherlands Defence Academy initiated a Network-Enabled C2 Systems research project entitled "Beyond Situation Awareness" aimed at addressing the research question: "Is it feasible and advantageous in terms of operational agility to construct NEC-

[1] Faculty of Military Sciences, Netherlands Defence Academy, P.O. Box 90.002, Breda, Netherlands. Email: **tj.grant@nlda.nl**.
[2] Also visiting Professor, Department of Computer Science, University of Pretoria, South Africa. Email: **tgrant@cs.up.ac.za**.

[3] Known in the USA as Network-Centric Operations.

era C2 system architectures based on the human users' needs, functions, and thinking processes?"

The starting point for the project was US Air Force Colonel John Boyd's [3] Observe-Orient-Decide-Act (OODA) model of the military commander's decision-making process. Although it is an empirical – rather than scientific – process model, OODA is taught widely in military academies throughout NATO. It is the *de facto* standard way of modelling the military C2 process, in effect having been extensively peer-reviewed.

The first step in the project was to benchmark OODA against similar process models in the psychological and cybernetics literatures. This disclosed shortcomings [4] that were addressed by rationally reconstructing OODA [5]. Two Masters students from the University of Liverpool (UK) developed an initial test-bed based on the rationally reconstructed OODA (OODA-RR) model, demonstrating it using the domain of a computer intruder ("hacker") versus a system administrator [6] [7] [8].

As the HackSim domain lacks any organizational features[4], the next step is to apply the test-bed to a larger, more representative domain, such as the US civil-military Air Traffic Control (ATC) / air defence (AD) organization that attempted (unsuccessfully) to shoot down Al Qaeda's hijacked airliners on September 11, 2001. Analysis [9] shows that this domain includes examples of communication up, down, across, and between hierarchies. Moreover, the communication flow differs according to whether events unfolded as *should have* happened given the Standard Operating Procedures (SOPs) then in force, as they *actually did* happen, and as they *could have* happened if the ATC / AD organization had been linked using NEC. Hence, the 9/11 domain is a worthy test for social network techniques.

The purpose of this paper is to show how social network constructs could be added to an OODA-RR-based multi-agent system to model network-enabled capabilities in military Command & Control. Coming from the multi-agent systems community, this paper addresses the first SNAMAS'09 research question: how to use social network analysis results for developing multi-agent systems. The key contribution is as a demonstration of how a complex, real-world domain (namely military C2) can be translated into theoretical constructs in the social network field. A major limitation is that the paper describes work in progress that has not yet been tested by implementing and evaluating it.

The paper consists of seven sections. Section 2 outlines the organizational aspects of present-day military Command & Control (C2). Section 3 summarizes the transformation to Network-Enabled Capabilities (NEC). Section 4 describes the rationally reconstructed OODA (OODA-RR) model and the HackSim test-bed. Section 5 illustrates the issues using the events of September 11, 2001, as a case study. Section 6 shows how social network constructs could be added to OODA-RR to model NEC. Section 7 draws conclusions and identifies directions for future work.

## 2 MILITARY COMMAND & CONTROL

Organizational structure allows authority, responsibilities, and processes to be allocated to organizational entities such as individuals, work groups, (project) teams, departments, sites, branches, organizations, consortia, and coalitions. The predominant relationship that defines organizational structure is one of power: a superior has power to give instructions and assign tasks to his/her subordinates. Power is often expressed in terms of resources. These resources may be measured in terms of money, manpower, or equipment. For example, the military commander who is assigned a squadron of 12 fighter aircraft has more power than one who has a flight of four fighters.

There is a variety of organizational structures. An organization may have a charismatic, bureaucratic, cooperative, functional, matrix, flat, network, or virtual structure. Present-day military organizations are mature, complex, and large-scale and are invariably hierarchical. Since military forces have become highly specialized, they often exhibit strongly bureaucratic traits. Because military commanders may have to order their subordinates to put their lives at risk, there are strong legal requirements attached to the superior-subordinate relationship. Operational communication in a military organization is highly formalized, with subordinates transmitting situation reports ("sitreps") up the hierarchy to their superior, and superiors disseminating operation orders ("op-orders") down the hierarchy to their subordinates. The path that sitreps and op-orders follow up and down the organizational hierarchy is known as the "reporting chain". The formats for sitreps and op-orders, the terminology to be used, and the circumstances in which they should be created is strictly prescribed by doctrine.

The fundamental functionality that C2 systems provide is one of communication between elements of a military force. Since modern military forces are highly mobile and may be widely spread over the operational area, the primary communications gap that a C2 system must bridge is one of geographical separation. The communications infrastructure in-theatre must be wire-less, with wired infrastructures only being feasible in fixed locations such as (air)bases, harbours, and headquarters. At the same time, C2 systems must bridge communications gaps between levels in the organizational hierarchy, between specializations, and between different services (i.e. army, navy, and air force).

Superimposed on top of the communications functionality, C2 systems must support the decision-making process at each level in the organizational hierarchy. Decision making involves aggregating, interpreting, and combining the sitreps a commander receives from his/her subordinates into an overall picture of the operational situation. Psychological research in the field of situation awareness shows that building up accurate and timely awareness of the prevailing situation is necessary (but not sufficient) for the commander to make good decisions [10]. The operational picture is a key input for the sitrep that the commander must make to his/her own superior. More importantly, the operational picture forms the basis for selecting responses to the situation, for generating courses of action that will bring about the selected response, and for developing the operations orders that incisively describe the commander's intentions.

Boyd's OODA loop is the canonical model for describing the military decision-making process. As Boyd drew it (see Figure 2), OODA is a cyclic model of four processes. By implication, the OODA processes are possessed by an agent that interacts competitively with other such agents in the environment. Boyd emphasized the Orient process both graphically and in text. He

[4] Other than the adversarial relationship between the intruder and the system administrator.

described the Orient process in the following terms [11, underlining in original]: "Orientation is an interactive process of many-sided implicit cross-referencing projections, empathies, correlations, and rejections that is shaped by and shapes the interplay of genetic heritage, cultural tradition, previous experience, and unfolding circumstances. ... Orientation is the schwerpunkt. It shapes the way ... we observe, the way we decide, the way we act." By contrast, he did not detail the three other processes. We interpret these three processes as follows:

- Observe is the process of acquiring information about the environment by interacting with it, by sensing it, or receiving messages about it.
- Decide is the process of making a choice among hypotheses about the environmental situation and among the possible responses to that situation.
- Act is the process of testing the chosen hypothesis by interacting with the environment and of generating expectations about the environmental response for subsequent Orientation purposes.

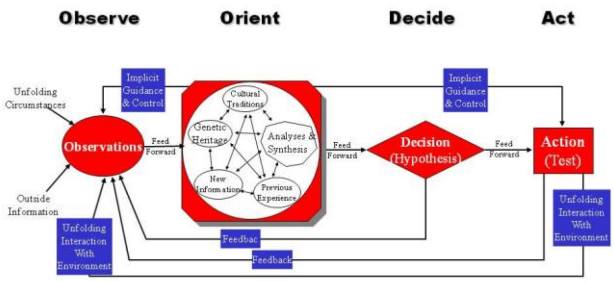


**Figure 2. Boyd's OODA loop.**

A unique feature of the OODA model is Boyd's emphasis on tempo, i.e. the decision cycle time. Boyd expressed this as follows [11]: "in order to win, we should operate at a faster tempo or rhythm than our adversaries or, better yet, get inside the adversary's Observation-Orientation-Decision-Act loop."

Several authors have pointed out that Boyd's OODA loop is purely reactive. Despite off-line, deliberative planning occupying a prominent place in military C2 doctrine and training, Boyd's OODA loop does not include a separate Plan process. Some authors attribute planning to the Orient process, others to the Decide process, and still others add a fifth Plan process to Boyd's original OODA loop.

## 3 NETWORK-ENABLED CAPABILITIES

In joint and combined operations, the flow of communications across the organizational hierarchy becomes essential to organizational performance. In coalitions and civil-military partnerships, communications must also flow between hierarchies. The pattern of communications flow adopts the form of a graph or network. In the 1990s, commercial organizations introduced flatter organizational structures supported by Internet technologies to exploit the power of informal communication flows (termed "eCommerce"). The US Department of Defense started to investigate analogous developments under the title of "Network-Centric Warfare" (NCW) towards the end of the 1990s. More recently, the term Network-Centric Operations (NCO) has gained currency in the USA, while NATO has adopted the term Network-Enabled Capabilities. The aim is the same: to increase combat power afforded by more robust and effective computer and communications networking.

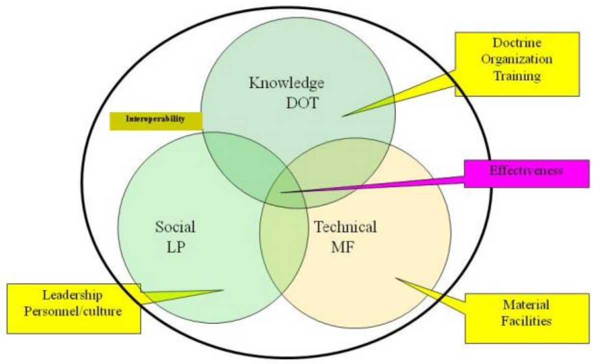


**Figure 3. Technical, cognitive, & social viewpoints.**

The original driver for NCW / NCO / NEC was technological. However, the community has accepted that, to be effective, any change in technology ("materiel") must be accompanied by simultaneous changes in doctrine, organization, training, materiel, leadership, personnel, and facilities: the "DOTMLPF factors". Moreover, networking can be seen from technical, cognitive (knowledge), and social viewpoints (see Figure 3).

There are four NEC tenets [12], often depicted as a value chain mapped onto the physical, information, cognitive, and social domains:

- A robustly networked force improves information sharing.
- Information sharing and collaboration enhance the quality of information and shared situation awareness.
- Shared situation awareness enables self-synchronization.
- These in turn increase mission effectiveness.

Organizational structure has been an important element of NEC research. Alberts and Hayes [12] argue that decision authority must be delegated from the centre of military hierarchies down to the edge of the organization. The US Naval Postgraduate School has studied such "edge" organizations extensively using computational experimentation techniques to bridge the gap between laboratory and field study [13]. Orr and Nissen [14] compare the Edge Organization against Mintzberg's [15] five archetypical organizational configurations: Simple Structure, Machine Bureaucracy, Professional Bureaucracy, Divisionalized Form, and Adhocracy. The Machine Bureaucracy corresponds to the traditional military organizational hierarchy. Orr and Nissen's results confirm that no single configuration is best for all circumstances. They elucidate seven specific performance measures that provide multi-dimensional insight into different aspects of organizational performance. The Edge Organization exhibits considerable agility, resisting performance degradation even under challenging conditions. They conclude that the military commander should vary the organizational structure according to the prevailing situation. One limitation of their research is that the communications flow pattern is dependent on the organizational structure.

Nissen [16] compares the communication structure of the hierarchical (Machine Bureaucracy) and edge organizations:

| **Hierarchy** | **Edge** |
|---|---|
| Vertical channels (stovepipes) | Horizontal network |
| "Wheel" structure | "Circle" structure |
| Few links | Many links |
| Low information exchange | High information exchange |
| Push | Post & smart pull |
| Meetings | Person-to-person |

**Table 1. Communication structure.**

## 4 OODA-RR TEST-BED

The rationally reconstructed OODA (OODA-RR) model was formalised using the SADT / IDEF0 notation and validated using a systematically generated set of use-cases. Boyd's original 4 processes were retained, with their names ending in "ing" to emphasize that they are continuous, concurrent processes. Two processes (Planning and Sensemaking) and two data structures (Goals and Prototypes) were added. Planning generates one or more new plans in real time for responding to each unexpected situation identified by Orienting. Deciding selects which plan will be enacted. Sensemaking is a Weickian [17] real-time learning process that creates new or modifies existing Prototypes from the experience gained in coping with novel situations. The agent boundaries and the environment through which agents interact were explicitly represented. More details on the rational reconstruction and on the resulting OODA-RR model (see Figure 4) can be found in [5].

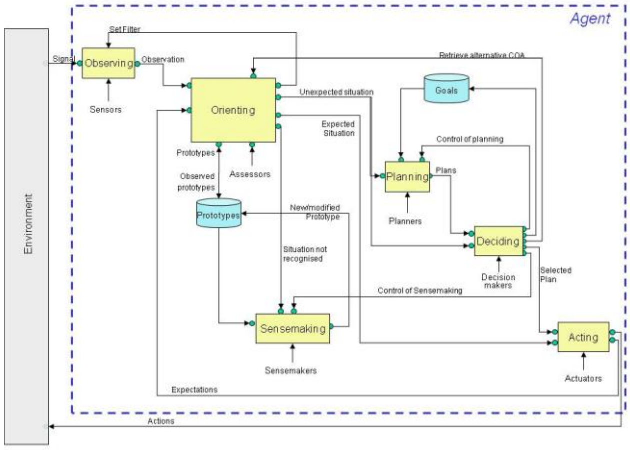


**Figure 4. OODA-RR model.**

To test the feasibility of the OODA-RR model, two Masters students designed and implemented the HackSim test-bed. To keep the work involved within the constraints of a Masters project, the Planning and Sensemaking processes were not implemented, and the user-interface and performance requirements were relaxed. The test-bed was designed in UML, with each OODA-RR process becoming an object-class in UML. Auxiliary object-classes were added. The test-bed was implemented in Java. Two instances of an OODA-RR agent were created, one representing the computer intruder and the other representing the system administrator. The intruder's target computer network was implemented as a set of instances of the Environment object-class. An initial pair of rule-sets representing the intruder's and system administrator's respective "doctrines" were generated by hand from one of the possible intruder-system administrator interaction scenarios [6]. The Masters students then iteratively extended the rule-sets. Their thesis [7] [8] include traces of several interaction scenarios.

The limitations of the implemented HackSim test-bed are:

- The Planning and Sensemaking processes have not yet been designed, implemented, and integrated with the object-classes representing the other OODA-RR processes.
- There is no explicit representation of organizational or other relationships between instances of the OODA-RR agent. The adversarial interaction between the intruder and system administrator instances was implicit in their respective rule-sets.

Finally, there is no representation in OODA-RR of collaborative processes, equivalent to the team-focused behaviours and outcomes in [18].

## 5 9/11 CASE STUDY

On September 11, 2001, the Federal Aviation Authority (FAA) was responsible for civil air traffic over the Continent US, and the North American Aerospace Defense Command (NORAD) was responsible for military air defence. The FAA's primary concerns were maintaining separation between aircraft and the nation-wide effects of severe weather or airport congestion. Operationally, the FAA was organized into 22 Air Route Traffic Control Centers, with an overall System Command Center (SCC) located in Herndon, Virginia. In the FAA headquarters, the Operations Center received notifications of incidents, including accidents and hijackings. NORAD's mission was to defend the airspace of Northern America from external attack by ballistic and cruise missiles or bombers. Such attacks were expected to come over the Atlantic or over the North Pole. NORAD was divided into three sectors, of which only one – the Northeast Air Defense Sector (NEADS) – was involved in the events that day. NEADS could call on two alert sites, each with one pair of ready fighters: two F-15s at Otis Air National Guard Base (ANGB) and two F-16s at Langley Air Force Base (AFB).

Prior to September 11, 2001, the FAA and NORAD had developed procedures for working together in the event of a hijacking. Military assistance would normally take the form of providing a fighter aircraft to escort the hijacked airliner, to report anything unusual, and to aid search and rescue in an emergency. Because the FAA and NORAD C2 systems were not interoperable, the relevant FAA control centre would relay tracking information to NORAD to vector the fighter to a position some 5 miles behind the hijacked airliner. Every attempt would be made to have the hijacked airliner switch to the hijack-in-progress transponder code so that it would become visible to NORAD. In short, the procedures assumed that:

- The airliner crew would continue to fly the airliner, albeit following the hijackers' directions.
- The hijacked airliner would be readily identifiable and would not attempt to disappear from radar.

- There would be time to handle the problem through the appropriate FAA and NORAD reporting chains, up to the President if it became necessary to shoot down the hijacked airliner.
- The hijacking would take the then-traditional form, i.e. it would not be a suicide mission.

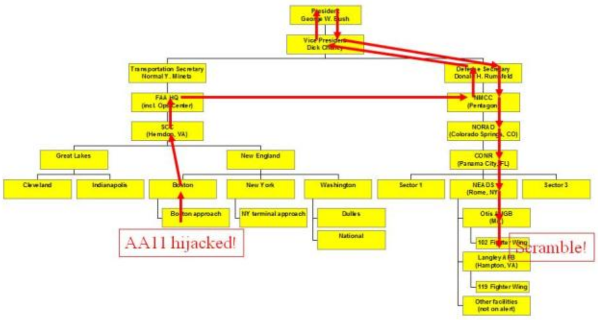


**Figure 5. 9/11: how reporting should have happened.**

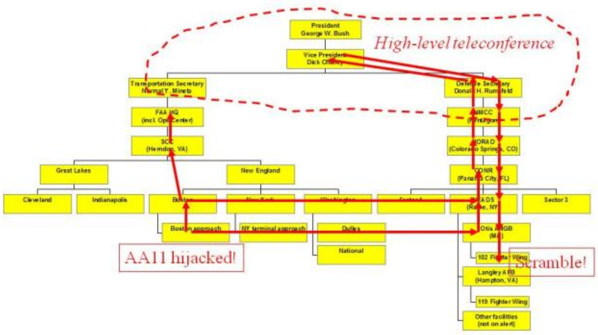


**Figure 6. 9/11: how reporting actually happened.**

The source material [19] [20] [21] [22] [23] [24] shows that the actual reporting chain (see Figure 6) on September 11, 2001, differed significantly from that called for in the prevailing procedures (Figure 5). Firstly, information was broadcast via the public media. Secondly, individuals within FAA's Boston approach and air route traffic control centres used their informal networks to inform military acquaintances in NEADS and at Otis ANGB, shortcutting the formal procedure [22]. This was fortuitous, because the hijack coordinator at FAA headquarters was on vacation and his deputy had not been trained in hijack procedures. Thirdly, military officials in NEADS and NORAD on their own initiative ordered fighters to take off, before they had obtained the required approval from their superiors [22]. The outcome was that the Otis fighters were airborne earlier than they would have been if the procedure had been followed meticulously.

If NEC had been available to the FAA and NORAD on September 11, 2001, all the players would have been networked to one another, enabling communications to take a direct line from "sensors" to "shooters". For the purposes of a thought experiment, a hijack network was assumed to have been superposed on the (unchanged) FAA and US military organizational hierarchies. As soon as FAA Boston approach informed the Boston air route traffic centre of the first hijack, this information would have been posted on the hijack network. It would then have been instantly available to all players involved, including NORAD, NEADS, Otis, and Langley (see Figure 7). While NORAD was giving approval for the airliners to be shot down, NEADS would already have been locating the fighter resources to do so. Appropriate recalculation of the declassified timeline [20] shows that the military air defence would have been notified between eight and 35 minutes earlier than actually happened. While this would not have changed the outcome for the first three hijacked airliners, the fighters would have been positioned well in time to shoot down United Airlines 93, had the passengers not stormed the cockpit, forcing the hijackers to crash short of their target.

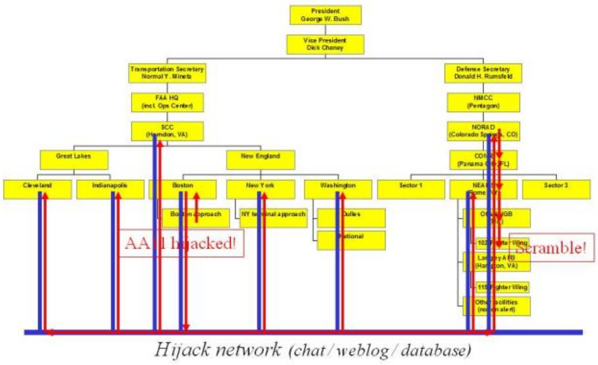


**Figure 7. 9/11: how it could have happened.**

This case study shows that the availability of a network linking civil ATC and military AD centres would have increased C2 tempo on September 11, 2001. This could well have altered the course of events if the passengers of United Airlines 93 had not forced the hijackers to crash short of their target. It is particularly noteworthy that the case study shows that the same ATC / AD organizational structure can support at least three different patterns of communication flow.

## 6 ADDING SOCIAL NETWORKS

The next step in the "Beyond Situation Awareness" research project is to apply the HackSim test-bed to the 9/11 case study. Each hijacked airliner and each of the civil ATC and military AD control centres would be modelled as an OODA-RR agent instance. To model the FAA and NORAD organizational hierarchies, the OODA-RR agent object-class would have to be given attributes for its superior agent and a set of subordinate agents. In addition, to model what actually happened and what could have happened with NEC, the OODA-RR agent object-class would also have to be given an attribute representing peer agents to be informed in the case of a hijacking.

In effect, two networks are superposed in the 9/11 case study: one representing the vertical, superior-subordinate reporting chains, and the other representing the horizontal peer-to-peer links. In all three cases (*should-have*, *actually-did*, and *could-have*) the FAA hijack coordinator would have had the Pentagon as a peer agent to be informed in the case of a hijacking. In the *actually-did* case, the agents representing the Boston approach and air route traffic control centres would also have had their Otis ANGB and NORAD acquaintances as peer agents to be informed in the case of a hijacking. In the *could-have* case, each of the FAA SCC, all FAA Air Route Traffic Control Centers, NORAD, NEADS, Otis ANGB, and Langley AFB would have had all the other control centres as peer agents.

In terms of social network constructs, the formal reporting chains running vertically in an organizational hierarchy could be represented as regular graphs or lattices [25]. The informal horizontal peer-to-peer links could be represented as small-world or scale-free networks.

## 7 CONCLUSIONS & FURTHER WORK

This paper describes work in progress in developing a C2 systems architecture based on the OODA-RR model of the military commander's thinking processes. It outlines some organizational aspects of present-day military C2, summarizes the transformation to NEC, and describes the HackSim test-bed implementing the OODA-RR agent model. The complex, real-world events on September 11, 2001, are used as a case study to show how social network constructs are needed to supplement a multi-agent system model.

Areas for further work include:

- Designing, implementing, and integrating OODA-RR's Planning and Sensemaking processes into the HackSim test-bed.
- Extending OODA-RR to include collaboration processes, and designing, implementing, and integrating this functionality into the HackSim test-bed.
- Implementing in the HackSim test-bed those attributes needed to represent agents' formal and informal social networks.
- Modelling the 9/11 case study in the HackSim so as to reproduce the *should-have*, the *actually-did*, and the *could-have* cases.

## REFERENCES


[1] US Department of Defense. Dictionary of Military and Associated Terms. Joint Publication 1-02, 12 April 2001, as amended through 12 July 2007 (2007).

[2] NATO. NATO Network-Enabled Capabilities Feasibility Study. Version 2.0, MCM-0032-2006 dated 19 April 2006 (2006).

[3] J.R. Boyd. The Essence of Winning and Losing. Unpublished lecture notes (1996).

[4] T.J. Grant & B.M. Kooter. Comparing OODA and Other Models as Operational-View C2 Architecture. Procs. 10th ICCRTS, Washington DC, USA. US DoD Command & Control Research Program, Washington DC, USA (2005).

[5] T.J. Grant. Unifying Planning and Control using an OODA-Based Architecture. Procs. SAICSIT 2005, White River, South Africa. University of Pretoria, South Africa (2005).

[6] T.J. Grant, H.S. Venter, and J.H.P. Eloff. Simulating Adversarial Interactions between Intruders and System Administrators using OODA-RR. Procs. SAICSIT 2007, Fish River Sun, South Africa. Nelson Mandela Metropolitan University, South Africa (2007).

[7] Maarten Dollenkamp. Examining OODA as an Architectural Basis for Intrusion Detection Systems. MSc thesis, University of Liverpool, UK, 25 December 2007 (2007).

[8] Michiel Dollenkamp. Research in the Feasibility of OODA as an Architectural Basis for Intrusion Detection Systems. MSc thesis, University of Liverpool, UK, 24 December 2007 (2007).

[9] T.J. Grant. Measuring the Potential Benefits of Network-Centric Warfare: 9/11 as case study. Procs. 11th ICCRTS, Cambridge, UK. US DoD Command & Control Research Program, Washington DC, USA (2006).

[10] M.R. Endsley. Theoretical Underpinnings of Situation Awareness. In: *Situation Awareness Analysis and Measurement*. M.R. Endsley and D.J. Garland (Eds.). LEA, Mahwah, NJ, USA (2000).

[11] J.R. Boyd. Organic Design for C2. Unpublished lecture notes (1987).

[12] D.S. Alberts and R.E. Hayes. *Power to the Edge: Command ... Control ... in the Information Age*. US DoD Command & Control Research Program, Washington DC, USA (2003).

[13] M.E. Nissen and R.R. Buettner. Computational Experimentation with the Virtual Design Team: Bridging the chasm between laboratory and field research in C2. Procs. 9th ICCRTS, Copenhagen, Denmark. US DoD Command & Control Research Program, Washington DC, USA (2004).

[14] R.J. Orr and M.E. Nissen. Hypothesis Testing of Edge Organizations: Simulating performance under industrial era and 21st century conditions. Procs. 11th ICCRTS, Cambridge, UK. US DoD Command & Control Research Program, Washington DC, USA (2006).

[15] H. Mintzberg. Structure in 5's: A synthesis of the research on organization design. *Management Science*, 26: 322-341 (1980).

[16] M.E. Nissen. Hypothesis Testing of Edge Organizations: Specifying computational C2 models for experimentation. Procs. 10th ICCRTS, Washington DC, USA. US DoD Command & Control Research Program, Washington DC, USA (2005).

[17] K.E. Weick. *Sensemaking in Organizations*. Sage, Thousand Oaks, CA, USA (1995).

[18] P. Essens, A. Vogelaar, J. Mylle, C. Blendell, C. Paris, S. Halpin, and J. Baranski. *Military Command Team Effectiveness: Model and instrument for assessment and improvement*. NATO RTO Technical Report AC/323(HFM-087)TP/59, April 2005 (2005).

[19] 9/11 Commission. The 9/11 Commission Report: Final report of the national commission on terrorist attacks on the United States. US Government Printing Office, Washington DC, USA (2004).

[20] 9/11 Commission. Staff monograph on the "Four Flights and Civil Aviation Security". US Government National Archives website, posted 12 September 2005, accessed 4 November 2005 (2005).

[21] Aviation Week & Space Technology. *Crisis at Herndon: 11 airplanes astray*. AWST, 155 (25): 96-99 (2001).

[22] Aviation Week & Space Technology. *Exercise Jump-starts Response to Attacks*. AWST, 156 (22): 48-52, 3 June 2002 (2002).

[23] Aviation Week & Space Technology. *NORAD and FAA Sharpen View Inside Borders*. AWST, 156 (23): 50-52, 10 June 2002 (2002).

[24] Aviation Week & Space Technology. *F-16 Pilots Considered Ramming Flight 93*. AWST, 157 (11): 71-74 (2002).

[25] M.E. Gaston and M. desJardins. Social Network Structures and their Impact on Multi-Agent System Dynamics. Procs. 18th FLAIRS, Clearwater, FL, USA (2005).

# Four Ways to Change Coalitions: Agents, Dependencies, Norms and Internal Dynamics

**Leendert van der Torre**[1] and **Serena Villata**[2]

**Abstract.** We introduce a formal social network approach to distinguish four ways in which coalitions change. First, the agents in the network change. Second, dependencies among the agents change, for example due to addition or removal of powers and goals of the agents. Third, norms can introduce normative dependencies for obligations and prohibitions. Fourth, coalitions can change due to internal processes. We propose a number of stability measures to identify each one of the four proposed sources of coalitions' dynamics and the consequences they induce on the stability of coalitions.

## 1 Introduction

Coalitions play a central role in social reasoning, and thus various theories have been used and developed in multiagent systems. For example, coalitional game theory has been adopted from economics and extended for multiagent systems [8, 9], and social networks have been adopted from social sciences and modified to represent dependence networks among agents [10, 6, 7]. These theories differ in various ways. For example, in the former, potential coalitions may be seen as sets of agents while in the latter, dependence networks can be seen as criteria for proposing/accepting to form coalitions [10], or *potential* coalitions are viewed as sets of dependencies (the dependencies represent the contract of the potential coalition) [7]. Moreover, in the former various notions of stability are defined, whereas in the latter they are not. In this paper, we address the question how to distinguish and model the different reasons behind the change of coalitions.

Possible reasons behind these changes are due to operations of addition and removal of the components of our model such as agents, dependencies among agents, normative dependencies concerning normative goals and powers. More precisely, how do we measure the evolution and the changes of a coalition over time in terms of:

**changes of the agents and dependencies.** We distinguish two kinds of uses for dependence networks: global use in software engineering where the designer models all stakeholders [4], and social simulation where no such assumption is made [10]. In the former, game theory can be used for reasoning about social interaction, in the latter simulation methods are used. We follow the tradition of TROPOS [4], as formalized by Sauro [7] and close to qualitative game theories developed by Wooldridge et al. [1], not the latter [10].

**changes of the dependencies related to norms.** Norms are used for the dynamics of dependence networks, which explained why they have not been considered thus far in the static dependence networks [11]. A norm analytically implies that agents (intend to) execute them, and therefore leads to dependencies among agents just like the original goal-based dependencies studied by Sichman and Conte [11]. More precisely, norms generate normative goals in the dependence networks, and these normative goals, i.e., obligations, are treated just like goals derived from the agent's desires. The coalitions which may emerge depend on the dependencies among the agents, so since norms change the dependencies among agents, they also change the coalitions which will emerge.

**internal dynamics.** Changes of the coalition itself in terms of goal-based and norm-based dependencies composing the coalition, e.g., an agent is excluded from a coalition because of a malicious behaviour.

We call the last kind of change *internal dynamics* to distinguish it from the other dynamics related to the addition or deletion of agents or goal-based and norm-based dependencies. They represent the case in which the network remains the same, involving the same agents and dependencies, but the composition of the coalition changes, including new dependencies or excluding the old ones. A simple and intuitive common sense example of the above presented changes can be the next one. Consider a soccer team as a coalition. It can change because new players come in, or players retire. It can change, because agents acquire new abilities or loose abilities, e.g., they loose their form, they break a leg, and so on, or get new goals, e.g., they want to play in the national team. Concerning norms, there can be the obligation set by the trainer for a player to play in the left wing position. Concerning internal dynamics, there may be a malicious behavior of a player, e.g., he gets too many red cards since he is too aggressive and he is no longer allowed to play. In the paper, we explain the changes using a grid-based running example.

From the multiagent systems field, we use the normative multiagent paradigm while from social network theory we take the idea of defining graph theoretic measures. Concerning measures, we define measures associated to the number of agents and the number of goal-based dependencies present in each time instant, counting the number of norm-based dependencies in each time instant and counting the changes in the dependencies composing coalitions. Our measures are unified in an average measure returning coalitions' stability depending on the differences between values associated to consecutive time instants.

In this paper, we do not give a formal ontology but we define indications of the possible changes of coalitions. Moreover, we do not perform any simulation as in Carley's dynamic networks analysis [5]. This paper is organized as follows. Section 2 presents a grid-based scenario. Section 3 and 4 present the key concepts of our metamodel and the three coalitions' changes in detail. Related work and conclusions end the paper.

[1] University of Luxembourg, Luxembourg
[2] University of Turin, Italy, email:villata@di.unito.it



The Society for the Study of Artificial Intelligence and Simulation of Behaviour Convention 6th - 9th April 2009, Edinburgh, Scotland

## 2 Changing coalitions in a GRID scenario

We use the following example of a coalition in a grid environment. Inside a virtual organization (VO), local coalitions may be formed in order to cooperate to achieve shared goals such as, i.e., computations and storage of satellites' data. We depict a section of the VO composed by five nodes, as in Figure 1.a, following the legend of Figure 3. The VO is composed by four nodes connected to each other by dependencies based both on goals and on norms and nodes $a$, $b$ and $c$ form a local coalition. Considering goal-based dependencies, node $b$ depends on node $a$ to save the file *satellite.jpg*, node $c$ depends on node $b$ to save the file *satellite.mpeg* and node $c$ depends on node $d$ to run the file *results.mat*, since they are not able to perform their goals alone. Considering norm-based dependencies, instead, node $a$ depends on node $c$ to have the permission to open the file *data-June.mat* while node $c$ is obliged to give to node $b$ the results of the running of file *mining.mat*.

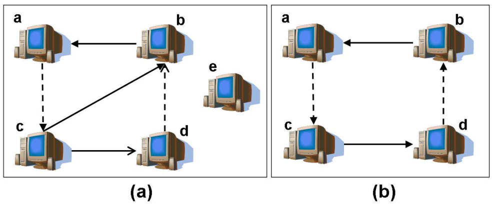


**Figure 1.** Grid network:a-Coalition $\{a, b, c\}$;b-Coalition $\{a, b, c, d\}$.

The first kind of change of coalitions in the grid scenario follows directly from the grid metaphor. Computers can be connected to the grid like electrical machines can be connected to the power net. So the computers connected to the grid changes frequently, e.g., node $e$. If they do so, then also the coalition changes. How frequently they change is our first measure.

The second kind of change concerns goal-based dependencies. Node $b$ fulfilled the goal of node $c$ to save the file *satellite.mpeg*. This dependency does not hold anymore and it is deleted, as shown in Figure 1.b. This deletion of dependencies changes the structure of the local coalition because of now the reciprocity involves also node $d$ inside the system. The deletion, as the addition, of a goal-based dependency may cause a change in the coalitions composed by these dependencies.

The third kind of change is related with security. A node has a number of private information, e.g., a unique access to its pc. If another node has the necessity to access to it, it has to ask the first node the permission, e.g., a login and a password, as in the norm-based dependency among nodes $a$ and $c$. Obligations, instead, are due to particular services provided by the nodes. The obligation is represented as a dependency, as in the case of the norm-based dependency among nodes $d$ and $b$, and it is removed if the obligation is no more active in the system. Figure 2.a shows the introduction of a norm-based dependency representing the obligation for node $b$ to give the access to file *finalres.txt* to node $a$.

The fourth kind of change, internal changes of coalitions, represents changes in the composition of the coalition because of internal reasons. In Grid networks, malicious behaviours can be recognized, e.g., in case of attacks or for not properly following the protocol, and malicious nodes can be excluded from further interactions with the other nodes, as shown in Figure 2.b.

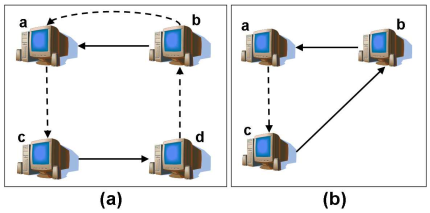


**Figure 2.** Grid network:a-Coalition $\{a, b, c, d\}$;b-Coalition $\{a, b, c\}$.

## 3 The model

Our modeling approach aims to provide a design methodology both for multiagent systems and social systems, based on the normative multiagent paradigm. The main concepts and relationships composing our models are agents, institution, roles, dependency, norm, power, goals. For more details on the conceptual metamodel, see [2, 3]. We divide our conceptual metamodel in three submodels: the agent model, the institutional model, and the role assignment model. Such a decomposition is common in organizational theory, because the organization can be designed without having to take into account the agents that will play a role in it. Likewise, agents can be developed without knowing in advance in which institution they will play a role. The notion of agent and all its features as goals, capabilities, facts are used in the conceptual modeling. Moreover, we add to these notions those related to the institution [3] such as the notion of role and all its institutional goals, capabilities and facts. Both these notions, combined in the combined view, are used in the conceptual modeling and to each agents it is possible to assign different roles depending on the organization in which the agent is inserted.

A coalition can be defined using the modeling technique of dependence networks, based on the idea that to be part of a coalition, every agent has to contribute something, and has to get something out of it. Roughly, a coalition can be formed when there is a cycle of dependencies (the definition of coalitions is more complicated due to the fact that an agent can depend on a set of agents, as we will see below). Since the processes involving coalitions dynamics are complex and costly social behaviors, agents have to maintain the stability of their own coalition, paying attention to the possible actions that can be performed by the other agents to strategically increase their profit, mining the position of the agents inside the coalition or, even worse, destroying the coalition itself.

### 3.1 The model definition

In this section, we present our model as a tuple composed by the concepts of agents, goals, norms, time. These components are linked by the relationship of dependency, which characterizes our dependency modeling activity. Our model can be represented as follows:

**Definition 1** $\langle A, G, N, T, D, D \subseteq A \times A \times G, T \rightarrow 2^A, T \rightarrow 2^D, N \rightarrow 2^D, C \subseteq 2^D, N \subseteq C\rangle$ *consists in a set of agents* $A$*, a set of goals* $G$*, a set of norms* $N$*, a set of time instants* $T$ *and a set of dependencies* $D$*. Every time instant is related to the set of agents and to the set of dependencies* $D$ *present in the system in that instant. Norms are represented as a subset of dependencies. A coalition is represented as a set of dependencies and a subset of the dependencies composing a coalition can be represented by norms.*

In this model, a coalition can be represented by a set of dependencies, represented by $C(a, B, G)$ where $a$ is an agent, $B$ is a set of agents and $G$ is a set of goals. Intuitively, the coalition agrees that for each $C(a, B, G)$ part of the coalition, the set of agents $B$ will see to the goal $G$ of agent $a$. Otherwise, the set of agents $B$ may be removed from the coalition or be sanctioned.

**Example 1** *Given a set of agents* $A = \{n_1, n_2, n_3, n_4, n_5\}$, *a set of goals* $G = \{g_1, g_2, g_3, g_4, g_5\}$, *a set of norms* $N = \{nr_1, nr_2\}$, *a set of time instants* $T = \{t_1, t_2, t_3\}$ *(these instants are associated to instants* $t_1$, $t_3$ *and* $t_5$ *of Figure 5), we have the following assignments showing what agents are present in each time instant* $(t_1, \{n_1, n_2, n_3, n_4, n_5\})$, $(t_2, \{n_1, n_2, n_3, n_4\})$, $(t_3, \{n_1, n_2, n_3, n_4\})$, *what dependencies are present in each time instant* $(t_1, \{(n_1, n_3, g_1), (n_3, n_4, g_2), (n_4, n_2, g_3), (n_2, n_1, g_4), (n_3, n_2, g_5)\})$, $(t_2, \{(n_1, n_3, g_1), (n_3, n_4, g_2), (n_4, n_2, g_3), (n_2, n_1, g_4)\})$, $(t_3, \{(n_1, n_3, g_1), (n_2, n_1, g_4), (n_3, n_2, g_5)\})$, *what dependencies are norm-based ones* $(n_1, \{(n_1, n_3, g_1)\})$, $(n_2, \{(n_4, n_2, g_3)\})$ *and how are coalitions composed* $C_1 = \{(n_1, n_3, g_1), (n_3, n_4, g_2), (n_4, n_2, g_3), (n_2, n_1, g_4)\}$, $C_2 = \{(n_1, n_3, g_1), (n_2, n_1, g_4), (n_3, n_2, g_5)\}$.

In a multiagent system, since an agent is put into a system that involves also other agents, he can be supported by the others to achieve his own goals if he is not able to do them alone. This leads to the concept of power representing the capability of a group of agents (possibly composed only by one agent) to achieve some goals (theirs or of other agents) performing some actions without the possibility to be obstructed. The power of a group of agents is defined as follows:

**Definition 2 (Agents' power)** $\langle A, G, power : 2^A \rightarrow 2^{2^G} \rangle$ *where* $A$ *is a set of agents,* $G$ *is a set of goals. The function* $power$ *relates with each set* $S \subseteq A$ *of agents the sets of goals* $G^1_S, \ldots, G^m_S$ *they can achieve.*

**Example 2** *Given a set of agents* $A = \{n_1, n_2, n_3, n_4, n_5, n_6\}$ *and a set of goals* $G = \{g_1, g_2, g_3, g_4, g_5, g_6\}$, *the function power relates each agent with the set of goals it can achieve:* $power(n_1) = \{\{g_5\}\}$, $power(n_2) = \{\{g_1\}\}$, $power(n_3) = \{\{g_2\}, \{g_6\}\}$, $power(n_4) = \{\emptyset\}$, $power(n_5) = \{\{g_4\}\}$, $power(n_6) = \{\{g_3\}\}$.

Definitions 1 and 2 have the aim to explain how social dependence networks can be seen as multiagent systems. The notion of power is relevant for our methodology since it represents the social basis for the development of our model based on the methodology of dependence networks as developed by Conte and Sichman [11]. In this model, an agent is described by a set of prioritized goals, and there is a global dependence relation that explicates how an agent depends on other agents for fulfilling its goals. For example, *dep*$(\{a, b\}, \{c, d\}) = \{\{g_1, g_2\}, \{g_3\}\}$ expresses that the set of agents $\{a, b\}$ depends on the set of agents $\{c, d\}$ to see to their goals $\{g_1, g_2\}$ or $\{g_3\}$. A dependence network is defined as follows:

**Definition 3 (Dependence Networks (DN))** *A dependence network is a tuple* $\langle A, G, dep, \geq \rangle$ *where:*

- $A$ *is a set of agents and* $G$ *is a set of goals;*
- *dep* $: 2^A \times 2^A \rightarrow 2^{2^G}$ *is a function that relates with each pair of sets of agents all the sets of goals on which the first depends on the second.*
- $\geq$: $A \rightarrow 2^G \times 2^G$ *is for each agent a total pre-order on goals which occur in his dependencies:* $G_1 \geq (a)G_2$ *implies that* $\exists B, C \subseteq A$ *such that* $a \in B$ *and* $G_1, G_2 \in depend(B, C)$.

The *dependency modeling* represents our modeling activity consisting in the identification of the dependencies among the agents. Our *dependency modeling* is represented as a directed labeled graph whose nodes are instances of the metaclasses of the metamodel, e.g., agents, goals, and whose arcs are instances of the metaclasses representing relationships between them such as goal-based dependency and norm-based dependency. A graphical representation of the model obtained following this modeling activity is depicted in the legend of Figure 3. The *dependency modeling* describes the agents (white circles), the dependency among agents (one arrowed line connecting two agents with the eventual addition of a label representing the goal on which there is the dependency), the norm-based dependency among agents (one arrowed striped line connecting two agents with the eventual addition of a label representing the goal on which there is the dependency), the goal-based or norm-based dependency composing a coalition (one arrowed [striped] line with plain arrow) and the goal-based or norm-based dependency not composing a coalition (one arrowed [striped] line with open arrow). Open and closed arrows are used to provide an immediate graphical representation of coalitions.

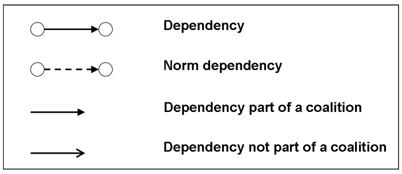


**Figure 3.** Legend of the graphical representation.

Example 3 presents the dependence network arising from the power relations of Example 2 with all dependencies belonging to a single coalition.

**Example 3** *Considering a Grid composed by six nodes, we can imagine to view each node as an agent and we can form the following dependence network:*

1. *Agents* $A = \{n_1, n_2, n_3, n_4, n_5, n_6\}$ *and Goals* $G = \{g_1, g_2, g_3, g_4, g_5, g_6\}$;
2. *dep*$(\{n_1\}, \{n_2\}) = \{\{g_1\}\}$*: agent* $n_1$ *depends on agent* $n_2$ *to achieve the goal* $\{g_1\}$*: to save the file comp.log;*
   *dep*$(\{n_2\}, \{n_3\}) = \{\{g_2\}\}$*: agent* $n_2$ *depends on agent* $n_3$ *to achieve the goal* $\{g_2\}$*: to run the file mining.mat;*
   *dep*$(\{n_3\}, \{n_1\}) = \{\{g_5\}\}$*: agent* $n_3$ *depends on agent* $n_1$ *to achieve the goal* $\{g_5\}$*: to save the file satellite.jpg;*
   *dep*$(\{n_4\}, \{n_6\}) = \{\{g_3\}\}$*: agent* $n_4$ *depends on agent* $n_6$ *to achieve the goal* $\{g_3\}$*: to run the file results.mat;*
   *dep*$(\{n_6\}, \{n_5\}) = \{\{g_4\}\}$*: agent* $n_6$ *depends on agent* $n_5$ *to achieve the goal* $\{g_4\}$*: to save the file satellite.mpeg;*
   *dep*$(\{n_5\}, \{n_3\}) = \{\{g_6\}\}$*: agent* $n_5$ *depends on agent* $n_3$ *to achieve the goal* $\{g_6\}$*: to have the authorization to open the file dataJune.mat;*

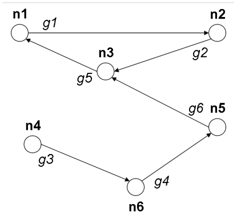


**Figure 4.** Dependence Network of Example 3.

## 4 Coalitions' Dynamics

In this section, we present a definition of coalition based on the structure of dependence network and how to use these different kinds of dependencies to model and measure coalitions' dynamics. In our model, a coalition is defined as follows:

**Definition 4 (Coalition)** *Let $A$ be a set of agents and $G$ be a set of goals. A coalition function is a partial function $C : A \times 2^A \times 2^G$ such that $\{a \mid C(a, B, G)\} = \{b \mid b \in B, C(a, B, G)\}$, the set of agents profiting from the coalition is the set of agents contributing to it. Let $\langle A, G, dep, \geq\rangle$ be a social dependence network, a coalition function $C$ is a coalition if $\exists a \in A, B \subseteq A, G' \subseteq G$ such that $C(a, B, G')$ implies $G' \in dep(a, B)$.*

As introduced before, we can model and measure coalitions' dynamics over time in terms of: changes of the agents and goal-based dependencies, changes of the dependencies related to norms and changes inside the coalition itself.

### 4.1 Agent and dependencies' changes

The first kind of change is due to agents entering or leaving the multiagent system we model or to the dependencies added or deleted depending on the fulfillment of the related goal or the presence of the power to fulfill this goal. In our model, we distinguish two different kinds of goals, achievement goals and maintenance goals. In contracts goals are typically achievement ones while, in game theoretical approaches, coalitions are typically concerned with maintenance goals. In this paper, we assume that goals are maintenance goals rather than achievement ones, which give us automatically a longer term and a more dynamic perspective to define the evolution of coalitions and thus their stability. Moreover, our model aims to distinguish and represent not only short term situations such as, for example, a virtual meeting on Second Life but also long term situations as, for example, the work of a particular department or office or, in the Grid scenario, the work of a virtual organization for e-Research.

We can define two measures associated to the number of agents and the number of goal-based dependencies present in each time instant. The first measure calculates the ratio between the number of agents added and removed in a particular time instant depending and the number of agents present at the previous time instant. The second measure calculates the ratio between the number of goal-based dependencies added and deleted in a particular time instant depending and the number of goal-based dependencies present at the previous time instant. The measures are defined as follows:

**Definition 5 (Agents and Dependencies Measures)** *Let $t_i$ be a time instant, $N_i^{Agent}$ is given by the number of agents entering the system $A_i^+$ and leaving the system $A_i^-$, depending on the total number of agents $A_{i-1}$ present at time instant $t_{i-1}$:*

$$N_i^{Agent} = \frac{\sum A_i^+}{\sum A_{i-1}} + \frac{\sum A_i^-}{\sum A_{i-1}}$$

*Let $t_i$ be a time instant, $N_i^{Dep}$ is given by the number of goal-based dependencies added to the network $D_i^+$ and deleted form the network $D_i^-$, depending on the total number of goal-based dependencies $D_{i-1}$ present at time instant $t_{i-1}$:*

$$N_i^{Dep} = \frac{\sum D_i^+}{\sum D_{i-1}} + \frac{\sum D_i^-}{\sum D_{i-1}}$$

**Example 4** *In Figure 5, we present the case of six time instants depicting the evolution of a system. In the first time instant, we have five agents and a coalition composed by agents $a$, $b$, $c$. Always in this first instant, there are two norm-based dependencies and three goal-based dependencies. The passage from the first instant $t_1$ to the second one shows the deletion of agent $e$ but all the dependencies remain the same. From instant $t_2$ to instant $t_3$, we can observe the deletion of the goal-based dependency connecting agents $c$ and $b$. This deletion can depend on a removal of the goal or of the power associated to the dependency or on the fulfillment of the goal. In the first case, if an agent has no more the power to fulfill the goal, the dependency has no reason to be maintained and the same thing happens if the agent has no more a particular goal. The coalition changes and it is formed by all the four agents. From instant $t_3$ to instant $t_4$, the situation changes back to the original configuration but the coalition is fixed. From instant $t_4$ to instant $t_5$, agent $d$ disappears and the coalition changes its actors and is composed by agents $a$, $b$ and $c$. From instant $t_5$ to instant $t_6$, the situation cames back to the situation of instant $t_4$. All these time instants can be represented as in Example 1. The measures vary as follows: [Agents] $t_1 : 0/5$, $t_2 : 1/4$, $t_3 : 0/4$, $t_4 : 0/4$, $t_5 : 1/3$, $t_6 : 1/4$; [Dependencies] $t_1 : 0/3$, $t_2 : 0/3$, $t_3 : 1/2$, $t_4 : 1/3$, $t_5 : 1/2$, $t_6 : 1/3$.*

### 4.2 Norms' changes

The second kind of change is due to norms and, in particular, to obligations. An obligation is a requirement which must be fulfilled to take some course of action, whether legal or moral. Normative reasoning is strictly related to norms' changes and the definition of a representation and a measure for them allows to do it. The norm sets a particular kind of dependency among two agents. This dependency can be deleted if the obligation is fulfilled or a new obligation can be inserted into the system to regulate its behaviour. In our model, we distinguish, represent and measure both short term contracts, e.g., a transaction on e-Bay such as an agreement carried out between separate entities involving the exchange of items of value as goods and money, and long term contracts, e.g., the marriage contract which hopefully lasts forever.

We can define a measure associated to the number of norm-based dependencies present in each time instant. This measure calculates the ratio between the number of norm-based dependencies added and deleted to each time instant depending and the total number of norm-based dependencies present in that time instant. The measure is defined as follows:

**Definition 6 (Norms Measure)** *Let $t_i$ be a time instant, $N_i^{Norm}$ is given by the number of norm-based dependencies added to the network $N_i^+$ and deleted form the network $N_i^-$, depending on the total*

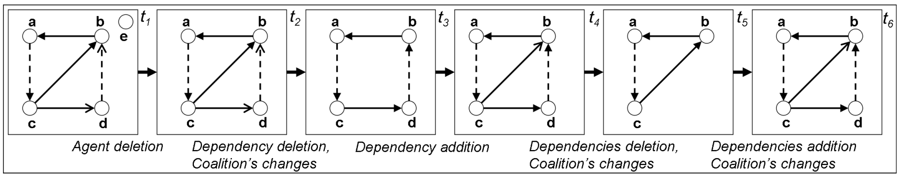


**Figure 5.** Agents and dependencies' change.

*number of norm-based dependencies $N_{i-1}$ present at time instant $t_{i-1}$:*

$$N_i^{Norm} = \frac{\sum N_i^+}{\sum N_{i-1}} + \frac{\sum N_i^-}{\sum N_{i-1}}$$

.

**Example 5** *In Figure 6, we model three time instants. In the first time instant $t_1$, we have a coalition formed by all the four agents, three goal-based dependencies and two norm-based dependencies. From time instant $t_1$ to time instant $t_2$, the norm-based dependency involving agents $d$ and $b$ is removed due to the removal of the normative goal or the removal of the associated power. From time instant $t_2$ to time instant $t_3$, a new norm-based dependency is set due to the insertion of a new normative goal or the associated normative power. The measure varies as follows: [Norms] $t_1 : 0/2$, $t_2 : 1/1$, $t_3 : 1/2$;*

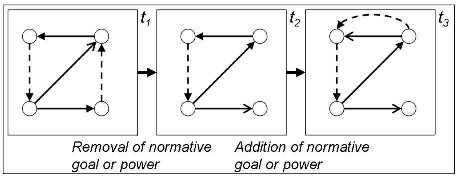


**Figure 6.** Norms' change.

## 4.3 Coalitions' changes

The third kind of change is related to changes inside the coalition itself, e.g., an agent is excluded from a coalition because of a malicious behaviour. This third kind of change is the only one related to the coalition itself and it has to represent and measure the changes in the composition of each coalition of the system. We define a measure which calculates the ratio between the number of the goal-based and norm-based dependencies composing the coalition in each time instant and the dependencies composing the coalition in the previous time instant, as follows:

**Definition 7 (Coalitions Measure)** *Let $t_i$ be a time instant, $N_i^{Coal}$ is given by the number of new norm-based and goal-based dependencies $D_i^+ \cup N_i^+$ belonging to a coalition $C_i$ added to the network and deleted form the network $(D_i^- \cup N_i^- \subseteq C_i)$ depending on the total number of norm-based and goal-based dependencies composing the coalition $(D_{i-1} \cup N_{i-1} \subseteq C_{i-1})$ at time instant $t_{i-1}$:*

$$N_i^{Coal} = \frac{\sum (D_i^+ \cup N_i^+ \subseteq C_i)}{\sum (D_{i-1} \cup N_{i-1} \subseteq C_{i-1})} + \frac{\sum (D_i^- \cup N_i^- \subseteq C_i)}{\sum (D_{i-1} \cup N_{i-1} \subseteq C_{i-1})}$$

.

**Example 6** *Consider the coalition depicted in time instant $t_1$ of Figure 7. The coalition is composed by agents $a$, $b$ and $c$. The passage from time instant $t_1$ to time instant $t_2$ sees the addition inside the coalition of agent $d$ due to the reciprocity-based principle of coalition formation. From time instant $t_2$ to time instant $t_3$, agent $d$ is excluded from the coalition, without any change in the number or type of the dependencies composing the coalition itself. This can depend, as said, on a malicious behaviour of the excluded agent. The measure varies as follows: [Coalitions] $t_1 : 0/3$, $t_2 : 1/4$, $t_3 : 3/3$;*

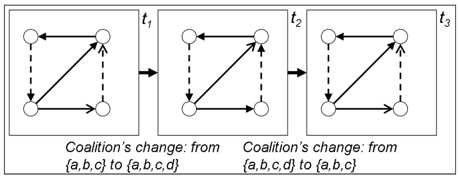


**Figure 7.** Coalitions' change.

The above measures are defined for one time moment only. We can unify these measures for a sequence of dependence networks associating to each time instant the average number of changes. We can define this measure as follows:

**Definition 8 (Changes Measures)** *Let $t_i$ be a time instant of a sequence of social dependence networks, the changes measure is given by the following formula:*

$$\frac{N_i^{Agent} + N_i^{Dep} + N_i^{Norm} + N_i^{Coal}}{4}$$

.

For example, considering instants $t_2$, $t_3$, $t_4$ and $t_5$ depicted in Figure 5, the average number of changes at each moment in time is: $t_2 : 1$; $t_3 : 3$, $t_4 : 1,3$; $t_5 : 11,3$. Thanks to this measure, we underline that the two time instants with the main changes in comparison with their previous time instant are $t_3$ and $t_5$, as can be supposed observing the relative figures. Always considering the measures of Figure 5 for time instants $t_4$, $t_5$ and $t_6$, it can be noted that in our measures the deletion of agents, norm-based and goal-based dependencies and dependencies composing coalitions increases the difference of the changes measure associated to two time instants in a row while the addition of these components causes a minor difference relatively to two consecutive time instants. This behaviour is due to the relation of our measure with the game theoretical approaches for defining stability: the stability is maintained in order to avoid the breaking off of the agents from the grand coalition and form their own group. Thus, the removal of agents and dependencies from coalitions is more relevant that their addition and this is represented in the changes measure.

We choose the simplest possible measures that capture the stability of the networks, because they represent all possible changes can be performed in the composition of coalitions and of the networks. When the average of the measures for a sequence of dependence networks presents a great difference in the values of two connected time instants, it underlines a lack of stability while when the average presents a small or inexistent difference between two connected time instants, the stability of the coalition and of the network in general is maintained. Moreover, the measures now only give a global indication of the stability of agents, dependencies, norms and coalitions. We could also measure whether changes in agents and dependencies coincides with changes in the coalition thanks to our four measures.

## 5 Related Work

In a multiagent perspective, a coalition can be viewed under two different representational frameworks. The first one regards cooperative game theory. Cooperative game theory studies those games in which players are able to make binding agreements with the aim to achieve a collective benefit. This approach is strictly related to the field of economics and various approaches of this kind have been presented in literature as, for example, the work of Shehory and Kraus [8]. The second perspective is based on the theory of the social power and dependence pioneered by Castelfranchi [6] as starting point and then developed in the context of coalition formation by Sichman [10] and Sauro [7]. This involves the development of a social reasoning mechanism that analyzes the possibility to profit from mutual-dependencies, e.g., two agents depend on each other for the satisfaction of a shared goal, or reciprocal-dependencies, e.g., two agents depend on each other for the satisfaction of two different goals. Both these two approaches present the following problems: they do not provide a modeling technique to represent coalitions' dynamics and to distinguish them. The comparison among these two approaches is summarized in Table 1.

| | **GAME THEORY** | **SOCIAL NETWORKS** |
|---|---|---|
| *Coalitions* | sets of agents | sets of dependencies |
| *Formation based on* | core | reciprocity |
| *Stability* | nucleous, Shapley value | none |

**Table 1.** Two models to represent coalitions.

## 6 Conclusions

We present a model to represent, at each time instant, the state of the system in terms of agents, goals, norms and the dependencies relating all these concepts. This model allows the distinction and measure of the possible coalitions' dynamics. In particular, we distinguish among three different kinds of coalitions' changes: changes based on addition or deletion of agents or goal-based dependencies, changes based on the addition or deletion of norm-based dependencies and changes on the internal structure of the coalition itself. It can be observed that with a more detailed model we could make more detailed and precise distinctions between the four kinds of changes. However, often we only have the given information, for example in systems' design, and we already would like to do kind of analysis on these models. This is precisely where graph-theoretical social network techniques are useful. We combine these techniques with the normative multiagent paradigm introducing in the networks norm-based dependencies. The strength of this combination consists in building a modeling technique able to represent in an intuitive way not only the inter-relationships among the actors of the system but also external constraints such as norms and, particularly, obligations, e.g., in our Grid scenario. The main difficulty of this approach consists in the creation of a common model without simplifying too much the two original frameworks.

Moreover, we introduce four measures aiming to measure these changes inside the networks to each time instant and an average measure to compute the stability of a sequence of dependence networks. Our model allows to measure coalitions' dynamics in terms of changing dependencies, agents and coalitions, distinguishing also among goal-based dependencies and norm-based ones. Using dependence networks as methodology to model a system advantages us from different points of view. First, they are abstract, thus they can be used for conceptual modeling, simulation, design and formal analysis. Second, they are used in high level design languages, like TROPOS [4], thus they can be used also in software implementation.

Concerning future work, we are working on a definition of coalitions' stability in our model, based on the presented measures, because of a lack of a definition of this notion in the field of social network theory. The notion of stability in our model can be identified intuitively in the absence of coalitions' changes we described but it is necessary to provide a formal definition of this notion and to associate it a measure able to represent it. Moreover, we start to simulate the use of our model and its associated measures in order to provide quantitative results based on our approach, similarly to social network theory approaches.

## REFERENCES

[1] Thomas Ågotnes, Wiebe van der Hoek, and Michael Wooldridge, 'Temporal qualitative coalitional games', in *AAMAS*, eds., Hideyuki Nakashima, Michael P. Wellman, Gerhard Weiss, and Peter Stone, pp. 177–184. ACM, (2006).

[2] G. Boella, L. van der Torre, and S. Villata, 'Social viewpoints for arguing about coalitions', in *PRIMA'08*, pp. 66–77. LNCS, Springer, (2008).

[3] Guido Boella, Leendert van der Torre, and Serena Villata, 'Changing institutional goals and beliefs of autonomous agents', in *PRIMA' 08*, pp. 78–85. LNCS, Springer, (2008).

[4] P. Bresciani, A. Perini, P. Giorgini, F. Giunchiglia, and J. Mylopoulos, 'Tropos: An agent-oriented software development methodology', *Autonomous Agents and Multi-Agent Systems Journal*, **8**, 203–236, (2004).

[5] K. M. Carley, 'Dynamic network analysis', in *Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers*, pp. 133–145, (2003).

[6] C. Castelfranchi, 'The micro-macro constitution of power', *Protosociology*, **18**, 208–269, (2003).

[7] L. Sauro, *Formalizing admissibility criteria in coalition formation among goal directed agents*, Ph.D. dissertation, University of Turin, 2005.

[8] O. Shehory and S. Kraus, 'Methods for task allocation via agent coalition formation', *Artificial Intelligence*, **101**, 165–200, (1998).

[9] Y. Shoham and K. Leyton-Brown, *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*, Cambridge University Press, 2008.

[10] J. S. Sichman, 'Depint: Dependence-based coalition formation in an open multi-agent scenario', *Artificial Societies and Social Simulation*, **1**(2), (1998).

[11] J. S. Sichman and R. Conte, 'Multi-agent dependence by dependence graphs', in *AAMAS'02*, pp. 483–490, (2002).

# Social evaluations and networks: a proposal for integration

**Francesca Giardini**[1]

**Abstract.** In social groups, individuals usually exchange information about their peers' actions, behaviors and attitudes. This exchange of information allows individuals to make more accurate and complete evaluations of other people; on the other hand, knowing facts about potential partners is pivotal to the establishment of new social links. In this work, a bottom-up approach will be applied to model the relationship between social networks and different kinds of social evaluations.

Results coming from Agent-Based social simulations will be reported and discussed in order to shed light on the emergence of complex social phenomena from the interaction among single agents' mental states and behaviours, namely the micro-macro link. More specifically, different types of social evaluations may affect the emergence of innovation and the network configuration of artificial firms working into an industrial cluster.

## 1 Cognition and the micro-macro link

Agents living in social systems are actually embedded in complex networks of relationships. Social groups can be described as networks with different sizes and configurations in which agents can occupy more or less peripheral positions, depending on the strength and the number of their links. In social groups, individuals usually exchange information about other agents, their actions, behaviours and attitudes, even if they have never met each other before. This exchange of information is essential for two reasons. On one hand, gathering information allows to make more accurate and complete evaluations of other people; on the other hand, knowing facts about potential partners is pivotal to the establishment of new social links and permits to enlarge the social group through the inclusion of far and distant nodes.

In a top-down perspective, social networks are described as complex systems of relationships among several nodes interconnected through diverse links in a variety of ways. This level of description applies to networks already given but does not allow to explain how networks emerge, evolve and change. If we want to understand how networks are created, an alternative approach is needed: following a bottom-up perspective, we claim that social networks are patterns of relationships among the goals of a given set of agents [8]. Heterogeneous agents, endowed with different beliefs, goals and resources are dependent upon each other to accomplish their tasks and achieve their goals. Basically, this means that agent x depends upon agent y, or upon its resources, to achieve its goals. Moving from this simple relationship between two agents it is possible to describe different macro-phenomena, their emergence and evolution.

[1] Institute of Cognitive Sciences and Technologies, ISTC CNR, Italy, email: francesca.giardini@istc.cnr.it

In fact, social complex phenomena, such as reputation, norms, and networks, emerge not only from agents' behaviours, but also derive from their cognitive representations and states. Cognition plays a crucial role in combining micro and macro levels: macro-social phenomena may emerge, unintentionally, from micro-elements and their interactions.

Aim of this work is to investigate how image and reputation, two different kinds of social evaluations, may affect the way in which agents are linked, i.e. their networks' configuration. Social evaluations and their transmission, namely gossip, are pervasive in human societies and can actually inform agents' networks of relationships. In order to unfold the connection between social evaluations and networks we need to model agents' minds and to understand why and about whom agents engage in gossip.

In what follows, main theories about evaluations and their transmission will be briefly reviewed with the purpose of highlighting the importance of social evaluations for human societies and of defining the framework of this research. Secondly, a cognitive account of social evaluations will be proposed in order to show how the macro level can emerge from agents' goals and beliefs. Some results coming from simulation experiments with artificial agents will be provided to support the model and to emphasize the importance of cognitive modeling to understand macro phenomena. Finally, future directions of research will be proposed.

### 1.1 Reputation and gossip: an introduction

The importance of social evaluations has been addressed by a variety of disciplines, with two distinct foci: reputation and gossip.

Reputation plays a fundamental role in social order, adding at the same time cohesiveness to social groups and allowing for distributed social control and sanctioning (plus a number of other functionalities, [6]). People use reputational information to make decisions about possible interactions, to evaluate candidate partners, to understand and predict their behaviours, and so on. The generation, transmission and manipulation of these beliefs contributes to regulate natural societies from the morning of mankind [12]). According to Frith and Frith [15], there are three ways to learn about other people: through direct experience, through observation and through *cultural information*. When the first two modalities are not available, agents turn to cultural information or, in our terms, to social evaluations in order to get some knowledge about potential partners and to predict their behaviours. This kind of information allows people to forecast, at least partially or approximately, what kind of social interaction they can expect and how that relationship could evolve, replacing personal experience in (a) identifying cheaters and isolating them, and in (b) easily finding trustful partners.



The Society for the Study of Artificial Intelligence and Simulation of Behaviour Convention 6th - 9th April 2009, Edinburgh, Scotland

Furthermore, in human societies reputation transmission facilitates the formation of groups [19]: gossipers share and transmit relevant social information about group members within the group, at the same time isolating out-group individuals. Besides, reputation contributes to stratification and social control, since it works as a tool for sanctioning deviant behaviours and for promoting, even through learning, those behaviours that are functional with respect to the groups goals and objectives. Reputation is also considered as a means for sustaining and promoting the diffusion of norms and norm conformity [28].

Social evaluations are crucial also in a developmental perspective, as demonstrated by Hamlin, Wynn and Bloom [20]. In their work with preverbal infants (6-, 9-, 10- and 12- month-old), they found that infants' social preferences are influenced by others' behaviors toward unrelated third parties, both in a choice paradigm and in a violation of expectation paradigm: when looking at social interaction events, infants preferred helpers and were independently inclined to avoid hinderers. The authors suggest that this early capacity for evaluations can be a biological adaptation evolved to allow humans to distinguish between cooperators and cheaters and then to engage in cooperative behaviors.

On the other hand, assuming an evolutionary perspective, theories of indirect reciprocity and costly signals show how cooperation in large groups can emerge when agents are endowed with or can build a reputation [22, 18]. As Alexander [1] pointed out, indirect reciprocity involves reputation and status, and results in everyone in the group continually being assessed and reassessed. According to this theory, large scale human cooperation can be explained in terms of conditional helping by individuals who want to uphold a reputation and so be included in future cooperation [23].

Another debated aspect regarding social evaluations is their transmission, i.e. gossip. The content of gossip, i.e. personal information about an absent third party, is quite uncontroversial, whereas the nature of the information, either evaluative or factual, is a matter of debate. Some authors require only the repetition of news about an absent third party to define a conversation as gossip, whereas other scholars consider necessary the presence of evaluative remarks to have a gossip talk. Baumeister, Zhang and Vohs [3] describe gossip as an exchange of useful information people can rely upon to face new situations and to behave properly when direct experience is impossible or too costly to acquire. Wert and Salovey [27] highlight the social character of gossip, and define it as an *evaluative talk* aimed at social comparison. Through gossip people can map the social environment and become conscious of their position whithin it. Rosnow [24] refers to the *secure standards of evidence* of a proposition, i.e. its veracity, to separate gossip from rumors, as also suggested by Noon and Delbridge [21]. The absence of the gossipee is one of the defining features of gossip, but it is also one of the reasons why gossip is condemned as immoral and is considered a malicious and harmful resource cowards use to criticize their peers [26].

Gossip serves many different social functions, both at the individual and at the group level. Gluckman [19] has been one of the pioneers in the study of gossip and scandal and one of the first to stress their positive virtues, among which their ability to maintain the unity, morals and values of social groups. Fine and Rosnow [13] draw attention to three main social functions served by gossip: information, influence and entertainment. Gossip is a valuable source of information about the community, its members, its norms, values and habits, but it is also useful to map the social environment and to make inoffensive comparisons. Regarding influence, this function is potentially prevalent when newcomers join the group or when there are conflicts between members. Foster [14] adds a fourth function: friendship, that refers both to dyadic relationships and to group bonded together by the sharing of norms and values. Other researchers consider gossiping as a means of knowledge: according to Ben-Ze'ev [5], gossip is a pleasurable way to gather information that is otherwise hard to obtain, but it also serves to satisfy the so-called *tribal need*, namely, the need to belong to the group and to be accepted by it.

Finally, one of the most debated issues in the literature about gossip is its being more or less purposeful. Some authors define gossip as idle-talk, giving a preminence to its relaxing and undemanding aspects [5, 26, 11], whereas others contend that gossip has a social purpose and people do not engage in it simply for entertainment, but mainly to achieve their goals [14, 13].

This brief review of main theories about reputation and gossip bear witness to the difficulty of analyzing this kind of complex social phenomena. A possible solution to this problem could be the assumption of a cognitive perspective to understand what happens in the agents' mind, and how cognitive underpinnings affect human behavior. On the other hand, we need to comprehend the emergent character of evaluation spreading and to take into account this complex dynamic: gossiping is a social activity rooted in individuals' beliefs, goals and preferences but it is also an emergent social phenomenon that, in turn, influences people's beliefs, behaviours and relationships.

## 2 Cognition at the intersection between evaluations and networks

Exchanging social information is a twofolded activity. It is a purposive act leading to specific and intended consequences, but it is also an emergent social phenomenon that has not been previously predicted and purposefully realized. In other words, its consequences can be unintended and functionally maintained, i.e. what emerged is independent of the agents' awareness and decisions, but it constrains their actions and determines their efficacy [7]. Once created and transmitted, social evaluations influence other agents' minds, changing their behaviors and goals. This can happen either intentionally, because the gossipee has the goal of influencing the receiver or even the target agent, or unintentionally, so that gossip effects are unintended but can equally give rise to functional phenomena. In this latter case, gossip has a twofold nature: it emerges from collective behaviours of information spreading but it needs to be represented into individuals' minds in order to work effectively.

Cognition works as a bond between the micro and the macro level, the so-called *micro-macro link* [8]: macro-social phenomena may emerge, unintentionally, from micro-elements and their interactions. In this view, social evaluations and networks derive from apparently autonomous social behaviours whose bases stand in the individual minds and in the relationships people are engaged in. In order to understand how social evaluations may affect social networks, we need to investigate how agents collect information about their peers, how they transform information in beliefs, and which goals drive agents actions, both when they are looking for information and when they pass on that news. Answering these questions is pivotal to the understanding of the relationship between evaluations and networks.

### 2.1 Modeling the micro-level

When dealing with social evaluations, a preliminary distinction is needed. Following Conte and Paolucci [9], we distinguish between *image* and *reputation*: the former refers to an evaluation regarding another agent's competence, behaviour, attitudes in which the source is clearly identified, whereas the latter designates an evaluation in which the source is missing. This difference is not inconsequential:

- if image is false, the addressee can punish the evaluator
- if image is false and negative, the target can punish the evaluator

Spreading inaccurate image, also involuntarily, exposes the evaluator to the risk of being reciprocated with false or inaccurate information or, even worse, of being ostracized. On the contrary, reputation is anonymous in itself, it circulates in the social network but its origin is unknown. Therefore, reputation spreading is easier than image transmission, and reputation is also more difficult to modify. Image, even when it is broadly transmitted, remains an evaluation coming from a specific and identified source, whereas reputation becomes an intangible mark floating within the social network. In cognitive terms, an image is a belief about a target coming from an identified source ("According to me, John is a nice guy"), while reputation necessitates a belief about the target, but also the belief that other agents believe that there is a specific evaluation about a given target.

Generally speaking, gossip has a triadic structure in which we can distinguish:

- A *gossiper*: an agent who has the goal to spread information. Informing another agent can be the only purpose of gossip, or it can be instrumental to other goals (influencing the receiver, punishing the gossipee, promoting gossipees image, enhancing groups feelings, ostracise someone, etc), more or less hidden.
- A *topic* or *third party*: an agent whose behaviours, attitudes, choices and emotions are the topic of the communication. The target belongs to the same group of gossiper and receiver and she is judged according to the group's rules and habits. Topic of the gossip talk can be an evaluation about an agent, not necessarily a report on her behaviours.
- A *receiver* (or more than one): one or more agents chosen from the gossiper to be told about the target. Receivers belong to the same social network, sharing the same knowledge and values of gossiper and gossipee. Choosing the receiver is pivotal to achieve gossiper's goals: the receiver can be the actual target of communication or she can serve as a vehicle to reach the intended target.

This triadic structure seems to be common to other forms of discourse, but the fact that the target is always an agent (and not an inanimate object) usually embedded in the same network of gossiper and receiver, is a peculiarity of evaluation transmission.

Relationships among the three roles above are neither symmetrical nor equal. First, there is an asymmetry of power: the gossipee finds herself in a position where she is helpless and vulnerable. The opposite is also true: DeSousa [11] challenges this view and considers gossip as a subversive form of power used by the weak, in this case the gossiper, to protect herself against more conventional powers. Looking at the relationship between the gossiper and the receiver, we find another asymmetry: the former influences the latter, providing the receiver with new knowledge that may change her goals, beliefs and intentions. Generally speaking, gossip is empowering to its participants because gives them access to knowledge

> Like that ill-fabled tree in the Garden of Eden, gossip promises us knowledge of good and evil. Like that same tree, it threatens us with expulsion if we are caught. The more vital the information exchanged through gossip, the more potentially damaging such gossip is both to those who are the topic of the conversation and to those who do the conversing ([2] p. 99).

Circulating social evaluations is a truly social action, namely an action that achieves a social goal. Social goals involve, at some point, another agent or some mental attitudes of that agent (goals, beliefs or emotions) [8]. Informing another agent about a third party is not the only goal gossip allows to achieve; it can imply other goals, such for instance the goal of influencing the other or changing her mind with regard to the gossipee. Distinguishing between the gossiper's goals is not unconsequential and allows to draw some hypotheses, although very preliminary.

There are two main goals the gossiper aims to achieve: influencing other agents or informing them (information can be used also to influence but here we do not care for this second-order motivation). In the former case, evaluation transmission is meant to induce new beliefs or goals, through *cognitive influencing*, that

> consists of providing the addressee with information that is pretended to be relevant for some of her goals, and this is done in order to ensure that the recipient has a new goal ([8] p. 32).

This mechanism is really powerful: an agent can induce another one to perform an action that was neither intended nor planned, simply by making her know that, for instance, the topic has been unfaithful to her. In this way the gossiper acts on the receiver's belief about the gossipee and this could lead the receiver to change her plan. For instance, the receiver can decide to avoid meeting the gossipee because she has a negative evaluation or, also, the receiver can use this information to influence another agent, even the gossipee himself (she can blackmail him in change for money).

While cognitive influencing is targeted (a particular agent is informed in order to achieve a given goal), the goal of informing is more general and can be rephrased as the goal of making information circulate in the network, no matter who the target is. This can be done because information is deemed useful to the group and disseminating it in the widest possible way allows the group to learn in a fast and efficient way, as when talking about the consequences of someone's choice. Knowing what happened to someone who faced a difficult situation can be useful for the listener as a guidance for her behavior or for learning how to avoid committing the same mistake. In this view, the goal of the gossiper is that of making the receiver(s) know how to behave in a certain situation or what consequences can follow, more than simply reporting what happened to another individual. Information spreading can also be pivotal to the strengthening of social bonds among group members. [2]

The final distinction regards the real addressee of the comunication: the receiver or the gossipee. This choice is not trivial and leads to two communication dynamics. In the former case, the gossiper collects, selects, and transmits information about the target with the purpose of reaching the receiver. This path is the most typical and it is "gossip" in the classical sense. When who is told about becomes the actual addressee of gossip talk, the receiver's involvement is only a ploy but the receiver should be a suitable, chatty agent to ensure that she will report to the gossipee, as planned. This "indirect gossip" can be motivated by several reasons. First, the gossiper wants to flatter the gossipee making her aware of her approval or, on the contrary, the gossiper wants to transmit to the gossipee her blame and condemn for her actions. In this latter case, the gossipee may react to what is considered malicious gossip, thus triggering a succession of gossiping with several consequences, both for her and for the gos-

[2] The path going from individual minds to collective phenomena can be followed also in the opposite direction, in which reputation, as a collective and emergent phenomenon, affects agents' cognition, but this issue will not be addressed here.

siper, ranging from ostracism to deep rearrangements of the social network.

## 2.2 Linking micro elements with macro phenomena

Once described what happens at the lowest level, we can put forward some hypotheses about social evaluations and networks at the macro level. The cognitive model of reputation and gossip can be related to social networks under several respects, but here the focus will be on the difference between networks based on reputation and networks based on image. In fact, the distinction between image and reputation appears to be relevant also in terms of network's structure and enlargement potentiality: image-based networks are based on familiarity and have a low potential for enlargement and innovation. On the other hand, reputation-based networks are wider, flexible and can be innovated more easily.

The difference between image-based and reputation-based networks has been proposed by Conte, Paolucci and Sabater-Mir [10]. They developed a computational system, RepAge, a REPutation and imAGE tool [25], implemented on an agent architecture and tested in an artificial market. In this setting, agents were allowed to exchange image only or image plus reputation. Their results show that social networks based upon image perform more poorly than networks based upon reputation at least when partner selection is a common goal of the network members. This is mainly due to two distinct effects. On the one side, when only image, i.e. evaluation coming from an identified agent, is available, the presence of cheaters triggers a mutual defeat strategy, leading the system to collapse. This happens also with informational error: once agents find out that they received a false information, the informer can not be trusted any more. In the authors' words:

> An image-based social network is expected to be rigid, meaning rather sensitive to errors: if a given threshold of error is overcome, the whole system is probably bound to fall apart, and the network will be fatally affected by distrust.

On the other side, results coming from simulation experiments with Repage show that reputation-based networks are more flexible. Since evaluations are not immediately tested and they can not be attributed to any specific agent, the chain of retaliations is prevented and the network is more error tolerant.

Similar results have been obtained also by Giardini, Di Tosto and Conte [16, 17] with a computational model of industrial clusters. Industrial clusters are usually defined as networks of interactions among heterogeneous and complementary firms embedded into a specific geographic area. In the district, the form of production requires a high degree of cooperation between firms and the lack of formal agreements could lead actors to behave in an opportunistic manner, but the merging between social community and firms [4] helps preventing this result. In industrial districts the interplay between economic dimensions and social relationships is very close, and there is a rich social structure made of informal and personal connections among people working in the cluster. In this context, social evaluations are really important and play a fundamental role in economic exchanges. Actors in the cluster select their partners also relying on evaluations received by other partners and peers, and these evaluations may heavily affect also the economic performance of single firms and of the cluster as a whole. The importance of the social dimension and the possibility of assessing the importance of evaluations in terms of economic performance make the study of industrial clusters especially appropriate to test the cognitive model of reputation and gossip.

In our agent-based model, agents are firms (for a detailed description of the model, see [16, 17]) that can perform two kinds of exchanges: informational exchange and material exchange. Agents have to choose the best available supplier in order to deliver high quality products. When their known suppliers are not available, they can rely on informers, i.e. other agents transmitting evaluations, in order to avoid the costs of direct interaction and to acquire useful information. Two different settings were tested: an image only setting, in which agents transmitted their own evaluations and retaliation against bad informers was allowed, and a reputation setting. In this latter condition, evaluations were not tested immediately and evaluator's identity was undisclosed, so that agents exchanged reputation without the fear of retaliation. Fixed percentages of cheaters, i.e. agents always providing false information, were implemented in both conditions.

Our results show that the quality of production was higher in the cluster with reputational information, compared to the cluster with image, for the same percentages of cheating. In other words, social information gives rise to different network configurations: image-based networks perform better when cheaters are few but quality of production dramatically decreases for higher levels of cheating. Conversely, reputation-based networks are more flexible and resistent when the number of cheaters is high, and the cluster's quality of production is only partially affected. In addition, agents in these networks explore the environment both spreading and using untested evaluations, so that innovation is promoted through the inclusion of new partners and informers.

# 3 Conclusions and future work

Social evaluations are widespread in human societies and serve many distinct functions: they are a valuable source of information about the community and its members, but they are also essential to map the social environment, to promote membership, and to sanction deviant behaviours in a public way. This list is far from being complete and many other functions can be attributed to this pervasive human activity.

In order to deal with the complexity of social evaluations, we need to understand what happens in the agents' minds, what the causes are of this specific kind of social action, making an attempt to take into account both mental and social aspects of it. Linking the micro and the macro level through cognitive modeling could be proven to be really effective to unfold complex social phenomena and to understand how they emerge from individuals' representations and actions. Gossip and reputation originate in people's minds, spread in the social environment and affect other individuals' minds, in a continuous loop of emergence and immergence.

In this work a preliminary cognitive account of reputation and gossip has been put forward but it needs to be further refined and also to be enriched with testable hypotheses. Once developed a more detailed cognitive model of the actors involved in gossip, two possible directions for future work could be envisioned. The first direction could be to apply Multi-Agent Based (MAS) methodology to investigate through computer-based simulations how gossip generates, how agents use it and with which effects, testing hypotheses about the difference between gossip with a direct addressee, the receiver, and gossip with an indirect addressee. The second direction of research would be more focused on networks' structure and will be devoted to add on the distinction between reputation-based and image-based networks. A further direction of research could deal with the integration of simulation experiments with experiments with humans in the laboratory, trying to replicate, at least partially, the dynamics of gossip spreading and change, in order to test the cognitive model of reputation and gossip with human participants.

## ACKNOWLEDGEMENTS

This work was partially supported by the Italian Ministry of University and Scientific Research under the Firb2003 programme (Socrate project, contract number RBNE03Y338) and by the European Community under the FP6 programme (eRep project, contract number CIT5-028575).

## REFERENCES


[1] R. D. Alexander, *The Biology of Moral Systems (Foundations of Human Behavior)*, Aldine, July 1987.

[2] M. Ayim, *Knowledge through the grapevine: Gossip as inquiry*, 85–99, University Press of Kansas, 1994.

[3] R. F. Baumeister, L. Zhang, and K. D. Vohs, 'Gossip as cultural learning', *Review of General Psychology*, **8**, 111–121, (2004).

[4] G. Becattini, *The marshallian industrial district as socio-economic notion*, International Institute of Labour Studies, 1990.

[5] A. Ben-Ze'ev, *The vindication of gossip*, 11–23, University Press of Kansas, 1994.

[6] C. Boehm, *Hierarchy in the Forest: The Evolution of Egalitarian Behavior*, Harvard University Press, February 2000.

[7] C. Castelfranchi, *Progress in artificial intelligence*, chapter Emergence and Cognition: Towards a Synthetic Paradigm in AI and Cognitive Science, 13–26, Springer, Berlin, 1998.

[8] R. Conte and C. Castelfranchi, *Cognitive and social action*, Londra: London University College of London Press, 1995.

[9] R. Conte and M. Paolucci, *Reputation in Artificial Societies: Social Beliefs for Social Order*, Springer, October 2002.

[10] R. Conte, M. Paolucci, and J. Sabater Mir, 'Reputation for innovating social networks', *Advances in Complex Systems*, **11(2)**, 303–320, (2008).

[11] R. De Sousa, *In praise of gossip: Indiscretion as a saintly virtue*, 25–33, University Press of Kansas, 1994.

[12] R.I.M. Dunbar, *Grooming, gossip and the evolution of language*, Harvard University Press, 1997.

[13] G.A Fine and R.L. Rosnow, 'Gossip, gossipers, gossiping', *Personality and social psychology bulletin*, **4**, 161–168, (1978).

[14] E. K. Foster, 'Research on gossip: Taxonomy, methods, and future directions', *Review of General Psychology*, **8**, 78–99, (2004).

[15] C. D. Frith and U. Frith, 'How we predict what other people are going to do', *Brain Research*, **1079**, 36–46, (March 2006).

[16] F. Giardini, G. Di Tosto, and R. Conte, 'A model for simulating reputation dynamics in industrial districts', *Simulation Modelling Practice and Theory (SIMPAT)*, **16(2)**, 231–241, (2008).

[17] F. Giardini, G. Di Tosto, and R. Conte, 'Reputation and economic performance in industrial districts: Modelling social complexity through multi-agent systems', in *World Congress on Social Simulation 2008 (WCSS-08)*, p. . George Mason University, Fairfax, USA, (2008).

[18] H. Gintis, E. A. Smith, and S. Bowles, 'Costly signaling and cooperation', *Journal of Theoretical Biology*, **213**(1), 103–119, (November 2001).

[19] M. Gluckman, 'Papers in honor of melville j. herskovits: Gossip and scandal', *Current Anthropology*, **4**(3), (1963).

[20] J.K. Hamlin, K. Wynn, and P. Bloom, 'Social evaluation by preverbal infants', *Nature*, **450**, 557–559, (2007).

[21] M. Noon and R. Delbridge, 'News from behind my hand: Gossip in organizations', *Organization studies*, **14**, 23–36, (1993).

[22] M. A. Nowak and K. Sigmund, 'Evolution of indirect reciprocity by image scoring', *Nature*, **393**(6685), 573–577, (June 1998).

[23] K. Panchanathan and R. Boyd, 'Indirect reciprocity can stabilize cooperation without the second-order free rider problem.', *Nature*, **432**(7016), 499–502, (November 2004).

[24] R.L. Rosnow, 'Psychology of rumor reconsidered', *Psychological Bulletin*, **87**, 578–591, (1980).

[25] J. Sabater, M. Paolucci, and R. Conte, 'Repage: Reputation and image among limited autonomous partners', *Journal of Artificial Societies and Social Simulation*, **9(2)**, (2006).

[26] G. Taylor, *Gossip as moral talk*, 34–46, University Press of Kansas, 1994.

[27] S. R. Wert and P. Salovey, 'A social comparison account of gossip', *Review of General Psychology*, **8**, 122–137, (2004).

[28] D. S. Wilson, C. Wilczynski, A. Wells, and L. Weiser, 'Gossip and other aspects of language as group-level adaptations', in *The evolution of cognition*, eds., C. Heyes and L. Huber, MIT Press, Cambridge, (2000).

# The Design of Convivial Multiagent Systems

**Patrice Caire and Leendert van der Torre**[1]

**Abstract.** In this paper, we consider the design of convivial multi-agent systems. Conviviality has recently been proposed as a social concept to develop multi-agent systems. In this paper we introduce temporal dependence networks to model the evolution of dependence networks and conviviality over time, we introduce epistemic dependence networks to combine the viewpoints of stakeholders, and we introduce normative dependence networks to model the transformation of social dependencies by hiding power relations and social structures to facilitate social interactions. We show how to use these visual languages in design, and we illustrate the design method using an example on virtual children adoptions.

## 1 Introduction

The focus of this paper is the social/organizational structure of a multiagent system. In particular, we are interested in the design of *convivial* multiagent systems, which is directly related to well studied issues such as groups and teams, norms and normative behavior, and coalition formation. First, we discuss the determining factors and the decisions we have to make concerning the actual convivial characteristics of the system. Following the TROPOS methodology, this process leads us to our dependence network model. A crucial step in this phase is to manage conflicting requirements such as reconciling freedom with exclusion and missing or incomplete specifications such as implicit agents goals. Second, we propose a representation of our model and present our formalism, initially expressing dependencies with static dependence network. We then express the sequence of different actors point of views, temporal dynamic networks. Third, we define the actors interactions and model a protocol.

We study the following research questions:

1. How to design the evolution of convivial social relations?
2. How to combine viewpoints from stakeholders?
3. How to incorporate normative aspects of conviviality?

The description level of this paper is methodologies and languages. To answer these questions we develop temporal dependence networks to model the evolution of dependence networks and conviviality over time, we introduce epistemic dependence networks to combine the viewpoints of stakeholders, and we introduce normative dependence networks to model the transformation of social dependencies by hiding power relations and social structures to facilitate social interactions.

The inspiration source of our work is political and social science. Empathy and reciprocity were foregrounded by Polanyi in 1964. "Individual freedom realized in personal interdependence" was tooled up by Illich in 1974 [12]. And in 1988, Putnam considered conviviality as a condition for civil society and social capital, a concept referring to the collective values of all social networks. One of the four themes of the European Community fifth framework program was entitled the "societe de l'information conviviale" (1998-2002) [18], which was translated as "the user-friendly information society." Today, a number of research fields such as computer supported cooperative work and social software aim at supporting users to interact and share data. Conviviality has recently been proposed also as a social concept to develop multi-agent systems [5].

As a running example, we use the design of a virtual adoption agency for instance on Second Life (SL). Adopting virtual children is a successful experience and a flourishing business on SL. Parents wishing to adopt a child must pay a fee to the adoption agency. The procedure typically involves that parents list themselves to advertise their profile to prospective children who can select them. The agency then matches children and parents and organizes a try-out period. There is no pressure. Once parents and children have made their decision, they simply come back to the agency to cancel the adoption if unhappy or otherwise to confirm it and get their adoption certificate and a ceremony. The experience must be convivial.

The conviviality literature discusses many definitions and relations with other social concepts, which we do not introduce in the formal model in this paper, referring to qualities such as trust, privacy and community identity. Also, in this paper we do not consider Polanyi's notion of empathy, which needs trust, shared commitments and mutual efforts to build up and maintain conviviality.

The layout of this paper is as follows. In Section 2 we discuss the social focus of this paper by explaining how the social concept "conviviality" can be used to develop multiagent systems in general, and their design in particular. In the following four sections we answer the research questions. In Section 3 we introduce temporal dependence networks to model the evolution of dependence networks and conviviality over time. In Section 4 we introduce epistemic dependence networks to combine the viewpoints of stakeholders. In section 5 we introduce normative dependence networks to model the transformation of social dependencies.

## 2 Convivial multiagent systems

In this section we discuss the use of social concepts in general, and "conviviality" in particular, for the development of multiagent systems.

Illich defines conviviality as "individual freedom realized in personal interdependence" [12]. We therefore model it using dependence networks [6, 16], representing on which agents and agent depends to fulfill its goals. An agent depends on a set of agents to fulfill one of its goals, when the set of agents has the power to fulfill the goal.

We define conviviality masks based on Taylor's idea that conviviality "masks the power relationships and social structures that govern

[1] University of Luxembourg, Luxembourg, email: patrice.caire@uni.lu and leendert@vandertorre.com

societies." [17, 5, 3]

> A conviviality mask is a transformation of social dependencies by hiding power relations and social structures to facilitate social interactions.

## 2.1 Conviviality requirements

Requirements for multiagent systems say that systems must be convivial, whereas system researchers and developers use other concepts. To model the requirement, the developers may interpret the conviviality requirement as being autonomous to make suggestions, to react the discussion in the meeting to reach their goals, being proactive to take the initiative and being goal-directed, and most importantly being social by interacting with others to reach their goals.

When writing down requirements for user friendly multiagent systems, it is crucial to understand the inherent threads of conviviality, such as deception, group fragmentation and reductionism [5]. Whereas conviviality was put forward by Illich as a positive concept, also negative aspects were discussed. People are often not rational and cooperative to achieve conviviality [15] and unity through diversity [11] may lead to suppression of minorities. Taylor explores the contradiction that conviviality cannot exist outside institutions: i.e., the question "whether it is possible for convivial institutions to exist other than by simply creating another set of power relationships and social orders that, during the moment of involvement, appear to allow free rein to individual expression. Community members may experience a sense of conviviality which is deceptive and which disappears as soon as the members return to the alienation of their fragmented lives."

## 2.2 Conviviality ontology

The use of conviviality as a computer science concept ensures that considerations on the user-friendliness of multiagent systems get the same importance and considerations on the functionality of the system. For example, our experience with the development of a digital city in Europe is that computer engineers are focussed on filling in forms and developing menu structures and other interface issues, and do not take into account that a digital city should be a meeting place for human and artificial agents.

Conviviality is a useful high level modeling concept for organizations and communities, emphasizing the social side of them rather than the legal side. Erickson and Kellogg [9] say: "In socially translucent systems, we believe it will be easier for users to carry on coherent discussions; to observe and imitate others' actions; to engage in peer pressure; to create, notice, and conform to social conventions. We see social translucence as a fundamental requirement for supporting all types of communication and collaboration". Taylor studies conviviality in British pantomime and observes that: "conviviality masks the power relationships and social structures that govern societies."

## 2.3 Design of convivial systems

In this paper we study how convivial multiagent systems can be designed using our operationalized concept of conviviality. We illustrate our arguments and contributions with a running example on multiagent systems for virtual adoptions, where typically physical reality such as multiagent technologies interact with virtual and social realities.

The aim of social scientists to create conviviality by creating the desired conditions for social interaction, coincides with the aim of designers of multiagent systems. For example, Illich defines a convivial learning experience in which the teacher and the student switch roles, such that the teacher becomes the student and the student becomes the teacher. This role swapping emphasizes the role of reciprocity as a key component for conviviality. Parallelely the importance of reciprocity in conviviality was shown for instance in [10]. As a result, such role swapping scenarios can directly be used in multi-agent systems.

# 3 Temporal dependence networks

In this section, we propose a design methodology for convivial multiagent systems based on the agent-oriented software development process, Tropos [1]. Key ideas in Tropos are first, that throughout the process phases, e.g. from early requirements to implementation, agents are endowed with intentionality. Second, the importance of very early phases of requirement analysis to allow for a profound understanding of the environment and of the interactions for the software to be built. This methodology guides designer through an incremental process, from the initial model of stakeholders, to refined intermediate models that, at the end, becomes the code.

## 3.1 Dependence networks

Multiagent systems technology can be used to create tools for conviviality. Illich defines conviviality as "individual freedom realized in personal interdependence" [12]. Dependence network is a tool that allows us to model this interdependence [6, 16]. In a recently published paper [5] dependence networks were formally defined as in Def. 1.

**Definition 1 (Dependence networks)** *A dependence network is a tuple* $\langle A, G, dep, \geq \rangle$ *where:*

- $A$ *is a set of agents*
- $G$ *is a set of goals*
- $dep : A \times 2^A \rightarrow 2^{2^G}$ *is a function that relates with each pair of an agent and a set of agents, all the sets of goals on which the first depends on the second.*
- $\geq$: $A \rightarrow 2^G \times 2^G$ *is for each agent a total pre-order on goals which occur in its dependencies:* $G_1 \geq (a)G_2$ *implies that* $\exists B, C \subseteq A$ *such that* $a \in B$ *and* $G_1, G_2 \in depend(B, C)$.

Nevertheless, this representation of conviviality is static and therefore has a limited field of application. In the next sub-section, we present our extension to encompass the temporal aspect of conviviality.

## 3.2 Temporal dependence networks

Before proposing our definition, we introduce our virtual adoption running example. The procedure typically involves that parents list themselves to advertise their profile to prospective children who, if they like the parents, can select them. The agency then matches children and parents and organizes a try-out period. Once parents and children have made their decision, they simply come back to the agency to cancel the adoption if unhappy or otherwise to confirm it and get their adoption certificate and a ceremony.

We start by informally listing critical stakeholders. We then identify the relevant goals and the social dependencies of the stakeholders represented as actors. In particular, the actor **Parent** is associated with the goal: adopt child, while the actor **Child** is associated with the goal: get adopted and **Virtual Agency** with the goal: provide adoption service.

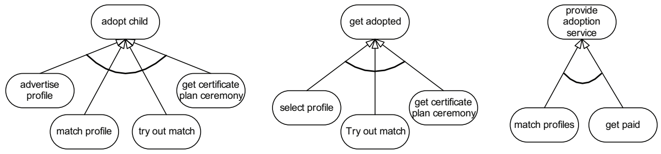


**Figure 1.** Decomposition of goals.

To enrich the model with a finer goal structure and elicit dependencies, we decompose each root goal into sub-goals. For instance, **Child** goal: get adopted, is decomposed into three sub-goals: select profile, try out match and get certificate - plan ceremony. In Fig. 1, a graphical representation of goal modeling is given through a goal diagram; AND decomposition only are shown, no OR decomposition, e.g. no alternate sub-goals.

The UML sequence diagram (Fig. 2), illustrates the interactions among the stakeholders and how operations are carried out. The diagram shows time incrementing vertically. In particular, the diagram models the interaction among the three Users: **parent**, **agency** and **child**. The interaction starts with the advertise profile request by the **parent** to the **agency** and ends with the pay fee by the **parent** to the **agency**. We note that the match ok sent by both **parent** and **child** can be asynchronous. Moreover, the **agency** sends the adoption certificate and the plan ceremony to both **child** and **parent**.

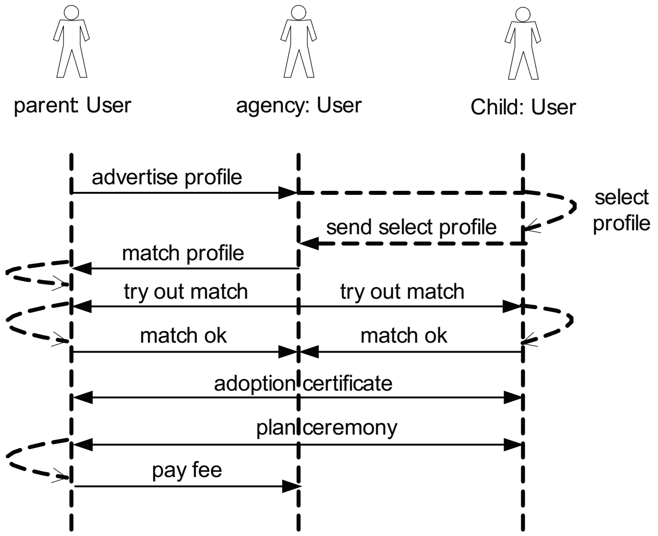


**Figure 2.** Actor diagram modeling the stakeholders for the virtual adoption domain.

Based on actor diagrams and goal decomposition, we proceed with a goal analysis taking each actor point of view. The objective is to obtain a set of strategic dependencies among the actors. We therefore perform an iterative analysis on each goal until all are analyzed. We build a succession of dependence networks from each actor point of view.

With temporal dependence networks, we aim at analyzing the evolution of dependence networks and conviviality over time. We identify the most relevant interactions in our running example and build a model with the key succession of dependence networks.

**Definition 2 (Temporal dependence networks)** *A dependence network is a tuple $DP = \langle A, G, goals, dep\rangle$ where:*

- *$A$ is a set of agents*
- *$G$ is a set of goals*
- *$T$ is the set of natural numbers*
- *$goals : T \times A \to 2^G$ is a function that relates with each pair of a sequence number and an agent, the set of goals the agent is interested in.*
- *$dep : T \times A \times 2^A \to 2^{2^G}$ is a function that relates with each triple of a sequence number, an agent and a set of agents, all the sets of goals on which the first depends on the second if the third creates the dependency.*

We use this structure to model our example (Fig. 3). Note that the set of agents does not change, but the goals of the agents and the dependencies among them, changes over time.

Agents $A = \{P, C, VA\}$ and

Goals $G = \{g_1, g_2, g_3, g_4, g_5, g_6, g_7, g_8, g_9, g_{10}\}$

We thus have the following sequence of dependence networks:

$DP_4 = \langle A, G, goals_4, dep_4\rangle$, where:

- $goals(4, VA) = \{\{g_5, g_6, g_7\}\}$: In $dep_4$, the goals of agent $VA$ are to provide adoption service, to get paid and to match parent-child profiles.
- $goals(4, P) = \{\{g_1, g_{10}\}\}$: In $dep_4$, the goals of agent $P$ are to adopt a child and to try out match.
- $goals(4, C) = \{\{g_8, g_{10}\}\}$: In $dep_4$, the goals of agent $C$ are to get adopted and to try out match.
- $dep(4, VA, \{P, C\}) = \{\{g_7\}\}$: In $dep_4$, agent $VA$ depends on agents $P$ and $C$ to achieve goal $g_7$: match parent-child profiles.
- $dep(4, P, \{C\}) = \{\{g_{10}\}\}$: In $dep_4$, agent $P$ depends on agents $C$ to achieve goal $g_{10}$: try out match.
- $dep(4, C, \{P\}) = \{\{g_{10}\}\}$: In $dep_4$, agent $C$ depends on agents $P$ to achieve goal $g_{10}$: try out match.

In our notation, $dep_i$ refers to the temporal dependence network where $i \in T$ and denotes the $i^{th}$ sequence, $P$ refers to agent **Parent**, $C$ to agent **Child** and $VA$ to agent **Virtual Agency**.

## 4 Epistemic dependence networks

In our running example, we use the Tropos methodology [1], with the difference that we include neither plans nor resources. However similarly to Tropos, we identify actors which depend on each other to achieve their hardgoals, simply referred to as goals, and softgoals, the latter being typically used to model non-functional requirements and "having no clear -cut definition and/or criteria for deciding wheter they are satisfied or not" [1]. In Fig. 4, we show an *actor diagram* for the virtual adoption. In particular, **Parent** is associated with the goal: adopt child, and the softgoal: get nice child. Similarly, **Child** is associated with the goal: get adopted and the softgoal get nice parents while **Virtual agency** wants to provide adoption service and has the

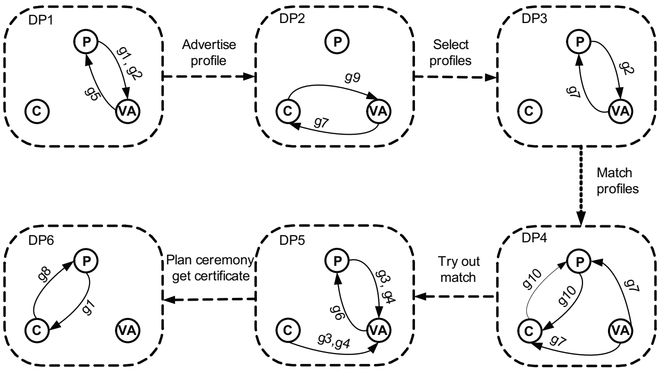


**Figure 3.** DP sequences

softgoal to provide a good service. Finally, the diagram includes one softgoal dependency where **Parent** depends on **Virtual agency** to fulfill the softgoal: adoption fee well spent.

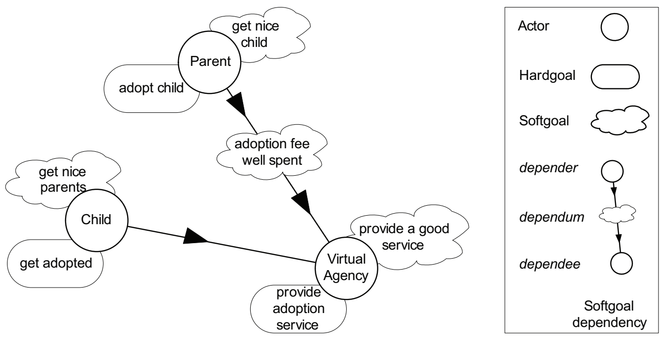


**Figure 4.** Actor diagram modeling the stakeholders for the virtual adoption.

Temporal dependence networks allow us to capture a relation from a specific point of view and at a specific time. Unfortunately, it is not sufficient for the situation we want to model, so in the next section, we try to answer this question by introducing a new model that will allow us to capture a more global view from the system point of view.

In order to model such system, we use the epistemic dependence network formally defined as Def. 3.

**Definition 3 (Epistemic dependence networks)** *An epistemic dependence network is a tuple* $DP = \langle A, G, T, goals, dep \rangle$ *where:*

- $A$ *is a set of agents*
- $G$ *is a set of goals*
- $T$ *is the set of natural numbers*
- $goals : T \times A \rightarrow 2^G$ *is a function that relates with each pair of sequence number and an agent, the set of goals the agent is interested in.*
- $dep : A \rightarrow T \times A \times 2^A \rightarrow 2^{2^G}$ *is a function that expresses from the point of view of an agent* $a \in A$*, the dependence relation between another agent* $b \in A$ *and a set of other agents regarding the goals of agent* $b$ *in a sequence* $t \in T$.

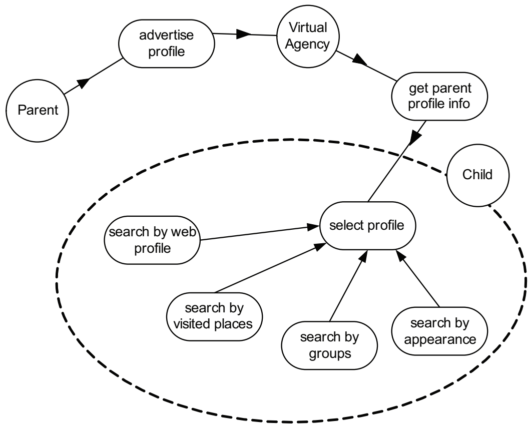


**Figure 5.** Goal diagram for the goal select profile and dependencies between the actor Child and other environment's actors.

If we consider Fig. 5 the starting *goal diagram*, the three steps of this design process are:

1. Goal delegation: Each goal of any actor may be delegated to any other actor, already existing or new. It proceeds with the analysis of goals from the point of view of each actor. This generates a network of delegation between stakeholders, external actors and the system. The inclusion of new actors and sub-actors and subsequently, the delegation of sub-goals to sub-actors continues until all goals have been analyzed. Actors that contribute to the requirements are also included.
2. Goal decomposition: Goals and softgoals are further decomposed into sub-goals or found not reachable. Through this refinement process a goal hierarchy is created where leaf goals represent alternatives to root goals. Moreover, some identified sub-goals become reasons for new dependencies with new actors. Therefore, dependencies in actors diagrams must be revised.
3. When all actors fulfill their goals, all the goals have been analyzed and the root goals are satisfied then, this design process is complete.

## 4.1 Example

In our running example, let's consider the set of agents

$A = \{P, C, VA, AS\}$, where $AS$ is the **Adoption System**.

$dep(P) = (2, VA, \{C\}) = \{g_9\}$: **Parent** believes that in sequence 2, **Adoption System** depends on **Child** to achieve goal $g_9$: select profile.

We express Fig. 6 as follows: $dep(AS) = (2, P, \{C\}) = \{g_9\}$: **Adoption System** believes that in sequence 2, **Parent** depends on **Child** to achieve goal $\{g_9\}$: select profile. We note that there is no dependency from **Adoption System** towards **Adoption System** for the goal: select profile.

With Fig. 5 and 6, we explain the iterative design process from the Tropos methodology that are tool supported [14].

To explain what is the delegation process, and as an example, we here give a partial view on goal: select profile.

To start, we have the goal of **Child**: select profile. After analyzing the rational for this goal from each actor point of view, we delegate this goal to the new actor, the system-to-be **Adoption System** . We continue by analyzing each sub-goal.

We then identify the capabilities needed by **Adoption System** to fulfill all the four identified sub-goals: search by web profile, search by visited places, search by groups and search by appearance. In order for this latter sub-goal to be fulfilled, we add a new goal: provide photo/video and a new dependency from **Adoption System** towards **Parent**. Similarly, in Fig. 5 the sub-goal: search by web profile has no dependency while in 6 a new dependency from **Adoption System** towards **Child** has been created to fulfill the subgoal: know web address. Of course, each dependency must be mapped to a capability. We then define a set of agent types and assign each of them one or more capabilities. The specification of agent's goals, beliefs, capabilities and the communication between the agents depends on the adopted platform and the chosen programming language. We therefore leave this part for further work.

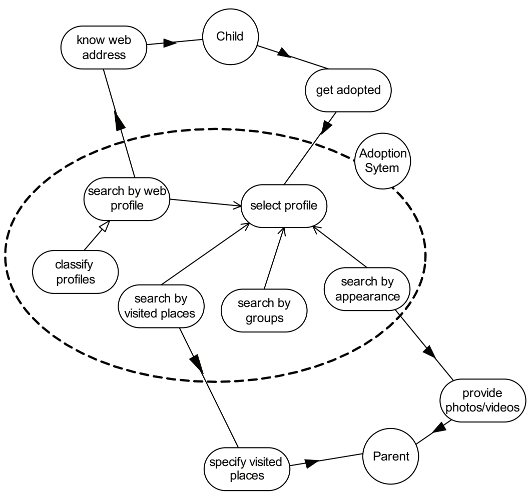


**Figure 6.** Goal diagram for the goal select profile and dependencies between the actor Adoption System and other environment's actors.

## 4.2 Nested dependencies

We first mention that by *nested* we simply mean a belief produced and only accessible by an agent $a$ and about another agent $b$, e.g. inaccessible to all others. For instance, empathy provides a way to know what another agent's preference is, and therefore to better adapt to it, allowing for a convivial relation, whereby agents contribute to each other. In our running example, let's assume that **Parent** believes that **Child** depends on it, **Parent**, for its goal: select profile. Let's further assume that **Child** believes that **Parent** depends on it to advertise parent profile, for example if **Child** first had to publish an announcement on a board that it is seeking parents to be adopted by. We write:

$dep(P) = (1, C, \{P\}) = \{g_9\}$: agent $P$ believes that in sequence 1, agent $C$ depends on it, $P$ to achieve its goal $g_9$: select parents' profile.

$dep(C) = (1, P, \{C\}) = \{g_2\}$: agent $C$ believes that in sequence 1, agent $P$ depends on it, $C$, to achieve its goal $g_2$: advertise its profile.

## 5 Norms and masks

There are many different kinds of goals, some goals may be considered normative, others personal. Agents do not only have personal goals, they also have normative goals, e.g. goals imposed by the procedures. We propose a further extension of epistemic dependence networks that we call "Normative epistemic dependence networks" in order to take into account the differences in the two kinds of goals as well as obligations and violations.

**Definition 4 (Normative epistemic dependence networks)** *A dependence network is a tuple*
$DP = \langle A, G, N, O, V, T, goals, dep\rangle$ *where:*

- $A$ *is a set of agents*
- $G$ *is a set of goals*
- $N$ *is a set of norms*
- $T$ *is the set of natural numbers*
- $O : N \times A \rightarrow 2^G$ *is a function that associates with each norm and agent the goals the agent must achieve to fulfill the norm; We assume for all* $n \in N$ *and* $a \in A$ *that* $O(n, a) \in power(\{a\})$*;*
- $V : N \times A \rightarrow 2^G$ *is a function that associates with each norm and agent the goals that will not be achieved if the norm is violated by agent a; We assume for each* $B \subseteq A$ *and* $H \in power(B)$ *that* $(\cup_{a\in A} V(n, a)) \cap H = \emptyset$*.*
- $goals : T \times A \rightarrow 2^G$ *is a function that relates with each pair of sequence number and an agent, the set of goals the agent is interested in.*
- $dep : A \rightarrow T \times A \times 2^A \rightarrow 2^{2^G}$ *is a function that expresses from the point of view of an agent* $a \in A$*, the dependence relation between another agent* $b \in A$ *and a set of other agents regarding the goals of agent* $b$ *in a sequence* $t \in T$*.*

## 5.1 Example 1

We explain with an example how to use our formalism and model normative situations. In sequence 2 of our running example, while **Child**'s obligation to select profiles is a normative goal, **Child**'s desire to select the parents it prefers is a personal goal. In this case, personal and normative goals coincide:

The goal $g_9$, to select parents' profile, is both a personal goal and a normative goal, that is, $goals(2, C) = g_9 \cup O(2, C) = g_9$, where $g_9 \in PG_C$: in sequence 2, agent $C$ has the goal and the obligation to select parents' profiles $g_9$, where *PG* is personal goal.

$G_C = \cup O(n, C) \cup PG_C$, where $G_C \in G$ is the set of normative goals of agent $C \in A$, $n \in N$ is an adoption norm, $O(n, C)$ is the obligation for $C$ to respect norm $n$ resulting in its normative goals, and $PG_C \in G$ are the personal goals of $C$.

## 5.2 Example 2

In this paragraph, we explain the notions of positive and negative consequences to a norm violation. A positive consequence is adding a goal to the existing ones whereas a negative consequence forbid the realization of a goal. We further explain with our example. Let's assume that the parent believes that, in sequence 2, the child depends on the virtual agency to hide its information to parents. However, the parent violates its obligation to respect it and looks up the child's information. One possible sanction is that the parent cannot advertise its profile at the agency any longer, which means that this goal is unrealizable. In the case of the violation sanctioned by the removal of the goal $g_2$, the obligation $O(n_2, P)$ is not possible any longer as agent $P$ cannot advertise its profile at the agency, it cannot depend on the agency to get the child information any longer. Moreover, agent $P$ cannot achieve its personal goal $g_1$: adopt a child, any longer as $g_2$ is a normative goal needed for agent $P$ to achieve $g_1$. And the violations are: $V^{-}(n_2, P) = g_2$: agent $P$ violating norm $n_2$ will not be able to achieve goal $g_2$, advertise its profile, because $g_2$ is removed.

As a consequence, the parent cannot adopt a child. Another possible sanction is that the parent must make a donation, e.g. pay a fee, in which case a new goal is added to the parent. As a result, until the parent has fulfill this new obligation, it cannot continue the process.

$dep(P) = (2, C, VA) = g_{14}$: agent $P$ believes that in sequence 2, agent $C$ depends on agent $VA$ to achieve its goal $g_{14}$: no child look up. Where the obligations are:

$O(n_1, C) = g_9$: agent $C$ has the obligation to fulfill norm $n_1$ to achieve goal $g_9$, select parent profile.

$O(n_2, P) = g_{14}$: agent $P$ has the obligation to fulfill norm $n_2$ to achieve goal $g_{14}$, no look up child.

$V^{+}(n_2, P) = g_{15}$: agent $P$ violating norm $n_2$ will not be able to achieve goal $g_2$, advertise its profile, because a new goal $g_{15}$, make a donation, is added. Until this new goal is achieved, $g_2$ cannot be achieved.

In the case of the violation sanctioned with the addition of the goal $g_{15}$, we note that a mechanism is needed to make sure that the new goal is fulfilled before agent $P$ can further proceeds.

## 6 Related work

Castelfranchi [6] introduces concepts like groups and collectives from social theory in agent theory, both to enrich agent theory and to develop experimental, conceptual and theoretical new instruments for the social sciences. For further work on the use of the concept of conviviality in computer science and multiagent system see [3, 2, 4]. A large body of work on design has been produced, to only cite a few: the AOSE methodology [13], GAIA [7], the PASSY methodology [8].

## 7 Summary

Conviviality has recently been proposed as a social concept to develop multi-agent systems. In this paper we introduce temporal dependence networks to model the evolution of dependence networks and conviviality over time, we introduce epistemic dependence networks to combine the viewpoints of stakeholders, and we introduce normative dependence networks to model the transformation of social dependencies by hiding power relations and social structures to facilitate social interactions. We show how to use these visual languages in design, and we illustrate the design method using an example on virtual children adoptions.

## ACKNOWLEDGEMENTS

We would like to thank the referees for their comments which helped improve this paper.

## REFERENCES

[1] P. Bresciani, A. Perini, P. Giorgini, F. Giunchiglia, and J. Mylopoulos. Tropos: An agent-oriented software development methodology. *Autonomous Agents and Multi-Agent Systems*, 8(3):203–236, 2004.

[2] P. Caire. How to import the concept of conviviality to web communities. *Web communities*, 2009.

[3] P. Caire and L. van der Torre. Convivial ambient technologies: Requirements, ontology, and design. *The Computer Journal*, 2009.

[4] P. Caire and L. van der Torre. Temporal dependence networks for the design of convivial mul- tiagent systems (short paper). In *Proc. of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AA- MAS 2009)*, 2009.

[5] P. Caire, S. Villata, G. Boella, and L. van der Torre. Conviviality masks in multiagent systems. In L. Padgham, D. C. Parkes, J. Müller, and S. Parsons, editors, *Proc. of 7th Int. Conf. on Autonomous Agents and Multiagent Systems (AA- MAS 2008)*, pages 1265–1268. IFAAMAS, 2008.

[6] C. Castelfranchi. The micro-macro constitution of power. *Protosociology*, 18:208–269, 2003.

[7] L. Cernuzzi and F. Zambonelli. Dealing with adaptive multi-agent organizations in the gaia methodology. In J. P. Müller and F. Zambonelli, editors, *AOSE*, volume 3950 of *Lecture Notes in Computer Science*, pages 109–123. Springer, 2005.

[8] A. Chella, M. Cossentino, L. Sabatucci, and V. Seidita. Agile passi: An agile process for designing agents. *Comput. Syst. Sci. Eng.*, 21(2), 2006.

[9] T. Erickson and W. A. Kellogg. Social translucence: an approach to designing systems that support social processes. *ACM Trans. Comput.-Hum. Interact.*, 7(1):59–83, 2000.

[10] E. R. Gomes, E. Boff, and R. M. Vicari. Social, affective and pedagogical agents for the recommendation of student tutors. In *Proceedings of Intelligent Tutoring Systems*, 2004.

[11] W. Hofkirchner. Unity through diversity.dialectics - systems thinking - semiotics. *Trans, Internet journal for cultural sciences*, 1(15), 2004.

[12] I. Illich. *Tools for Conviviality*. Marion Boyars Publishers, London, August 1974.

[13] J. Odell, P. Giorgini, and J. P. Müller, editors. *Agent-Oriented Software Engineering V, 5th International Workshop, AOSE 2004, New York, NY, USA, July 19, 2004, Revised Selected Papers*, volume 3382 of *Lecture Notes in Computer Science*. Springer, 2004.

[14] L. Penserini, A. Perini, A. Susi, and J. Mylopoulos. High variability design for software agents: Extending tropos. *TAAS*, 2(4), 2007.

[15] M. D. Sadek, P. Bretier, and E. Panaget. ARTIMIS: Natural dialogue meets rational agency. In *International Joint Conferences on Artificial Intelligence (2)*, pages 1030–1035, 1997.

[16] J. S. Sichman and R. Conte. Multi-agent dependence by dependence graphs. In *Procs. of The First International Joint Conference on Autonomous Agents & Multiagent Systems, AAMAS 2002*, pages 483–490. ACM, 2002.

[17] M. Taylor. Oh no it isn't: Audience participation and community identity. *Trans, Internet journal for cultural sciences*, 1(15), 2004.

[18] C. Weyrich. Orientations for workprogramme 2000 and beyond. Information society technologies report, Information Society Technologies Advisory Group, September, 17 1999.

# Micro-Social Systems: Interleaving Agents, Norms and Social Networks

**Moez Draief** and **Jeremy Pitt** and **Daniel Ramirez-Cano**[1]

**Abstract.** Ad hoc networks can be formed from arbitrary collections of individual people (forming online computer-mediated communities), mobile routers (forming data communication networks) or electronic business processes (forming virtual enterprises).

One way to deal with common features of dynamism in the network topology and membership, conflicts, sub-ideal operation, security, and the general need for continuous operation in the absence of a centralised facility, is to treat the ad hoc network as a norm-governed multi-agent system and use participatory adaptation as the mechanism for achieving autonomic capability (i.e. a global system response derived from the collective local behaviours and interactions of the individuals comprising the system).

Therefore, complementing the formal representation of organisational behaviour defined in terms of roles, rules, norms, etc., this autonomic capability is at least partially derived from an underlying social network which plays a significant role in determining how, for example, conflicts are resolved and how the organisation itself is run.

This position statement presents initial developments in what we call *micro-social systems*, which arise from interleaving a logical model of norm-governed systems with a mathematical model of social networks, and its application to issues of resource allocation, security, conflict resolution and self-adaptation in ad hoc networks.

## 1 INTRODUCTION

Ad hoc networks can be formed from arbitrary collections of individual people (forming online computer-mediated communities), mobile routers (forming wireless data communication networks) or electronic business processes (forming virtual enterprises).

All three types of ad hoc network exhibit similar features which require attention, for example:

- Dynamism: the network topology and nodes can vary rapidly and unpredictably.
- Conflicts: the network consists of heterogeneous nodes which may be competing rather than co-operating. This might give rise to conflicts over opinions, priorities, contracts, and so on.
- Sub-ideal operation: the nodes themselves may fail to comply according to the system specification, by accident, necessity, or design.
- Security: a successful operation may be subject to attack, either from malice or for profit. Examples include deception in online communities, denial of service in ad hoc networks, untrustworthiness in virtual enterprises, and so on.
- Continuity: even if all the network nodes change, it may be desirable for the network itself to be recognisably and identifiably "the same", offering the same support, functionality, services, and so on.

However, we anticipate that there is likely to be no permanently-connected, centralised facility (e.g. a moderator in an online community, a network server, or a central database, file manager or other co-ordinating mechanism in a virtual enterprise) to resolve any of these issues; unsurprisingly, of course, since these are essentially peer-to-peer networks.

Various proposals to deal with these issues have been presented, which analyse the ad hoc network as an open system (see [4] for a survey). In the same vein, we have advocated the use of norm-governed multi-agent systems [5] for formal specification of open systems, and more recently to facilitate adaptation by voting to change the specification [8]. In particular, following ideas of Ostrom [25] we proposed to use *participatory adaptation* as the mechanism for achieving autonomic capability (i.e. a global system response derived from the collective local behaviours and interactions of the individuals comprising the system).

However, norm-governed system specification provided the framework for (provably) correct procedures for casting and counting votes [26], and declaring the winner, while mechanism design provided the algorithm(s) for mapping expressed preferences onto a collective choice (computational social choice [9]). This formal analysis, from an external perspective, needs to be complemented by data structures and protocols which enable individual agents to determine a preference to express, usually based on information exchanged with their peers. Therefore, complementing the formal representation of organisational behaviour defined in terms of roles, rules, norms, procedures, preferences, etc., the requisite autonomic capability is at least partially derived from an underlying social network which plays a significant role in determining how, for example, conflicts are resolved and how the organisation itself is run.

This position statement presents initial developments in what we call *micro-social systems*, which arise from interleaving a logical model of norm-governed systems with a mathematical model of social networks, and its application to issues of resource allocation, security, conflict resolution and self-adaptation in ad hoc networks. We begin in Section 2 with a review of relevant background, and continue in Section 3 with a discussion and definition of micro-social systems. We then discuss (Section 4) a sample scenario in security in ad hoc networks, paying attention to the problem that this is a resource-constrained environment (i.e. power is a limited resource). We then present a proposed adaptive network security scheme, before drawing some conclusions and outlining our future programme of research in Section 5, in particular considering the inter-play and interleaving of overt organizational structures and the underlying social networks.

[1] Department of Electrical & Electronic Engineering, Imperial College London, UK, SW7 2BT; email: {m.draief,j.pitt,d.ramirez}@imperial.ac.uk

## 2 BACKGROUND

There is a rich tradition of research in multi-agent systems which is both informing and motivating the current proposal. This includes norm-governed systems, socio-cognitive systems, and information flow in social networks.

### 2.1 Norm-Governed Systems

A norm-governed multi-agent system can be expressed in terms of a set of agents (the members of a society), a set of social constraints on a society (norms, and other constraints, such as physical and logical constraints), a set of roles that members can play, a communication language, relationships between the members, including power, ownership and representation relations, and the structure of the society, including hierarchies, sub-structures, and ontologies for domains of responsibility.

We maintain the standard and long established distinction (in legal, social, and organizational systems research) between physical capability, institutionalised power and permission. Accordingly, a specification of the social constraints of a norm-governed multi-agent system expresses four aspects of agent activity: (i) the physical capabilities; (ii) institutionalised powers; (iii) permissions, prohibitions and obligations of the agents; and (iv) the sanctions and enforcement policies that deal with the performance of prohibited actions and non-compliance with obligations.

On this basis, we have developed a framework for executable specification of open systems [5], and an accompanying suite of protocols for norm-governed interaction, including resource allocation [2], negotiation [5], and voting in deliberative assemblies for virtual organizations [26].

In the most recent work, we have addressed the issue of organized adaptation in norm-governed systems, and have developed a norm-governed meta-protocol for changing object-level protocols and parameters [3]. The approach has also been applied to develop legal processes for virtual organizations; of particular relevance to the current context were protocols for conflict prevention and alternative dispute resolution [27].

### 2.2 Socio-Cognitive Systems

A socio-cognitive multi-agent system is a system in which the social relationships between agents are conditioned by cognitive, economic, or emotive factors. For example, in [23] we defined a framework which accounted for socio-cognitive and socio-economic factors to inform a trust decision in e-commerce. This showed how 'first encounter' decisions were best dealt with by a complex calculation based on risk exposure and recommendations ('risk' trust); but how '$n$th encounter' decisions were short-cut based on personal experiences (so-called 'reliance' trust). In [31] we combined both socio-cognitive and socio-emotive approaches to develop a formal model of forgiveness complementary to the trust model. Forgiveness turns out to be an essential element of autonomic agent systems as a repair mechanism for when a trust decision goes wrong: similarly, when a security breach occurs, the most effective response might not necessarily to engage a 'stronger' security level.

We have also been concerned with applications of multi-agent systems which depend on a set of distributed nodes exchanging information to reach some collective decision [28]. In open systems, though, self-interested agents might resort to different strategies, such as 'lying', in order to shift the collective decision towards that which will benefit their individual goals. We address this problem by defining such systems in terms of social exchange. Then, by making certain assumptions about the individual nodes and the collective decision-making, we define an opinion formation model of the system. We can then show how characteristics of the model can be matched with parameters of the system to achieve certain desirable properties of the information exchange, e.g. consensus, non-dominance, etc.

The study of norm-governed and socio-cognitive systems converged in the idea of participatory adaptation [7, 8]. This work used an iterated 'tragedy of the commons' scenario, with a dynamic network, partial knowledge, no central control, and self-interested agents. Two votes were taken: one for whom to allocate resources, and one to decide how many votes should be received in order to be allocated resources. The idea was that co-operative agents should manage the system by voting 'fairly'. Initial experiments showed that 'responsible' agents performed better than selfish or cautious ones [7] and that social networking (gossiping) algorithms can be used on an individual and group basis to protect the system from self-interested behaviour [8].

### 2.3 Information Flow in Social Networks

The dissemination of information is a ubiquitous process in human social networks and computer systems. It plays a fundamental role in numerous settings including the penetration of technological innovations [29], the word-of-mouth effect [22], the propagation of news and opinion [18], and distributed problem-solving [21]. The principle underlying how information diffuses from person to person mimics the spread of an epidemic or a rumour [11], expanding to a substantial proportion of the population in a short number of steps according to the 'small-world' phenomenon [30].

In our recent work we analysed many aspects of information flow in social networks. Firstly we have been concerned with the mathematical modelling of social selection whereby individuals tend to form relationships with others who are already similar to them. We explored two distinct dynamics that yield such phenomena: social proximity [13, 14] and social utility [15] in the context of opportunistic communication networks. Moreover we have investigated the process of social influence that leads people to adopt behaviours exhibited by those they interact with. We have analysed the impact of the logical structure of social networks on the spread of rumours [12]. We have also investigated the dynamics of the word-of-mouth effect through networks of people [1].

The two forces of social influence and selection are common place in a wide range of social settings: when deciding to adopt an activity individuals imitate the people they are currently interacting with; and they simultaneously form new interactions as a result of their existing activities. These two aspects of social networks are crucial to our understanding of how people form and update their opinions by interacting with a small sample of the population, namely their neighbours.

## 3 MICRO-SOCIAL SYSTEMS

Based on this background, in this section, we give an informal definition and brief discussion of the concept of micro-social systems.

The motivation for micro-social systems stems from the commonplace observation that networked computing devices are the driving force of modern industry, entertainment and commerce, providing powerful, and inter-connected, infrastructures for business organizations (through eCommerce, holonic manufacturing and agile enterprises), computer-mediated communication, and mobile ad hoc networks (i.e. MANETs, vehicular ad hoc networks (VANETs), etc.).

Research in networked computing also needs to take into consideration the three following features:

- Local information, partial knowledge and inconsistent union: What each network node 'sees' is the result of actions by (possibly millions) of actors, some of which are not known, and even those actions which are known, the actors motive may be unknown. Moreover, what a node 'thinks' it 'sees' may not be consistent with the 'opinion' of other nodes.
- Decentralised control: there is no single central authority that is controlling or coordinating the actions of others. The emphasis is on local decision-making based on locally available information and the perception of locally witnessed events.
- Social organization: from this it follows that in the absence of perfect knowledge there is no perfect form of government, therefore the next best thing is a government prepared to modify its policies according to the needs, requirements and expressed opinions of its 'population'.

In other words, social organization is both the requirement for and consequence of any networked computing which impacts on personal, legal or commercial relationships between real-world entities (people or organizations). The social sciences, together with legal and organizational theory, provide rich models of social organization, from which we can 'cherry pick' appropriate concepts which can be formalised in corresponding logical and computational models. These computational models are, to some extent, a scaled-down or simplified version of the original social system or theory – hence micro-social system.

Therefore we define (informally) a micro-social system as a distributed computer system or network where the interactions, relationships and dependencies between components is a microcosm of aspects of a human society. We will use the term 'agent' to denote one of these components, both to indicate membership of such a social system (e.g. as a node in a network) and in the computational sense of 'software agent', i.e. as an encapsulated, embedded, autonomous process responsible for its own state and decisions.

The aspects of human society in which we are interested includes communication protocols, organizational rules and hierarchies, network structures, inter-personal relationships, and other processes of self-determination and self-organization. In particular, though, we pick out the following three primary set of rules which underpin the social intelligence required to realise a micro-social system which can be applied to the issues in ad hoc networks detailed earlier (i.e. dynamism, conflicts, sub-ideality, security and continuity):

- Rules of Social Order. Micro-social systems (MSS) consist of agents whose actions have a conventional significance (according to the social rules of an institution); actions are therefore norm-governed. This requires characterising the permissions, obligations and (institutional) powers of each agent to determine which actions are valid, meaningful, legal, and so on.
- Rules of Social Choice. MSS consist of heterogeneous, self-interested agents that can have conflicting preferences in decision-making situations; these preferences can be aggregated by taking votes over potential outcomes. In practice, an election is held, and the winning candidate is declared to be the agreed choice.
- Rules of Social Exchange. MSS, being both open and local, will require agents to gain knowledge over time by exchanging information with each other. Each agent must therefore be capable of reliable opinion formation, based on the opinions gathered from the contacts in their own social networks. Processes of belief revision, belief merging, judgement aggregation and truth tracking are therefore important.

The development of a micro-social system is illustrated in Figure 1. We start from first principles, and analyse specific aspects of existing networked societies, i.e. human social systems, namely rules of social order, social exchange and social choice (i.e. rules, interactions and group decision-making). From this we derive the formal framework for specification of micro-social systems, which we apply to design, implement and run a computational system. Moreover, we aim visualise the operation of such systems in the same, human-understandable terms (There remains the possibility of using micro-social systems to simulate human societies, but we do not address this topic further in the current work.)

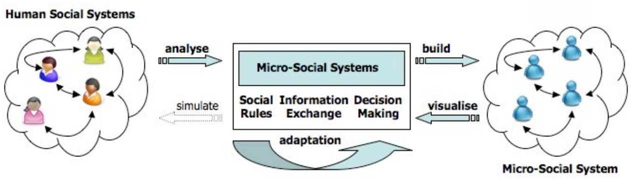


**Figure 1.** Developing micro-social systems

One of the key aspects of this development cycle are the mechanisms for run-time (self-)adaptation of the micro-social system. In particular, we seek to go beyond emergent behaviour seen, for example in swarm intelligence for emergent systems, i.e. the non-introspective application of hard-wired local computations, with respect to physical rules and/or the environment, which achieve unintended or unknown global outcomes, to the social intelligence for micro-social systems, which involves the introspective application of soft-wired local computations, with respect to physical rules, the environment and conventional rules, in order to achieve intended and coordinated global outcomes.

This is illustrated in Figure 2. Note that the formation of an organization involves hierarchies of groups, each with their own roles (chair, head, etc.), conventional rules (institutional regulations), constituency, remit (as defined by some ontology), and so on.

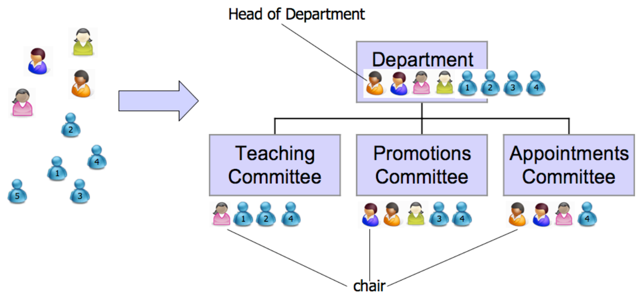


**Figure 2.** Formation and Adaptation of Organizations

Many other issues are raised by this form of self-organization, but in particular, we need to understand: firstly, how an arbitrary collection of agents can self-organise into an organisational structure; how an arbitrary collection of agents can self-organise its social network (which is structurally distinct from the organizational structure); and thirdly, what is the interplay between the explicit formal organization and the implicit social network. We illustrate this issue in the next section, with a discussion of a scenario, developing an adaptive security model for mobile ad hoc networks.

## 4 ADAPTIVE NETWORK SECURITY

In this section, we discuss an application scenario for micro-social systems, in the context of developing an adaptive network security policy for mobile ad hoc networks (MANETs).

### 4.1 Problem Description

Users access networked information and services from a variety of mobile devices over a variety of ad hoc networks. The devices have a wide range of characteristics, from processor speed to battery power; while the networks are open and volatile, implying that the network nodes cannot necessarily be trusted, either to be present or even to operate correctly. For applications to operate in a non-trusted computer network, there are many different security mechanisms, which differ in terms of cost, algorithmic complexity, types of license, and so on. Similarly there are many proposed trust/reputation frameworks, which also vary in the computational complexity of computing the trust decision. In a resource-constrained, i.e. power challenged, computing environment, It is essential to choose the most appropriate security mechanisms to match the features of content sensitivity, service delivery, device typology, and network topology.

This choice is too fast, frequent and complex for users: they will simply ignore security altogether, or worse default to the strongest level of security, to the detriment of resource utilisation, network throughput and quality of service. Instead, the choice should be delegated to the devices and applications themselves; but clearly there is a trade off involved, between algorithmic complexity (and so power cost), and the level of security and/or the (potential) accuracy of the trust decision. We propose to implement this choice via an Adaptive Network Security (ANS) scheme using the framework of micro-social systems. In the next two sections, we look at two models, one for rules of social exchange, and another for rules of social order (specifically for adaptation).

### 4.2 Social Exchange

In specifying a model of social exchange, we start from social science theories, such as [16, 17], which we formalise using graph theory, and represent ideas of influence, opinion and affinity to define a formal model. The properties of this model, e.g. in polarisation, fragmentation, and consensus, etc., are investigated in [28].

#### *4.2.1 Social Network*

We define the agents as the nodes and *confidence relations* between agents as the edges of a social network. We define a social network at timepoint $t$ as a directional, weighted graph, $G = (N, R)_t$, consisting of a non-empty set of agents $N$ of cardinality $n$, and an incidence relation $R \subseteq N \times N \times [0, 1]$.

Each agent is characterised by:

**A task:** purpose for which the agent was designed. A task defines what the agent should achieve while being part of the system.
**A strategy:** action(s) which the agent follows in order to fulfil its task.
**A mind-set:** hides the *true preference* of the agent about the issue under discussion.
**An opinion:** communicates the *expressed preference* of the agents about the issue under discussion.
**A level of self-confidence:** confidence that an agent has in its own opinion. This value might change as new opinions are received.
**A level of confidence on others' opinions:** weights the relation that an agents has in each of its acquaintances.

#### *4.2.2 Modeling Confidence*

From $R$, for any given timepoint $t$, we can derive a family of confidence functions $W = \{w_1, w_2, \ldots, w_n\}$ where each function $w_i$ is of the form $w_i : N \times T \to [0, 1]$. Accordingly, each confidence function is expressed as $w_i(j, t)$. For simplicity, we use the reduced form $w_{i,j}(t)$ to express the confidence (function) that agent $i$ assigns to agent $j$ at time $t$.

Thus, each confidence function $w_{i,j}(t)$ assigns a real value between 0 and 1 to the confidence relation between the ordered pair $\langle i, j \rangle$, indicating how much confidence $i$ has in $j$ at a specific time point $t \in T = \{0, 1, \ldots\}$. When $j = i$ the confidence function $w_{i,i}(t)$ yields a measure of *self-confidence*.

Following the well-established tradition in the field of opinion formation (see e.g. Friedkin [17] and Hegselmann [20]), we consider that for each agent the sum of the confidence in its acquaintances is always 1 (including itself), $\sum_{j=1}^{n} w_{i,j}(t) = 1$. This ensures that the measure of confidence is not based on absolute inter-agent judgements but on relative intra-agent ones, and each agent's relative confidence in the others increases and decreases only with respect to its own value judgements.

#### *4.2.3 Modeling Opinion Formation Dynamics*

Starting from the simplest assumption that agents are able to communicate very simple, concrete pieces of information with one another, and that one agent might receive this information from different sources, we are concerned with the problem of aggregating information from many sources into one.

The type of information that we deal with in this study is that of subjective nature. We conceptualise information as "information-as-knowledge" as defined by Buckland in [6]: "A key characteristic of "information-as-knowledge" is that it is intangible: one cannot touch it or measure it in any direct way. Knowledge, belief and opinion are personal, subjective, and conceptual". We refer to "information-as-knowledge" as opinions.

Each agent holds information in the form of opinions, $o_i : T \to [0, 100]$. We adopt a continuous opinion approach, in line with [10, 17, 20], and consider an agent $i$'s opinion at time $t$, $o_i(t)$, as a real-valued representation of an opinions (rather than a specific measure).

We assume that each agent holds an initial opinion (i.e. $o_i(1) \neq \perp$) on every issue about to be discussed. However this opinion can change with time as agents are influenced by opinions exchanged with other agents. The influence that one agent's opinion exerts on another's is given by how much confidence, relative to other agents, the latter has in the former.

Correspondingly, the opinion formation dynamics consists of simultaneous opinion exchanges between pairs of agents and a subsequent individual opinion revision. The main objective of an agent is to collect opinions from other agents in order to revise (i.e. consolidate or modify) its own opinions. A secondary objective is to share its own opinions and influence other agents towards them.

The opinion formation dynamics occurs at discrete time points and on a per issue basis. At each time point each agent exchanges opinions with other agents. An agent $i$'s opinion changes at time $t+1$ by weighting each received opinion at time $t$ with the confidence in the corresponding source (including its own opinion weighted by its self-confidence) such that:

$$o_i(t+1) = \sum_{j=1}^{n} w_{i,j}(t) o_j(t) \quad (1)$$

### 4.2.4 *Modeling Affinity Between Agents*

We specify the matching between an agent $i$'s mind-set and another agent's opinion by defining an *affinity function* $a_i : N \times T \rightarrow [0,1]$. This function evaluates the linear similarity between an opinion and a given constant $\mu$ which is a representative reference value of an agent's *mind-set* for a given issue. Correspondingly, we express the affinity function as $a_i(j,t)$. Again for simplicity, we denote this affinity (function) between $i$'s mind-set and $j$'s opinion as $a_{i,j}(t)$ and we define it as:

$$a_{i,j}(t) = 1 - \frac{|o_j(t) - \mu_i|}{\max(\mu_i, 100 - \mu_i)} \quad (2)$$

$\mu_i$ is agent $i$'s mind-set, which is constant per issue per agent, although no two agents need to have the same mind-set on any issue. Thus in each time step, the *affinity* between agents can be different for each ordered pair of agents corresponding to the fitness between opinions and *mind-sets*.

### 4.2.5 *Modelling Social Influence*

One of the main characteristics of our model is that we assume that agents rely differently on other agents. Thus agents can have more confidence in some agents than others and this can change with time. Agents start the opinion formation protocol with an initial value of confidence in each of the agents within their social network. As the first opinion exchange happens and the opinion of each agent changes according to equation 1, the confidence changes accordingly.

In our model, an agent $i$ increases its confidence in another agent $j$ based on how well $j$'s opinion coincides with $i$'s *mind-set*. Assuming a positive evaluation for those opinions matching agent $i$'s *mind-set* and a negative for those contradicting it, then it can be said that the confidence in an exchange partner $j$ increases as $j$'s opinion matches $i$'s *mind-set*.

Therefore, confidence changes in time differently for each agent, based on the *affinity* between agents. Agents increase the confidence in those agents whose opinions fit their *mind-set*. Thus the confidence in other agents is redistributed according to the following equation:

$$w_{i,j}(t+1) = \frac{w_{i,j}(t) + w_{i,j}(t) a_{i,j}(t)}{\sum_{k=1}^{n} (w_{i,k}(t) + w_{i,k}(t) a_{i,k}(t))} \quad (3)$$

Equation 3 contains the principle that an increase in confidence in those agents with an initially given high confidence value is higher than in those with an initially given low confidence value.

## 4.3 Social Order

We can now use the rules of social exchange in conjunction with rules of social order as the basis for adaptation, providing mechanisms with which the specification of a micro-social system may be modified during its execution (cf. [3]).

System participants may require to change at run-time their institutional powers, permissions, obligations, rights, etc. We envisage a system architecture composed of several protocol levels: the protocol at the bottom level (0-level protocol) corresponds to the procedure for conducting the business of the system – for example, in a MANET the 0-level protocol could be a security protocol.

At any point in time during the execution of the 0-level protocol the participants may start a 'meta' protocol in order to modify the rules of the 0-level protocol. A meta-protocol can be a procedure for voting, that is, the system participants vote for or against a proposed modification in the rules of the 0-level protocol. Moreover, the participants of the meta protocol may initiate a meta-meta protocol to modify the rules of the meta protocol, or they may initiate a meta-meta-meta protocol to modify the rules of the meta-meta protocol, and so on. In general, assuming an infrastructure of maximum $k$ protocol levels, the protocol participants of a $n$-level protocol $(0 \leq n \leq k)$ may start a $n+m$-level protocol $(1 \leq m \leq k)$ in order to modify the rules of the $n+m-1$-level protocol (see Figure 3).

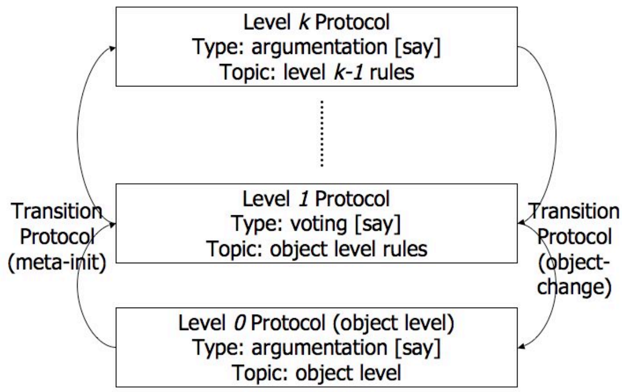


**Figure 3.** Protocol Stack for Rules of Social Order

The key problem is to judge precisely when to start the 'meta' protocol. Recall, that in a power-challenged environment, there is scant value in each agent re-assessing the 'fitness' of the network given the current security policies and protocols vs. the power-cost to implement them. Rather, we use the rules of social exchange to filter the opinion "it is necessary to change" or "it is not necessary to change" through the network. For example, the security policy and protocols at the base level (level 0) may be traded off against the available power and/or the trustworthiness of the nodes in the network: at low power but with fully trusted nodes, a low-level security policy may be appropriate; at 'high' power with a risk that some nodes may have been compromised by some network attack, a higher-level security policy may be required.

Depending on the population profile and the model of social exchange, as defined in the previous section, at some point, if an opinion leader starts to advocate the opinion that "it is necessary to change", this will be disseminated throughout the network (in a computable time), will modify the opinions of other agents, the pressure of which (i.e. the current state being $X$ and enough agents 'think' (express the opinion) that $X$ is unsatisfactory), and eventually induce a self-confident agent to initiate the meta-protocol.

## 5 SUMMARY AND CONCLUSION

We have introduced the concept of *micro-social systems*, scaled down agent societies based on a computational microcosm of human societies. Micro-social systems are an attempt to bring together principles of social intelligence – including rules of social order and rules of social choice – with principles of social networks – including rules of social exchange, and other factors such as proximity, influence, and utility.

We propose to use micro-social systems as a formal, computational framework on which to design and specify agent societies for various types of ad hoc network, whether computer-mediated communication, mobile networks, or virtual organizations. We considered a scenario involving an adaptive network security scheme, in which rules of social order and rules of social exchange were interleaved to prompt proposals for run-time self adaptation. This self-organization could be applied to develop an adaptive network security scheme, for example by changing the security protocols and procedures to suit the available power and trustworthiness of the nodes. A more substantive model than that presented here would accommodate a wider range of security policy conflicts, e.g. following the comprehensive classification of [19].

It is evident then that this work is at a formative stage and there are many details yet to be worked out in the interleaving of rules of social order and rules of social exchange, not to mention rules of social choice. When this is formulated, we can start applying the framework to deal with other issues, in particular conflicts. Conflicts are inevitable in open systems; however experience in (human) legal and commercial systems shows that those systems which facilitate internal resolution of conflicts (i.e. without recourse to time-consuming and expensive litigation) emerge with stronger social and inter-personal relationships. Therefore one immediate avenue of further research is the interleaving of rules of social order for argumentation and conflict resolution protocols [5, 27] with rules of social exchange used in juries to reach consensus and decisions. Rules of social choice are also of particular significance here, to prevent manipulation of the outcome.

In terms of experimentation, we propose to use the PreSAGE multi-agent simulation platform [24], which provides substantial support for investigating the effect of agent interactions, network properties and organizational rules on individual agent behaviour and long-term collective global performance.

## ACKNOWLEDGEMENTS

We would like to thank Alexander Artikis (Demokritos, Athens) for his contribution to the research.

## REFERENCES


[1] H. Amini, M. Draief, and M. Lelarge, 'Marketing in a random network', in *Springer LNCS, Network Control and Optimization*, p. To appear, (2008).

[2] A. Artikis, L. Kamara, J. Pitt, and M. Sergot, 'A protocol for resource sharing in norm-governed ad hoc networks', in *Second International Workshop on Declarative Agent Languages and Technologies (DALT'04)*, eds., J. Leite, A. Omicini, P. Torroni, and P. Yolum, volume 3476, pp. 221–238, (2005).

[3] A. Artikis, D. Kaponis, and J. Pitt, 'Dynamic specifications of norm-governed systems', in *Multi-Agent Systems: Semantics and Dynamics of Organisational Models*, ed., V. Dignum. IGI Global, (2009).

[4] A. Artikis and J. Pitt, 'Specifying open agent systems: A survey', in *Engineering Societies in the Agents World: Proceedings ESAW08*, eds., G. Picard, L. Vercouter, and A. Artikis, volume LNAI, p. to appear. Springer Verlag, (2009).

[5] A. Artikis, M. Sergot, and J. Pitt, 'Specifying norm-governed computational societies', *ACM Transactions on Computational Intelligence*, **10**(1), to appear, (2009).

[6] M.K. Buckland, 'Information as thing', *Journal of the American Society for Information Science*, **42**(5), 351–360, (1991).

[7] H. Carr and J. Pitt, 'Adaptation of voting rules in agent societies', in *Proceedings AAMAS Workshop on Organised Adaptation in Multi-Agent Systems*, eds., G. Vouros, A. Artikis, K. Stathis, and J. Pitt, volume LNAI5368, p. to appear. Springer Verlag, (2008).

[8] H. Carr, J. Pitt, and A. Artikis, 'Peer pressure as a driver of adaptation in agent societies', in *Proceedings AAMAS Workshop on Organised Adaptation in Multi-Agent Systems*, eds., A. Artikis, G. Picard, and L. Vercouter, p. to appear. Springer Verlag, (2008).

[9] Y. Chevaleyre, U. Endriss, J. Lang, and N. Maudet, 'A short introduction to computational social choice', *SOFSEM: Theory and Practice of Computer Science*, 51–69, (2007).

[10] Guillaume Deffuant, David Neau, Frédéric Amblard, and Gérard Weisbuch, 'Mixing beliefs among interacting agents', *Advances in Complex Systems*, **3**(1-4), 87–98, (2000).

[11] P. Dodds and D. Watts, 'Universal behavior in a generalized model of contagion', *Physical Review Letters*, **92**, 218701, (2004).

[12] M. Draief, 'Epidemic processes on complex networks', *Physica A: Statistical Mechanics and its Applications*, **363**(1), 120–131, (2006).

[13] M. Draief and A. Ganesh, 'Efficient routing in poisson small-world networks', *Journal of Applied Probability*, **43**(3), 678–686, (2006).

[14] M. Draief and A. Ganesh. Spread of epidemics and rumours with mobile agents, arXiv:0810.3128, http://eprintweb.org/S/authors/All/dr/Draief/1.* (under revision).

[15] M. Draief, A. Ganesh, and L. Massoulié. Exponential random graphs as models of overlay networks, arXiv:0810.3173, http://eprintweb.org/S/authors/All/dr/Draief/2 (under revision).

[16] J. French, 'A formal theory of social power', *Psychological Review*, **63**, 181–194, (1956).

[17] Noah E. Friedkin and Eugene C. Johnsen, 'Social influence and opinions', *Journal of Mathematical Sociology*, **15**, 193–205, (1990).

[18] D. Gruhl, D. Liben-Nowell, R.V. Guha, and A. Tomkins, 'Information diffusion through blogspace', in *Proceedings of the 13th International World Wide Web Conferencee*, eds., S.I. Feldman, M. Uretsky, M. Najork, and C.E. Wills, pp. 491–501. ACM Press, (2004).

[19] H. Hamed and E. Al-Shaer, 'Taxonomy of conflicts in network security policies', *IEEE Communications Magazine*, **44**(3), 134–141, (2006).

[20] Rainer Hegselmann and Ulrich Krause, 'Opinion dynamics and bounded confidence: models, analysis and simulation', *Journal of Artificial Societies and Social Simulation*, **5**(3), (2002).

[21] M. Kearns, S. Surit, and N. Monfort, 'An experimental study of the coloring problem on human subject networks', *Science*, **313**, 824–827, (2006).

[22] J. Leskovec, L. Adamic, and B. Huberman, 'The dynamics of viral marketing', in *Proceedings of the 7th ACM Conference on Electronic Commerce*, eds., J. Feigenbaum, J.C-I.Chuang, and D.M.Pennock, pp. 228–237. ACM Press, (2006).

[23] B. Neville and J. Pitt, 'A computational framework for social agents in agent mediated e-commerce', in *AAMAS Trust Workshop*, ed., R. Falcone, pp. 83–91, (2004).

[24] B. Neville and J. Pitt, 'Presage: A programming environment for the simulation of agent societies', in *Proc. AAMAS Workshop on Programming Multi-agent Systems (ProMAS)'08*. Springer Verlag, (to appear).

[25] E. Ostrom, *Governing The Commons: The Evolution of Institutions for Collective Action*, Cambridge University Press, 1990.

[26] J. Pitt, L. Kamara, M. Sergot, and A. Artikis, 'Voting in multi-agent systems', *Computer Journal*, **49**(2), 156–170, (2006).

[27] J. Pitt, D. Ramirez-Cano, L. Kamara, and B. Neville, 'Alternative dispute resolution in virtual organizations', in *Proceedings ESAW 2007*, eds., A. Artikis, G. O'Hare, K. Stathis, and G. A. Vouros, volume 4995 of *Lecture Notes in Computer Science*, pp. 72–89, (2007).

[28] D. Ramirez-Cano and J. Pitt, 'Follow the leader: Profiling agents in an opinion formation model of dynamic confidence and individual mind-sets', in *Proceedings of the 2006 IEEE/WIC/ACM International Conference on Intelligent Agent Technology*, pp. 660–667, (2006).

[29] E. Rogers, *Diffusion of Innovations*, Free Press (4th Edition), 1995.

[30] J. Travers and S. Milgram, 'An experimental study of the small world problem', *Sociometry*, **32**, 425–443, (1969).

[31] A. Vasalou, J. Pitt, and G. Piolle, 'From theory to practice: Forgiveness as a mechanism to repair conflicts in cmc', in *Proceedings iTrust 2006*, eds., K. Stølen, W. Winsborough, F. Martinelli, and F. Massacci, volume LNCS3986, pp. 397–411. Springer, (2006).

# The emergence of shared social representations in complex networks

**Davide Donetto** [1] and **Federico Cecconi** [2]

**Abstract.** We introduce a model based on an assimilation algorithm to study the emerging of a shared social representation in a virtual community structured on a scale-free network. Individual representations are modelled as randomly extracted bit strings and evaluated in terms of fitness, coherence with respect to underlying culture. Simulation dynamic is based on this fitness value and on most represented opinions in the agent's immediate social environment. In a short time the virtual population converges towards a shared representation and the average fitness remains stable around middle values.

## 1 INTRODUCTION

The theorization by the social psychologist Serge Moscovici about *social representations* origins from a study about diffusion of psychoanalytic theory beyond the bounds of official circles and towards a much more vast audience [12]. The author highlights how, in this transition from a scientific context to a *naïve* one, the theory undergoes changes that reflect the differences between the two fundamental ways of approaching reality that deeply characterize western thought and the different constraints they pose. In the shift from a *reified universe* to a *consensual* one, to use the author's terminology, the original strength of internal logical coherence gets lost but what is created through informal interactions between individuals is a much more understandable view, an almost tangible experience: a shared social representation. These constructions have the function to cope with the unfamiliar, be it an object, an event, a theory or whatever, bringing it back to the soothing dimension of familiar reality [13].

In this work we try to simulate the emerging of a shared social representation in an agent-based model built upon the structure of a complex network. Agents will be driven by the need of giving sense to an hypothetical sudden event to get to a satisfactory representation of it through social interactions and opinions exchange.

## 2 RELATED WORKS

The themes of cultural assimilation and opinion spreading through social interactions in artificial societies have been addressed in many other studies before. Moscovici himself relied on computer simulation in the study of group polarization [8, 7]. Opinion shifting has been modeled by Nowak, Szamrej and Latané [14], Hegselmann and Krause [9] and Deffuant et al. [6] and cultural assimilation by Axelrod [3], Kennedy [10] and Parisi, Cecconi and Natale [15]. All these works are based on some kind of assimilation dynamic between individuals or groups interacting with their neighbors and influencing each other through an algorithm that relies on similarity, Latané's theory of *social impact* [11], or rather Boyd and Richerson's [5] *frequency bias*.

In our model a virtual society has to cope with the need of representing an event of growing importance in the life of the community. We imagine that each individual of this community will have a tendency to reach a more structured and coherent representation (i.e. one that better fits in the underlying shared culture) and that will try to satisfy this need adopting the most represented opinions in his neighborhood. What characterizes our algorithm from a plain frequency bias one though, is the introduction of relative impact of individuals, that will be proportional to their own fitness, and variable need for change of individuals, that will be inversely proportional to fitness. As a matter of fact we could imagine that an individual with a more coherent representation could exert a stronger social influence than others and feel more satisfied with his representation, thus less willing to change it. We chose to model the structure of individuals relationships network upon a *scale-free* graph as it seems to be the topology that better represents real social networks [4]. This class of networks shares the two fundamental properties of the so called *small-world* networks, namely high clustering coefficients (i.e. the neighbors of two linked nodes form two largely overlapping groups) and low diameter value (i.e. the average shortest distance between two nodes increases logarithmically with the number of nodes). But the main property of scale-free networks — the one to which their name is due — is that their degree distribution follows a power-law so that a few nodes, the *hubs*, will have many links while the vast majority will be poorly connected. In our simulation we chose not to consider the additional factors that may intervene in real social networks formation thus influencing their resulting topology, such as *aging* or *linking costs* [2], because we felt they were not relevant to the study of the specific topic of social representations emergence.

## 3 THE MODEL

The nodes of our scale-free network represent individuals. To each one of them is assigned a binary string composed by 450 randomly extracted values that stands for the representation of event $X$ of the individual $j$. Each string is interpreted in blocks of 15 elements through *gray code* in a vector of real values. This coding is built so that two consecutive real numbers are represented by strings that differ from one another for just one bit. In our case this means that the changing of a string element will produce little incremental changes in the position of that string in the fitness hyperspace.

The 30 real values will then be transposed in the interval $[-600, +600]$ and used as coefficients of the *Griewank* function:

[1] Dept. of Relational Sciences, Univ. of Naples Federico II, Via Porta di Massa, 1, 80133, Naples, Italy. Email: davide.donetto@unina.it
[2] ISTC - CNR, Via S. Martino della Battaglia, 44, 00185, Rome, Italy. Email: federico.cecconi@istc.cnr.it

$$C = \sum_{i=1}^{n} \frac{x_i^2}{4000} - \prod_{i=1}^{n} \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$$

The resulting fitness index C will be interpreted as the coherence of the representation in its cultural context (that we suppose as static during the time of the simulation). The Griewank function is a widely used benchmark tool in the context of global optimization but two of its properties that we found particularly fit to our simulation were determinant in our decision to rely on it:

- a subtle change in a string may cause an even strong fitness variation;
- two (or more) different strings may have the same fitness score.

In setting up the network we relied on the algorithm proposed by Albert and Barabási [1, p. 71] based on *growth* and *preferential attachment*: we started with a small number of nodes $m_0$ and step by step we added a new node with $m \leq m0$ connections linking it to $m$ different nodes with a probability $\Pi$ to be linked to the node $i$ proportional to its degree $k_i$ so that:

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}$$

We repeated the procedure until we obtained a 100 nodes network and then we checked that the nodes degree distribution followed a power-law curve. Then we computed for each node the initial string fitness value C and started the dynamic of the simulation that went on for 100 cycles.

## 3.1 Simulation Dynamic

At each time step for each node we will select a number of binary elements of its string, based on its fitness value, to submit to the revision process. This number will be inversely proportional to $C_{MAX}$, that for a 30 values vector like ours will be about 2701. Which elements will be revised is randomly decided so that the probability remains constant for each element (i.e. an element may be selected more than once).

The revision process of a node string element consists of three different phases (Figure 1): first of all the values (0 or 1) of respective elements in directly linked nodes strings are read; then the category (0 or 1) each of them represents is given a score between 0 and 1 proportionally to the fitness of the string it belongs; finally the element is accorded to the category with the higher score. Should categories end up with the same score the element value will be randomly extracted.

It's worth noting that in any case this procedure does not assure the extracted element value will change. Any change however is stored in a temporary buffer and the strings are actually updated only after all nodes have been processed.

At the end of each cycle we collect three parameters:

- average strings fitness
- maximum strings fitness
- dispersion (average hamming distance between strings)

To test the network topology impact we decided to run the same simulation on a random graph with a connection probability $p_c$ = 0,1 between each possible pair of nodes. As graph construction is based on different arbitrary parameters ($m$ and $m_0$ for the scale-free graph and $p_c$ for the random graph) no guarantee is given that resulting

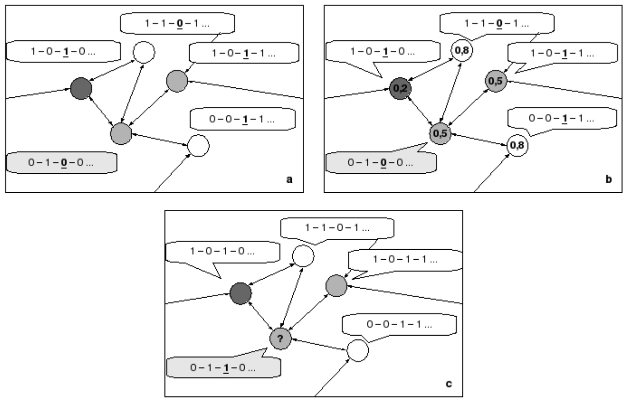


**Figure 1.** Revision process phases: a) an element from processed node string (in grey) is extracted and respective elements in directly linked nodes strings are read; b) the category each element represents (0 or 1) is added a value between 0 and 1 proportionally to the fitness of the string it belongs (showed as the tone of grey of the node); c) the extracted element is accorded to the most represented category.

architectures will be comparable. To maximize validity we chose to build the random graph before the scale-free one so that we could try to get the same number of total connections for the two graphs by choosing appropriate values for $m$ and $m_0$.

Simulations have been run for different values of $p_c$ and with the introduction of different percentages of mutations to test the impact of these factors on the assimilation dynamic. Mutations consist of value re-extraction of up to the chosen percentage of randomly selected binary values for each string at each time step. They could account for changes in individuals representations not specifically due to social interactions. Each experimental setup has been tested 50 times with different randomly generated starting conditions.

# 4 RESULTS

In the base setup, with no mutations and $p_c$ = 0,1, dispersion decreases quickly, yet not reaching zero, in both topologies as can be seen in Figure 2.

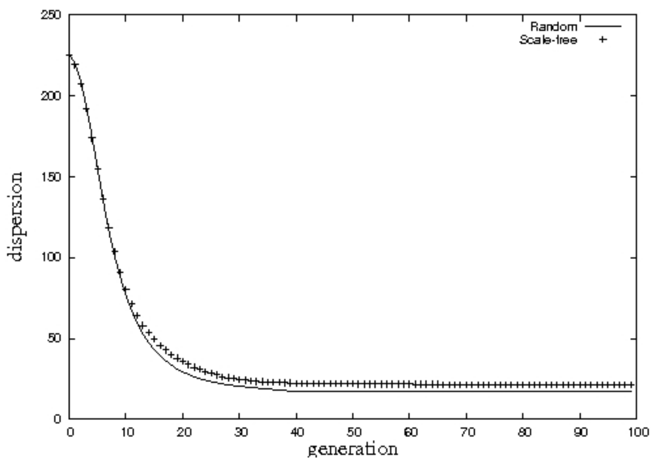


**Figure 2.** Dispersion through time in the two considered topologies: random graph (line) and scale-free graph ("+" points), with $p_c$ = 0,1

We applied the Student's *t-test* to dispersion values of last simulation cycle for the two different graphs and the result was a significant difference for $p < 0{,}05$. More in detail for the scale-free graph we had an average value of about 15,7804 and a standard deviation of about 7,7786, while for the random graph we had an average value of about 13,1765 and a standard deviation of about 3,7111. The $t$ value was consequently of about 2,1363.

Increasing the number of connections though, we noted that this difference started decreasing so that with $p_c = 0{,}2$ we found no more a significant result.

Adding a little percentage of mutations (1%) to the base setup we had an even stronger decreasing of dispersion and a more significant difference (Figure 3): average value for last cycle dispersion in the scale-free graph over 50 replications was about 10,7980 with a standard deviation of 4,6018, while in the random graph it was about 8,0026 with a standard deviation of 2,5320. The $t$ value was about 3,7635 and so the probability that the difference between topologies could depend on chance was less than 0,01 percent.

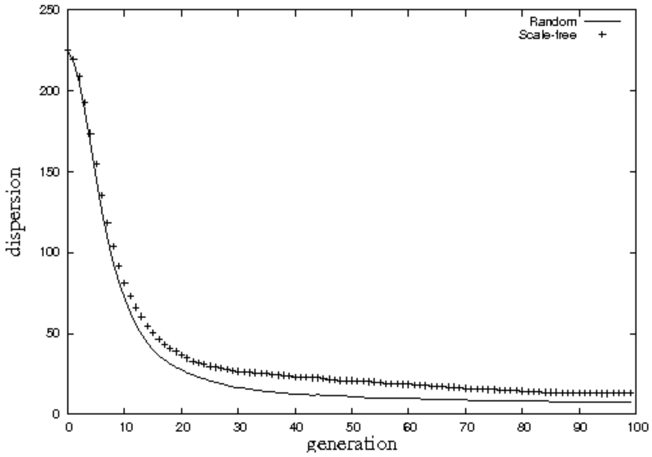


**Figure 3.** Dispersion through time in the two considered topologies: random graph (line) and scale-free graph ("+" points), with $p_c = 0{,}1$ and mutations at 1%

Even in this case the increasing of $p_c$ had the effect of hiding the effect of topology in the assimilation dynamic. The same result was given by the introduction of an higher percentage of mutations (5%).

A result that remains constant through conditions is the fitness trend that in a few generations reaches an intermediate value and tends to stay stable throughout time (an example is reported in Figure 4).

## 5 DISCUSSION AND FUTURE WORK

We formalized the social construction of a representation of an element of novelty as a process that requires finding a satisfactory solution to a complex equation. The less coherent, with respect to the individual's view of reality, this solution will be the more he will feel the urge to come to a different solution and integrate the new element in a more stable structure. Moreover we imposed that the only way to reduce this angst would have been social exchange of opinions.

In all experimental conditions, based on these premises, we assisted to the emerging of a shared representation. We were surprised that the assimilation dynamic, though being subjected to the pressure towards high coherence, did not succeed in taking the fitness up. We would have expected it to behave like a genetic algorithm, instead it showed a peculiar course that deserves further in-depth studies.

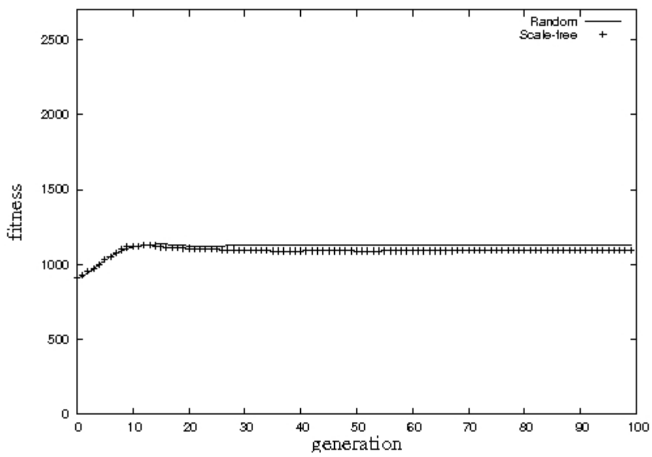


**Figure 4.** Average fitness through time in the population for the random graph (line) and the scale-free graph ("+" points)

However this result is perfectly in line with Moscovici's theorization which postulates, in the process of construction of a social representation, the sacrifice of a certain amount of internal coherence in favour of the probably more essential dimension of sharing.

## ACKNOWLEDGEMENTS

We would like to thank the ESF-EUROCORES-programme TECT (The Evolution of Cooperation and Trading) for supporting our work.

## REFERENCES

[1] Reka Albert and Albert L. Barabasi, 'Statistical mechanics of complex networks', *Reviews of Modern Physics*, **74**(1), (2002).

[2] Luis A. Nunes Amaral, Antonio Scala, Marc Barthelemy, and Harry Eugene Stanley, 'Classes of small-world networks', *Proceedings of the National Academy of Sciences of the United States of America*, **97**(21), 11149–11152, (2000).

[3] Robert Axelrod, 'The dissemination of culture: A model with local convergence and global polarization', *Journal of Conflict Resolution*, **41**(2), 203–226, (April 1997).

[4] Albert-Laszlo Barabási and Eric Bonabeau, 'Scale-free networks.', *Scientific American Magazine*, **288**(5), 60–69, (May 2003).

[5] Robert Boyd and Peter J. Richerson, *Culture and the Evolutionary Process*, University Of Chicago Press, June 1988.

[6] Guillaume Deffuant, Frédéric Amblard, Gérard Weisbuch, and Thierry Faure, 'How can extremism prevail? a study based on the relative agreement interaction model', *Journal of Artificial Societies and Social Simulation*, **5**(4), 1, (2002).

[7] Serge Galam and Serge Moscovici, *Compromise versus polarization in group decision making*, 40–51, Defense Decision Making, Springer-Verlag, 1991.

[8] Serge Galam and Serge Moscovici, 'Towards a theory of collective phenomena: Consensus and attitude changes in groups', *European Journal of Social Psychology*, **21**(1), 49–74, (1991).

[9] Rainer Hegselmann and Ulrich Krause, 'Opinion dynamics and bounded confidence: models, analysis and simulation', *Journal of Artificial Societies and Social Simulation*, **5**(3), (2002).

[10] James Kennedy, 'Thinking is social: Experiments with the adaptive culture model', *Journal of Conflict Resolution*, **42**(1), 56–76, (February 1998).

[11] Bibb Latané, 'The psychology of social impact', *American Psychologist*, **36**(4), 343–356, (April 1981).

[12] Serge Moscovici, *La Psychanalyse, son image et son public*, Presses Universitaires de France - PUF, February 1976.

[13] Serge Moscovici, *The phenomenon of Social Representations*, Cambridge University Press, 1984.

[14] Andrzej Nowak, Jacek Szamrej, and Bibb Latané, 'From private attitude to public opinion: A dynamic theory of social impact.', *Psychological Review*, **97**(3), 362–376, (1990).
[15] Domenico Parisi, Federico Cecconi, and Francesco Natale, 'Cultural change in spatial environments: The role of cultural assimilation and internal changes in cultures', *Journal of Conflict Resolution*, **47**(2), 163–179, (April 2003).

# Utility Seeking in Complex Social Systems: An Applied Longitudinal Network Study on Command and Control

**Joshua Lospinoso[1], Ian McCulloh[1&2], and Kathleen M. Carley[2]**

**Abstract.** Humans are autonomous, intelligent, and adaptive agents. By adopting social network analysis techniques, we submit a framework for the study of dynamic networks and demonstrate the use of actor-oriented specifications in longitudinal networks. Through the use of a unique command and control dataset from experiments run at the US Military Academy, we illustrate the power of testing hypotheses on actor utility profiles. We frame static, covariate factors onto communication networks, and find that statistical hypothesis testing indicates edge networks truly motivate soldiers to seek information, collaborate, and modify the social network around them into more comfortable configurations of triad closure and edge reciprocity, when compared to hierarchical networks: a finding with profound implications to the study of complex, adaptive social systems.

## 1 INTRODUCTION

Multi-agent simulation is rapidly emerging as a popular tool for understanding complex social and organizational structures. Historically, these models have been either very simple, or have contained few agents due to issues of computational complexity. As the power of computers continues to increase rapidly, more complex multi-agent simulation models are needed. Social network analysis has become equally popular for understanding social and organizational structures. This paper applies methods in longitudinal social network analysis to multi-agent simulation.

Human organizations and social groups are composed of individuals. The individuals can be related in a number of different ways: friendship, trust, ethnicity, shared ideology, shared goals, and more. Some of these relationships are important in understanding the behaviour and actions of the organization or social group. Other relationships are unimportant. Furthermore, some relationships affect others, creating very complex dynamic behaviour.

Multi-agent simulation is used to model individual agents that can act, interact, and learn. The agents exist in an environment where their interaction is constrained by their position in various social networks defined by the aforementioned relationships among others. Group behaviour emerges as a result of the complex interaction between agents.

Understanding network structure is very important for modelling social groups and organizations in a realistic manner. For example, Valente [1] was interested in modelling the diffusion of contraceptive innovations in the Cameroon. He found that real-world adoption rates did not follow simulation models when the network relationships were ignored. An individual's decision to adopt an innovation is highly dependent on the decisions of adjacent individuals in a social network. Assumptions of random mixing of individuals, therefore, generate inaccurate adoption rates since trust and friendship networks are important factors. When the simulation accurately models the underlying social networks of people in the Cameroon, more accurate diffusion models are obtained. For a more thorough review of the diffusion of innovations, see Valente [1].

Understanding social networks is not only important for modelling diffusion processes. Social networks are important for modelling any social group or organization involving humans. Multi-agent simulation modellers should be familiar with important theories in social network analysis that govern relationships between individual agents. Incorporating some of these theories into simulation models will contribute to more realistic models.

It is also important to be able to identify what social theories are applicable to certain problems and situations. Relationships that may be important in one context may be unimportant in another. Social network analysts are able to statistically test for the significance of various social theories in longitudinal network data. Equipped with significant theories governing network formation in empirical data, the multi-agent simulation modeller can include these factors in their simulation, thereby creating more realistic agent interactions.

This paper will present a novel approach to multi-agent simulation and demonstrate it on a real-world network data set. Longitudinal network data is collected in a natural experiment focused on studying shared situational awareness and communication. An actor-oriented model [2] is fit to the data to determine significant social theories contributing to network dynamics. These theories can then be incorporated in a multi-agent simulation model to create more accurate organizational behaviour.

The paper is organized as follows. First, we describe a theory of network dynamics used in social network analysis. Next, we describe the concept of network utility. In Section 4 we describe network data collected from a natural experiment conducted at the U.S. Military Academy. Section 5 describes a longitudinal analysis of that data, with the results presented in Section 6. In Section 7, we highlight implications for multi-agent simulation modellers and provide directions for future work.

## 2 NETWORK DYNAMICS

Network dynamics is a term used in social network analysis to describe the behaviour of networks over time [3,4,5]. Social network analysts have been conducting research in this area for quite some time [2,6,7,8,9,10,11,12,13,14,15]. There are four

[1] Network Science Center, United States Military Academy, West Point, NY 10996. Email: {Joshua.lospinoso, ai6873}@usma.edu.
[2] Center for Computational Analysis of Social and Organizational Systems (CASOS), Carnegie Mellon University, Pittsburgh, PA 15213. Email: {imccullo, Kathleen.carley}@cs.cmu.edu.

behaviours that can occur in a network over time: *Stability, Evolution, Random Change,* and *Mutation.*

Network *Stability* occurs when the underlying relationships that connect agents in a network remain the same over time [15]. The observed data may contain error. Some relationships may not be observed, while some observed connections may be inadvertent and no relationship exists. Consider email communication. An agent may communicate with some friends every day, others sporadically, and they may even accidentally email someone they do not know by hitting the wrong name in a distribution list or replying to all in an email. While the observed networks may fluctuate from day to day, the underlying relationships remain unchanged. They have reached a dynamic equilibrium for at least the short term.

Network *Evolution* occurs when agent interaction over time changes the underlying relationships [3]. Furthermore, evolution assumes that there is some underlying stochastic process that causes change over time. There are two leading approaches for modelling network evolution. One general class of approach is to use Markov chains [16,17,18,19,20,21,22]. Under this approach, the network transitions from one network state to the next over time. The future state of the network is conditioned only on the current time step and not previous time steps. Research has focussed on the structure of the transition matrix that governs the evolution of the networks.

An alternate approach for modelling network evolution is multi-agent simulation [3,23,24,25,26,27]. Under this approach, agent based models are created in which agents interact according to some established social theory. Interactions allow the agents to change in some important way that may affect future interaction.

*Random Change* in a network occurs when the future behaviour of the network is independent of the current state [5]. In other words, the agent interaction is affected by something external to the network. For example, an Army platoon may evolve as individual agents interact and communicate. When that same Army platoon comes under attack by the enemy, there is something fundamentally different about their relationships. There is not anything inherent in the individual agent interactions that could have predicted the change in network behaviour as a result of the enemy attack.

It is also possible that a random change could initiate network evolution [5]. We call this type of behaviour a *Mutation.* In our Army example, it is possible that under the stress of enemy combat an individual agent displays remarkable courage or cowardice. This individual behaviour may improve or remove the status of an agent. Other agents in the network may respond differently to agent based on their actions during the random change.

One possible explanation of network dynamics is agent-driven optimization. Agents in a network attempt to optimize their utility subject to various costs and constraints. Under this concept, stability can be viewed as an equilibrium surrounding some local optima. Evolution can be viewed as the network converging on some new dynamic equilibrium. Random change is still exogenous to the network and changes the state of agents in the network. If this change results in some other local optima, then the network reaches some new stability states. Otherwise, the network experiences mutation as the network converges to a new equilibrium. This concept of agent-driven optimization is further explored in this paper as an approach for modelling complex adaptive social systems.

## 3 NETWORK UTILITY

The concept of actor-driven models for network evolution was proposed by Snijders [2,28]. Several applications of this model have been presented [29,30,31]. Snijders' concept of actor-driven models views a network from the perspective of individual agents. Each agent can control the set of outgoing links to other agents in the network. His seminal assumption is that actors perform myopic stochastic optimization in continuous time. These changes are Markovian and depend on network structure, attributes, and observed covariates.

Social network analysts use Snijders' actor-driven model to determine what pre-defined social factors are important in describing the evolution of empirical social network data. Snijders [2] defines 11 potential objective functions that have some sociological meaning:

1. The *density effect* is defined by the number of links an agent has to other agents in the network.
2. The *reciprocity effect* is defined by the number of links to other agents that are reciprocated, in that when an agent links to a target agent, that target also links back to the original agent.
3. The *transitivity effect* is defined by the number of transitive patterns among an agent's connections. A transitive pattern occurs when two of an agent's connections are connected themselves. This is also known as a transitive triplet. Transitivity follows the logic that two agents are more likely to know each other if they have a common friend.
4. The *balance effect* is defined by the similarity of outgoing links between an agent's connections. This theory is driven by the idea that there are positive and negative links and an agent is uncomfortable having both relations simultaneously. In other words the enemy of my friend should be my enemy and the friend of my friend should be my friend. If I am friends with my enemy's friend, I will feel uncomfortable. This effect is highly correlated with the density effect and transitivity effect. If both are included in a model a correction for the correlation between effects should be included.
5. The *number of geodesic distances of two effect* is defined by the number of other agents that an agent is indirectly connected to through an intermediary agent.
6. The *popularity effect* is defined as the number of links an agent has coming from other agents in the network.
7. The *activity effect* is defined as the number of other agents that can be reached by an agent in two steps.
8. The *main link effect* is a covariate effect for links in the network. The other objective functions might be weighted by certain relationships. For example, a link to an agent of high prestige or rank might be more valuable than a link to an agent with equivalent status.
9. The *related popularity effect* is a covariate effect for agents in the network. This is defined for an agent, *i*, as the sum of the popularity effect of all other agents connected to agent *i*.
10. The *related activity effect* is a covariate effect for agents in the network. This is defined for an agent, *i*, as the sum

of the activity effect of all other agents connected to agent *i*.

11. The *related dissimilarity effect* is a covariate effect for agents in the network. This is defined as the sum of the differences in some important attribute between an agent and its' direct connections.

Agents in a network can also experience constraints as well as objectives. Agents can be constrained in the number of links that they can maintain to other agents in the network. This constraint models cognitive limitations on individuals. A person is not capable of maintaining meaningful relationships with hundreds of people. Other constraints may be imposed on the agents in the network. Snijders does not consider constraints in his model to simplify computation. When estimating the effects, the density effect often has a negative coefficient. This is interpreted as an observed constraint on node degree. See Snijders [2] for a more thorough explanation. Our aim is to present considerations in multi-agent simulation based on social network analysis and not to generate a comprehensive model.

Under a network utility model, an agent will change its outgoing links in such a way as to increase its overall utility, which is equivalent to optimizing its objective function. It is important to note that the list of objective functions are suggestions and are non-exhaustive. When tested against empirical data, only a subset of the objective functions may be found to be significant. Undoubtedly, an analyst could consider other important social factors. Therefore, when using these objective functions in a multi-agent simulation, the modeller should use some intuition in determining important effects. Ideally, a modeller could record empirical data, use Snijders' actor-driven approach to determine significant objective function effects as his approach was intended, and then use those effects in a multi-agent simulation to make inference on the future behaviour of the network.

It is important to point out differences between network utility and classic game theory. Common applications of game theory intend to focus on trading scarce resources. The network utility approach does not consider the transfer of resources, rather agents attempt to optimize their position in their social network. This approach may not be common in multi-agent simulation, but it is supported in the social sciences.

## 4 DATA

Parity Communications in collaboration with the Higgins Trust Framework and the Social Physics project constructed the ELICIT software package. Installed on client computers, the software serves as the platform for studying organizational efficiency and effectiveness. The four phase experiment entails an introduction, practice round, a one hour exercise, and a wrap up. During both the practice round and the actual exercise, thirty four subjects are randomly assigned to one of two organizations: a typical hierarchically arrayed organization (C2) and a control-free, self-organizing organization (E). These two organizations operate independently for the duration of the exercises.

The goal of the organization is to identify a terrorist attack based on bits of information distributed around the organization. After ten minutes of the one hour experiment, all of the correct information has been issued to the organization. Among the correct bits of information, or factoids, are also distributed false factoids. Each agent receives four factoids, and they must corroborate within the organization to come up with the correct arrangement of who, what, where, and when. The C2 group is comprised of a squad leader, four team leaders, and twelve team members. Communications among these agents are restricted to the following graph in Figure 1:

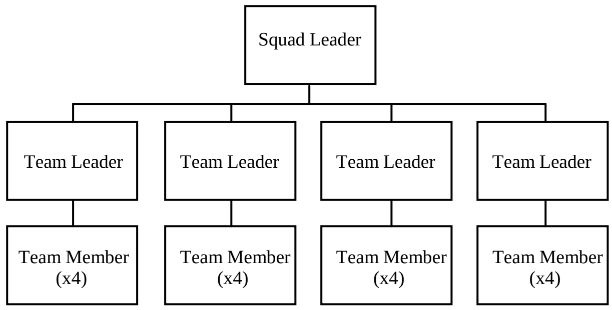


**Figure 1**. C2 Communications Hierarchy

Each team is dedicated to identifying one key element of the terrorist attack: who, what, where, and when.

The E group is comprised of seventeen agents with full communication capability across the organization. There are no defined teams, but the goal remains the same: positively identify the terrorist attack. All agents have the ability to post their information on their organization's website. Within the E group, this website is global to the organization. The C2 group has separate websites for each echelon (four teams and one squad site). The hierarchy in Figure 1 describes where each agent can post information. Agents can also share information with other individual agents. Once an agent believes that it knows any number of correct factoids, it can report its belief through the "identify" function to its immediate superior in the C2 group or to the entire network in the E group.

Data was collected on two iterations of ELICIT experiments conducted at West Point. During one iteration, the cadets were allowed to communicate within an edge-network configuration. In the other, the cadets were required to adhere to a strict hierarchy. Other than these systemic restrictions, the two iterations were run identically for an actual test run of two hours.

The human participants in this experiment were all cadets at the U.S. Military Academy between the ages of 17 and 23. The experiment was approved for ethics and safety by the West Point Institutional Review Board. All participants received a briefing on the experiment, consented to participate, and had the option to leave the experiment at any time without any adverse impacts. The investigators conducting the experiment were not in the participants' military chain of command, so no undue influence was exerted in this experiment.

## 5 METHOD

We use the social network software package SIENA [32] which implements an actor-oriented network model [2] to analyze data from two iterations of the ELICIT experiment. Adjacency matrices were constructed to reflect the structure of communication networks over time. These are un-weighted (dichotomous), directed, and non-reflexive square matrices. We must define time intervals in which to discretize or bin the data. Following the guidelines set out by Steglich and Snijders [33], we chose five bins. Each edge $e_{ijt}$ was assigned a positive value (of one) if one of two conditions was met: cadet *i* sent cadet *j* information during time bin *t*, or cadet *i* posted information on a team website sometime between the start of the experiment and time *t* which cadet *j* retrieved during time *t*.

Next, we defined covariates. This step is crucial and warrants special attention when conducting an actor-oriented model specification under the SIENA framework. Covariates are empirically derived values which are infused directly into four main objective functions (effects 8-11 above) and provide compelling parameter estimates which can potentially gain critical insight into important aspects of sociological systems. In the case of the ELICIT data, we identify two main link effect covariates corresponding to leadership and location. The *leadership-link effect* is modelled with dependence-style network. The leadership network consists of time-invariant relationships of who was in charge of whom. Note that the leadership network was completely empty for the edge-organization case, because there were no formally defined leadership roles. The statistically significant parameter estimates of the leadership-link effect indicate that formal leadership roles may play a significant part in driving agent behaviour. With low--or even negative--parameter estimates, agents in the network are averse to forming links with formal leaders. The *location-link effect* models geographical proximity. Within the ELICIT framework, geographic distance may play a significant role within the hierarchical network, since geographical locations coincide with team placements. It would seem to also be an important covariate for the agents in the edge network, since agents within the same geographical region post to the same website and are most likely to gain information from this site. The statistically significant parameter estimates of the *location-link effect* indicate a strong affinity or aversion across both the edge and hierarchical networks on the basis of team cohesion (whether enforced or not).

In addition to main link effect covariates defined on relationships between agents, we also defined a covariate for the information an agent possesses. As time progresses in the experiment, agents gain bits of information. Once an agent believes that the information is true, they will privately publish their belief to the ELICIT server, where the belief can be recorded by the experiment administrators. This is a time varying effect. We use the related popularity effect (number 9 above) to model this effect. Statistically significant parameter estimates of the *information effect* indicate that agents with more information attract more communication from other agents in the network.

We also modelled the density effect, the reciprocity effect, and the transitivity effect (effects 1-3 above), because they are commonly used in the literature. We elected to omit other objective functions to prevent over specification of the model. See Steglich and Snijders for a more comprehensive review [33].

## 6 RESULTS

In order to estimate the parameters of both the edge and hierarchical treatments simultaneously, we compiled both adjacency matrices and covariates into large matrices with structural holes where appropriate. We conducted estimation procedures within SIENA using default parameters and 1000 iterations of the three-stage Metropolis-Hastings Markov Chain Monte Carlo. Tables 1 and 2 display the parameter estimates of the E and C2 networks respectively.

| Measure | Parameter Estimate *(p-val)* |
|---|---|
| **Density Effect** | -.3693 *(.028)* |
| **Transitivity Effect** | .2054 *(.031)* |
| **Reciprocity Effect** | .1502 *(.070)* |
| **Location-link Effect** | .0513 *(.471)* |
| **Leadership-link Effect** | -- |
| **Information Effect** | .2146 *(.009)* |

**Table 1.** Parameter Estimates for Edge Network

| Measure | Parameter Estimate *(p-val)* |
|---|---|
| **Density Effect** | -.9976 *(.035)* |
| **Transitivity Effect** | .2007 *(.044)* |
| **Reciprocity Effect** | .0640 *(.36)* |
| **Location-link Effect** | .2632 *(.017)* |
| **Leadership-link Effect** | .1507 *(.023)* |
| **Information Effect** | -.1647 *(.019)* |

**Table 2.** Parameter Estimates for Hierarchical Network

We estimate six important objective functions to determine what sort of utility profiles are recurrent in each of the networks. After separating out the effects of each of the networks using individual covariate dummy variables, we find that the density effect measure is negative and statistically significant, which corresponds with our intuition that there is some sort of underlying cost to adding edges. Within the edge network, this effect is significantly diminished, which may indicated that agents in the edge network either have more cognitive capacity to form ties or that they are empowered by a lack of formal hierarchical structure. We find that the magnitude of this estimate (nearly -1) compared to the relative size of the other objective functions indicates that there are strong limitations to the cognitive capacity of the agents within the hierarchical network.

Transitivity effect has a strong and statistically significant, positive parameter estimate. Agents in both of these networks tend to close triads, which would confirm our intuition in the hierarchical network, where team members might be expected to close triads within their teams. The estimates are rather stable across the edge/hierarchical treatment, and it would appear that there is little difference between the two utility profiles.

Reciprocity effect has little effect within the hierarchical network, but it has a significant effect on the edge network. Reciprocity tells us how likely one node is to return information to the entity who sent them information. This supports our intuition that in an edge network, relationships are created on the basis of information necessity and all agents must cross-load information. Within the hierarchical network, team-leaders can ask for information and receive information without ever having to inform their teams what is going on; so the edges are not reciprocated (which is why we fail to have statistically significant results under the hierarchical network).

Location-link effect has a statistically significant effect on the parameters for the hierarchical network. This may be a result of location and team membership being highly correlated. When two agents in the hierarchical network are within a team, their team leader tasks them with determining one of the factoids, so it is natural that collaboration here should become important. Within the edge network, there is no statistically significant estimate for location. What this indicates is that within the edge network, covariates of initial team membership mean little and agents quickly breakout of their location to connect with the other locations and help contribute to their knowledge base.

Leadership-link effect was estimated for the hierarchical network and had a strong, positive estimate. This indicates that the leadership role could explain a large portion of variation in the communication patterns of the hierarchy. It both supports our intuition and supports the notion that leadership within the hierarchy was effective at promoting information sharing up and down the chain.

Information effect parameter estimates differed considerably between the edge and hierarchical treatments. Within the hierarchical network, there was actually a strong, negative correlation between people who had assembled information into some sort of conclusion and others. This means that there is information hoarding going on in the hierarchical network; the leadership is hoarding the information. Within the edge network, people who have assembled information seem to attract many edges. We cannot establish causality directly from this estimate (i.e. it could be that the entity has information because he is highly interconnected, or that he is interconnected because he has information), but it is certain that information sharing within the network is a largely significant behavioural engine.

There are some striking differences about the behaviour of these two networks. First, information sharing and collaboration occurs much more within the edge network, while leadership seems to drive much of the behaviour in the hierarchical network. Agents in the edge network tended develop sharing relationships much more than in the hierarchical network as evidenced by the high reciprocity and triad closure in the edge network. Finally, it appears that edge network agents had fewer constraints on collaboration *en masse* as indicated by the magnitude of their density effect estimates.

## 7 CONCLUSIONS & FUTURE WORK

Defence agencies of the future will increasingly rely on an understanding of complex systems. From understanding the asymmetrical nature (non-hierarchical) of armed adversaries to engineering net-centric systems that maximize efficiency and effectiveness, researchers have and will continue to benefit from empirical studies of complex systems--whether social, physical, or biological [33,34,35]. For a thorough review on this active area of research, the reader is referred to Alberts [34].

We utilized an actor-oriented specification of a complex social system as opposed to an aggregated, holistic assessment of the system, and as a result we were able to dig into the underlying behavioural mechanics of the network and truly understand what is driving the autonomous, intelligent behaviour of the cadets in the study. We now understand that soldiers within *net-centric* edge networks do collaborate across geographic and formal boundaries as expected, but more importantly--their behaviour is *driven* by the need to accumulate knowledge and settle into comfortable social patterns (like triad consensus, reciprocity, etc.).

Beyond contributing to sociological literature and the defence industry's understanding of net-centric operations and systems, this paper has introduced actor-oriented models in social network analysis which identify statistically significant utility seeking behaviour within empirical data. The study of complex, adaptive systems can benefit from this empirical framework by permitting the investigator a deep look into the underlying mechanics that drive network structure. Enabled with these tools, there is a considerable array of future directions that investigators can pursue to enrich our understanding of complex systems.

Parameter estimates from an actor-oriented specification as outlined in this paper can be used to drive a multi-agent simulation. Moreover, the approach laid out in this paper allows a modeller to use empirical data to determine factors driving agent interaction within a simulation. Building simulation based on statistically significant findings within empirical data is an important aspect of model verification.

This approach requires that multi-agent simulation frameworks are capable of modelling significant utility seeking behaviour. It is important to note that functions driving agent behaviour may differ among differing applications. In the ELICIT example, different objective functions were significant for the edge and hierarchical networks, even given highly homogeneous sets of agents. This implies that there is no one model that fits all applications.

An example of a flexible multi-agent simulation is *Construct*, which is a multi-agent simulation developed by the Center for Computational Analysis of Social and Organizational Systems [24,25]. *Construct* models agent interaction by assigning probabilities of link formation between agents at each time step. The probability of link formation is determined by a weighted function of homophily, socio-demographics, and proximity. Throughout the simulation, agents interact, share knowledge, and change in various attributes as a result of interaction with other agents. Within the framework laid out in this paper, homophily is equivalent to transitivity, reciprocity, balance, and the information effect. Socio-demographics are equivalent to the number of geodesics of two effect, the popularity effect and the activity effect as well as some covariate effects. The proximity is equivalent to a main-link effect. Other effects can be incorporated into the *Construct* model as well. While a detailed explanation of *Construct* is beyond the scope of this paper, we point out that it is an example of a multi-agent simulation framework that can be used to simulate empirically observed network data. The statistically significant parameter estimates of the actor-oriented model can be used to provide weights to the functions that determine the probability of link formation between agents. In this manner, the predictive power of the multi-agent simulation is enhanced due to it being closely tied to empirical data. Future work should explore the ramifications of resolving utility profiles into probability profiles.

An empirically grounded multi-agent simulation also contributes to better understanding network dynamics. This paper serves to unify competing approaches to modelling network evolution. Future work may explore opportunities to introduce random change into the simulated networks. Realistic simulation of networks allows investigators to explore network dynamics by introducing various forms of evolutionary and random change at known points in time and observing their behaviour. This is necessary for exploring networks over time.

The approach presented in this paper is still limited in several ways. The list of objective function effects outlined in Section 3 is not exhaustive. There are likely other important utility seeking functions governing agent interaction. Some effects are highly correlated and including too many effects may lead to over specified or degenerate models. Future work may investigate additional objective functions for actor-oriented models.

Hopefully, multi-agent system researchers will be motivated to apply an actor-oriented approach to empirical network data. The determination of statistically significant utility seeking behaviour in networks offers us a deep, complexity-preserving insight into the underlying behaviour of social systems. Whether the information is used at face value to draw inference on sociological, physical, and biological phenomena, or utilized as an intermediary to simulation analysis, empirical analysis of the utility seeking behaviour characterizing complex networks around us promises to deepen our understanding of them.

## ACKNOWLEDGEMENTS

This is a project of the U.S. Military Academy Network Science Center and the Center for Computational Analysis of Social and Organizational Systems (CASOS) at Carnegie Mellon University. This work was supported in part by the Army Research Institute for the Behavioral and Social Sciences, Army Project No. 611102B74F, the Army Research Labs Grant No. DAAD19-01-2-0009 , the Office of Naval Research (ONR), United States Navy Grant No.N00014-02-10973 on Dynamic Network Analysis, and the Air Force Office of Sponsored Research (MURI: Cultural Modeling of the Adversary Organization, 600322).

## REFERENCES


[1] T. Valente. Network Models and Methods for Studying the Diffusion of Innovations. In: *Models and Methods in Social Network Analysis.* P. Carrington, J. Scott, S. Wasserman (Eds.) Cambridge Press (2007).

[2] T.A.B. Snijders Models for longitudinal network data. In: *Models and Methods in Social Network Analysis.* P. Carrington, J. Scott, S. Wasserman (Eds.) Cambridge Press (2007).

[3] P. Doreian and F. Stokman. *Evolution of Social Networks.* Gordon and Breach, Amsterdam, (1997).

[4] I. McCulloh and K.M.Carley. Dynamic Network Change Detection. In: *Proceedings of the 26th Army Science Conference.* U.S.Army, Orlando, FL, (2008)

[5] I. McCulloh and K.M.Carley. *Detecting Change in Longitudinal Social Networks.* Unpublished manuscript, (2008).

[6] S.Sampson. *Crisis in a cloister*. Ph.D. Thesis. Cornell University, Ithaca, (1969).

[7] T. Newcomb. *The Acquaintance Process*. Holt, Rinehart and Winston, New York, (1961).

[8] A. Romney, A.K. Quantitative models, science and cumulative knowledge. *Quantitative Anthropology*, 1: 153-223 (1989).

[9] A. Sanil, D. Banks, and K.M. Carley. Models for evolving fixed node networks: Model fitting and model testing. *Social Networks*, 17. (1995).

[10] T.A.B. Snijders. Testing for change in a digraph at two time points. *Social Networks*, 12: 539-573 (1990)

[11] O. Frank. Statistical analysis of change in networks, *Statistica Neerlandica* 45: 283–293. (1991)

[12] M. Huisman, and T.A.B. Snijders. Statistical analysis of longitudinal network data with changing composition. *Sociological Methods and Research*, 32:253-287 (2003).

[13] J. Johnson, J. Boster, and L. Palinkas. Social roles and the evolution of networks in extreme and isolated environments. *Mathematical Sociology*, 27: 89-121 (2003).

[14] I. McCulloh, G. Garcia, K. Tardieu, J. MacGibon, H. Dye, K. Moores, J. Graham, and D. Horn. *IkeNet: Social network analysis of e-mail traffic in the Eisenhower Leadership Development Program. (Technical Report, No. 1218).* U.S. Army Research Institute for the Behavioral and Social Sciences. Arlington, VA (2007a).

[15] I. McCulloh, J. Lospinoso, and K.M. Carley. Social Network Probability Mechanics. In: *Proceedings of the World Scientific Engineering Academy and Society 12th International Conference on Applied Mathematics*, WSEAS, Cairo, Egypt. (2007b).

[16] P. Holland, and S. Leinhardt. A dynamic model for social networks. *Mathematical Sociology*, 5:5-20 (1977).

[17] S. Wasserman. *Stochastic Models for Directed Graphs*. Ph.D. dissertation, Harvard University, Department of Statistics, Cambridge, MA (1977).

[18] S. Wasserman. A stochastic model for directed graphs with transition rates determined by reciprocity. In: *Sociological Methodology*. K.F. Schuessler (Ed.) Jossey-Bass, San Francisco CA, 392-412 (1979).

[19] S. Wasserman. Analyzing social networks as stochastic processes. *American Statistical Association*, 75: 280-294 (1980).

[20] R. Leenders. Models for network dynamics: a Markovian framework. *Mathematical Sociology*, 20: 1-21 (1995).

[21] T. A. B. Snijders, and M.A.J. van Duijn. Simulation for Statistical Inference in Dynamic Network Models. In: *Simulating Social Phenomena.* R. Conte, R. Hegselmann, and P. Tera (Eds.) Springer, Berlin, Germany, 493-512 (1997).

[22] T.A.B. Snijders., The statistical evaluation of social network dynamics. In: *Sociological Methodology.* M.E. Sobel, and M.P. Becker (Eds.) Basil Blackwell, Boston, MA, 361-395 (2001).

[23] P. Doreian. On the evolution of group and network structures. II. Structures within structure. *Social Networks*, 8: 33-64 (1983).

[24] K.M. Carley. A theory of group stability. *American Sociology Review*, 56(3): 331–354 (1991).

[25] K.M. Carley. Group Stability: A Socio-Cognitive Approach. *Advances in Group Processes*, 7: 1-44 (1990).

[26] K.M. Carley. Communication Technologies and Their Effect on Cultural Homogeneity, Consensus, and the Diffusion of New Ideas. *Sociological Perspectives*, 38(4): 547-571 (1995).

[27] K.M. Carley. On the evolution of social and organizational networks. *Research in the Sociology of Organizations,* 16: 3-30 (1999).

[28] Snijders, T.A.B. (1996). Stochastic actor-oriented models for network change. Journal of Mathematical Sociology, 21, 149-172.

[29] G. Van de Bunt, M.A.J. van Duijin, and T.A.B. Snijders. Friendship networks through time: An actor-oriented statistical network model. *Computational and Mathematical Organization Theory,* 5: 167-192 (1999).

[30] W. de Nooy. The dynamics of artistic prestige. *Poetics,* 30:147-167 (2002).

[31] M.A.J. van Duijn, E.P.H. Zeggelink, J.M. Huisman, F.N. Stokman, and F.W. Wasseur. Evolution of sociology freshmen into a friendship network. *Mathematical Sociology*, 27: 153-191 (2003).

[32] T.A.B. Snijders, C.E.G. Steglich, M. Schweinberger, and M. Huisman. *Manual for SIENA version 3.1.* University of Groningen: ICS / Department of Sociology; University of Oxford: Department of Statistics (2007).

[33] C.E.G. Steglich, T.A.B. Snijders, and P. West. Applying SIENA: An Illustrative Analysis of the Coevolution of Adolescents' Friendship Networks, Taste in Music, and Alcohol Consumption. *Methodology*, 2: 48-56 (2006).

[34] Alberts, D. *Power to the Edge*. Washington, DC: CISSP, (2002).

[35] J.A. Lospinoso. Utility Maximizing Networks. In: *Proceedings of the 2008 International Conference on Information and Knowledge Engineering*, Las Vegas, NV (2008)

[36] J.A. Lospinoso. The ELICIT Experiment: Eliciting Organizational Effectiveness and Efficiency under Shared Belief. In: *Proceedings of the 12th International Command and Control Research and Technology Symposium*, CCRP, Washington D.C. (2007)

# Multi-Agent Systems and Virtual Producers in Electronic Marketplaces

**Isabel Praça[1], Maria João Viamonte[1], Hugo Morais[1], Zita Vale[1] and Carlos Ramos[1]**

**Abstract.** This paper presents an agent-based simulator designed for analyzing agent market strategies based on a complete understanding of buyer and seller behaviours, preference models and pricing algorithms, considering user risk preferences. The system includes agents that are capable of improving their performance with their own experience, by adapting to the market conditions. In the simulated market agents interact in several different ways and may joint together to form coalitions. In this paper we address multi-agent coalitions to analyse Distributed Generation in Electricity Markets.

## 1 INTRODUCTION

Market players and regulators are very interested in foreseeing market behaviour: regulators to test rules before they are implemented and detect market inefficiencies that should not encourage strategic behaviours that might reduce market performance; market players to understand market behaviour and operate in order to maximize profits. Simulation and Artificial Intelligence techniques may be very helpful under this context.

Multi-agent based simulation is particularly well fitted to analyze dynamic and adaptive systems with complex interactions among constituents [1]. Unlike traditional tools, agent based simulation does not postulate a single decision maker with a single objective for the entire system. Rather, agents, representing the different independent entities in electronic markets, are allowed to establish their own objectives and decision rules. Moreover, as the simulation progresses, agents can adapt their strategies, based on the success or failure of previous efforts.

We present a multi-agent market simulator designed for analyzing agent market strategies based on a complete understanding of buyer and seller behaviours, preference models and pricing algorithms, considering user risk preferences. Each market participant has its own business objectives, and decision model. The results of the negotiations between agents are analyzed by data mining algorithms in order to extract knowledge that gives agents feedback to improve their strategies. The extracted knowledge will be used to set up probable scenarios, analyzed by means of simulation and game theory decision criteria.

We intend to apply this platform to different market types, taking into account some previous work of our research group, where two different simulation platforms have already been developed, namely ISEM – Intelligent System for Electronic MarketPlaces [2], and MASCEM – Multi-Agent Simulator for Competitive Electricity Markets[3].

ISEM focuses specially on markets with finite time horizon. This simulator was selected as a worldwide case study in simulation of negotiation agents [4], while MASCEM focus particularly on market mechanisms usually found in liberalized electricity markets and was selected as a worldwide case study of agents technology applied to markets [5].

Our proposal is a Market Simulator that will act as a kind of What-if tool, trying to analyze what may occur if some decision is taken. However, some additional intelligence need to be placed in the system, otherwise we will have a kind of combinatorial explosion, since many scenarios need to be analyzed. Moreover, the Market Simulator will be used as the engine of a Market Participant (Seller or Client) in order to suggest him/her about the actions to have in the market.

Entities from real markets can use our tool to test several different negotiation mechanisms, different behaviours, strategies and risk preferences, and to analyze the future market evolution and other entities expected reactions. Our tool may also be used to understand the implications of agents' coalitions on markets.

Electricity Markets are an important area of application of our research. In this paper we present our developments in studying the increase in Distributed Generation, and its market influence, by means of agents' coalitions.

In a general way the formation of coalition can be seen as an enterprise example of the constitution of a social net, where the several elements of this net establish negotiation processes in order that a formal structure, with the capacity to supply goods and services to the society, can emerge. Each element of the coalition is not able to supply the services and goods with the desirable amount and quality, but the coalition presents itself as a credible structure to the eyes of the potentials consumers, that is, the structure (coalition) thus created worth more than the sum of its parts.

This article illustrates the constitution of coalitions for an important problem, subject to enormous transformations and with clear social and strategically impact: the establishment of electric energy supply contracts through competitive markets.

The rest of the paper is organized as follows: section 2 outlines the multi-agent model; section 3 addresses the negotiation mechanisms; section 4 explains the use of data mining in the scope of our tool; section 5 addresses agents strategic behaviour; and section 6 explores agents coalition to represent virtual power producers in the study of distributed generation penetration on electricity markets.

[1] GECAD – Knowledge Engineering and Decision-Support Research Group of the Engineering Institute of Porto – Polytechnic Institute of Porto (ISEP/IPP), Rua Dr. António Bernardino de Almeida, 431, 4200-072 Porto, Portugal. Email: {icp, mjv, hgvm, csr, zav}@isep.ipp.pt.

## 2 MULTI-AGENT MODEL

Our Simulator facilitates agent meeting and matching, besides supporting the negotiation model. In order to have results and feedback to improve the negotiation models and consequently the behaviour of user agents, we simulate a series of negotiation periods, $D = \{1, 2, ..., n\}$, where each one is composed by a fixed interval of time $T = \{0, 1, ..., m\}$. Furthermore, each agent has a deadline $D_{max}^{Agt} \in D$ to achieve its business objectives. At a particular negotiation period, each agent has an objective that specifies its intention to buy or sell a particular good or service and on what conditions.

The available agents can establish their own objectives and decision rules. Moreover, they can adapt their strategies as the simulation progresses on the basis of previous effort's successes or failures. The simulator probes the conditions and the effects of market rules, by simulating the participant's strategic behaviour.

The simulator was developed based on "A Model for Developing a MarketPlace with Software Agents (MoDeMA)" [4]. The following steps compose MoDeMA::

- Marketplace model definition, that permits doing transactions according to the Consumer Buying Behaviour Model;
- Identification of the different participants, and the possible interactions between them;
- Ontology specification, that identifies and represents items on transaction;
- Agents architecture specification, and information flows between each agents module;
- Knowledge Acquisition, defining the process that guarantees the agent the knowledge to act on pursuit of its role;
- Negotiation Model, defining the negotiation mechanisms to be used;
- Negotiation Protocol, specification of each negotiation mechanism rules;
- Negotiation Strategies, specification and development of several negotiation strategies;
- Knowledge Discovery, identification and gathering of market knowledge to support agents' strategic behaviour.

Multi-agent model includes a market administrator, buyers, sellers, traders and a market operator.

The market administrator agent has two main functions: coordinator and knowledge provider. On one hand it coordinates the simulated market and ensures that it functions correctly, according to market mechanisms and established rules. On the other hand, it plays the role of "power" agent, since it has access to market knowledge, which contains information about the organisational and operational rules of the market, as well as information about all different running agents, their capabilities and historical information. The market previsions and agent behaviour models are obtained through data mining algorithms, using data resulting from agent negotiations that support agents' market strategies.

Since we intend to cover several negotiation mechanisms, our model also includes a market operator agent, responsible to support negotiations based on an auction mechanism.

Seller and buyer agents are the two key players in the market, so we devote special attention to them, particularly to their business objectives and strategies to reach them. In order to be competitive in today's economic markets, buyer and seller agents need not only to be efficient in their business field, but also to be able to quickly react and adapt to new environments as well as to interact with other available entities. The control architecture adopted for the design of those agents meet these requirements, having a similar structure but with a kind of symmetrical behaviour (due to their antagonistic business objectives).

## 3 NEGOTIATION MECHANISMS

As a decision support tool, our simulator includes several types of negotiation mechanisms to let the user test them and learn the best way to negotiate in each one. So, we include bilateral contracts and a Pool, centralized mechanism based on an auction, and regulated by a market operator. Both types of negotiation may exist at the same time: Mixed Market. These implies each agent must decide whether to, and how to, participate in each market type.

Let $Agtb$ denote the buyer agent, $Agts$ the seller agent and let $[Pi_{\min}, Pi_{\max}]$ denote the range of values for price that are acceptable for agents.

A seller agent has the range $[Psi_{\min}, Psi_{\max}]$, which denotes the scale of values that are comprised of the minimum value that the seller is disposed to sell to the optimal value.

A buyer agent has the range $[Pbi_{\min}, Pbi_{\max}]$, which denotes the scale of values that are comprised of the optimal value to buy to the maximum value.

### 3.1 Bilateral Contracts

In bilateral contracting buyer agents are looking for sellers that can provide them the desired products at the best price. We adopt what is basically an alternating protocol [6].

Negotiation starts when a buyer agent sends a request for proposal. In response, a seller agent analyses its own capabilities, current availability, and past experiences and formulates a proposal.

Sellers can formulate two kinds of proposals: a proposal for the product requested; or a proposal for a related product, according to the buyer preference model.

$PPg_{i\,Agts \rightarrow Agtb}^{DT}$ represents the proposal offered by the seller agent $Agts$ to the buyer agent $Agtb$ at time T, at the negotiation period D for a specific product.

The buyer agent evaluates the proposals received with an algorithm that calculates the utility for each one, $U_{PPgi}^{Agtb}$; if the value of $U_{PPgi}^{Agtb}$ for $PPg_{i\,Agts \rightarrow Agtb}^{DT}$ at time T is greater than the value of the counter-proposal that the buyer agent will formulate for the next time T, in the same negotiation period D, then the buyer agent accepts the offer and negotiation ends successfully in an agreement; otherwise a counter-proposal $CPg_{i\,Agtb \rightarrow Agts}^{DT}$ is made by the buyer agent to the next time T.

The seller agent will accept a buyer counter-proposal if the value of $U_{CPgi}^{Agts}$ is greater than the value of the counter-proposal that the seller agent will formulate for the next timeT; otherwise the seller agent rejects the counter-proposal.

On the basis of the bilateral agreements made among market players and lessons learned from previous bid rounds, both agents revise their strategies for the next negotiation rounds and update their individual knowledge module.

### 3.2 Pool

In our simulator, agents also have the possibility of negotiating through a Pool, which is a centralized mechanism that functions according to an auction mechanism, and is regulated by a market operator. We have two different auction mechanisms: a double and a single uniform auction.

The process starts at the market operator, who sends a request for participation. The *call_for_participation* message triggers the negotiation process and is delivered to all agents in the simulated market. If the agent is interested, or capable, of participating in the Pool, it will formulate a bid and send it to the market operator, specifying for each requested parameter the value of its proposal.

The process of formulating bids, by buyer and seller agents, is related to agent strategies, addressed in detail in section 6. The market operator evaluates all the received bids, analyses them through the pool auction mechanism, defines the market price and accepted bids. Then a *reply_bid* message is sent to all pool participants, specifying the settled market price and if the bid was or not accepted and why.

### 3.3 Mixed Markets

The Mixed model combines features of Pools and Bilateral Contracts. In this model, a Pool isn't mandatory, and customers can either negotiate an agreement directly with sellers, at the pool market price or both. Agents must decide whether to try or not the Pool, whether to keep bilateral negotiations simultaneously with Pool negotiations or just after Pool results if bids were not accepted. For that agents use their past experiences, market knowledge and agents own negotiation strategies to support their decisions.

## 4 DATA MINING SUPPORT

The market previsions and agent behaviour models are obtained through data mining algorithms, using data resulting from agent negotiations that support agents' market strategies. In practice, usually, after a confidential negotiation period, the market administrator agent discloses information about past transactions and agents' characteristics (if possible); all agent interactions are logged at a transaction level of detail, which provide a rich source of business insight that can help to customise the business offerings to the needs of the individual buyers. With this functionality it is possible to discover sub-groups that behave independently and associations between products. For that, our market simulator uses clustering, classification and association operations.

To carry out the clustering operation a Two-Step clustering algorithm [7] is used to target buyers with similar characteristics in the same agent group. Then, to obtain more relevant information that describes the consumption patterns of each cluster population, a rule-based modelling technique, using C5.0 classification algorithm, an evolution of C4.5 algorithm [8], is used to analyse those clusters and to obtain descriptions based on a set of attributes, collected in the individual agents' knowledge module. These models are transferred to the market administrator agent and offer a set of market information, such as: preferred sellers; preferred marks; favourite products and reference prices, which support the process of agents' strategy implementation.

To discover associations between buyer details and purchases, data from multiple agent negotiations are manipulated to create "basket" records showing product purchases. This permits the observation of the behaviour of each buyer agent. This data is combined and manipulated by the "Apriori algorithm" [9], to discover associations between buyer details and purchases. The best association rules, those with a strong support and confidence, are extracted and transferred to the market administrator agent. With this kind of knowledge it is possible to provide insight into the sellers' agents about the profiles of buyer agents with certain purchase propensities, showing associations between products, prices, style, etc.

After these operations, to get confident data, agents can request the services provided by the market administrator agent, in order to support their strategic behaviour. Only players with more sophisticated behaviour will take advantage of this new knowledge; since the user can determine which seller agents have access to this facility. The user can also determine if the agents' information will be private or public; public information is available to market analysis with the data mining functionality. However the market can get knowledge about an agents' behaviour even if they are set as a private information agent. This situation occurs, by the simple fact of being on the market.

## 5 STRATEGIC BEHAVIORS

### 5.1 Bilateral Contracts

Agents use four time-dependent strategies to change their price during a negotiation period: Determined, Anxious, Moderate and Gluttonous, these strategies depending on both the point in time when the agent starts to modify the price and the amount it changes.

Although time-dependent strategies are simple to understand and implement [10], they are very important since they allow the simulation of important issues such as: emotional aspects and different risk behaviours. For example, an agent that gains utility, with the time, and has the incentive to reach a late agreement (within the remaining time until the end of a negotiation period) is considered a strong or patient player; an agent that loses utility with time and that tries to reach an early agreement is considered a weak or impatient player.

### 5.2 Behavior-dependent Strategies

In this work, we have also used the time-dependent strategies, based on the model proposed by S. Fatima [11], to model different attitudes towards time, during a negotiation period.

Agents use behaviour-dependent strategies to adjust parameters for the next negotiation period according to the results obtained in the previous ones. Buyers and seller agents develop their behaviour and strategies based on a combination of public information, available through requesting from market administrator services; and private information, available only to the specific agent at their individual knowledge module.

For Pool Negotiations we define two different behaviour-dependent strategies: one called Composed Goal Directed (CGD) and another called Adapted Derivative Following (ADF). The CGD strategy is based on two consecutive objectives, the first one is selling (or buying) all the available (or needed) units, and then increase the profit (reduce the payoff). The ADF strategy is based on the Derivative Following strategy proposed by Greenwald [12]. The ADF strategy adjusts its price by looking at the amount of revenue earned in the previous period as a result of the previous period's price change. If the last period's price change produced more revenue per good than the previous period, then the strategy makes a similar change in price. If the previous change produced less revenue per good, then the strategy makes an opposite price change.

For Bilateral Contracts Negotiations we also have several behaviour-dependent strategies. Buyer agents can use two complementary behaviour-dependent strategies: the Modified Goal Directed for Buyers (MGDB) and the Fragmented Demand (FD). The MGDB strategy is an adaptation of CGD for bilateral contracts. The FD strategy, adjusts the demand per day by attempting to reach the goal of buying its entire needs by the last day of the market, and not before, this strategy paces its purchases over the market, with the goal of buying all the units needed but with less costs. Seller agents can also choose from two different behaviour-dependent strategies: the Modified Goal Directed for Sellers (MGDS), that adjusts its price by attempting to reach the goal of selling the entire inventory by the last day of the market, by lowering prices when sales in the previous day are low and raising prices when the sales are high; and the Derivative Following (DF) strategy weighted by Seller Satisfaction (DFWS) or by the Previewed Demand for a specific product (DFWPD). The DFWS/PD is based on the ADF behaviour weighted by the referred issues. Seller agents can obtain these values through requesting for market administrator agent support.

## 6 AGENTS COALITIONS MODELLING VIRTUAL POWER PRODUCERS

Coalition formation is the coming together of a number of distinct, autonomous agents that agree to coordinate and cooperate, acting as a coherent grouping, in the performance of a specific task. Such coalitions can improve the performance of the individual agents and/or the system as a whole. It is an important form of interaction in multi-agent systems.

It has been advocated in e-commerce (where buyers may pool their requirements in order to obtain bigger group discounts), in grid computing (where multi-institution virtual organizations are viewed as being central to coordinated resource sharing and problem solving), and in e-business (where agile groupings of agents need to be formed in order to satisfy particular market niches). In all of these cases, the formation of coalitions aims to increase the agents' abilities to satisfy goals and to maximize their individual or the system's outcomes.

Most work on coalition formation in multi-agent systems and game theory has focus on payoff distribution, where it is usually assumed that a coalition structure has been formed, and the question is then how to divide the payoff so that the coalition structure is stable. In this context, many solutions have been proposed based on different stability concepts. Transfer schemes have also been developed to transfer non-stable payoff distributions to stable ones (while keeping the coalition structure unchanged).

Research is giving attention to the coalition structure generation [13], [14]. The work of Shehory and Kraus [15] considers a somewhat broader environment, where the coalitions can be overlapped but the complexity is reduced by limiting the size of the coalitions.

Some other researchers address both coalition structure generation and payoff distribution in competitive environments. Ketchpel [16] presents a coalition formation method with cubic running time in the number of agents, but his method can neither guarantee a bound from the optimal nor stability. Shehory and Kraus's protocol guarantees that if the agents follow it, certain stability (kernel-stability) is met. In the same paper, they also present an alternative protocol that offers a weaker form of stability with polynomial running time. However, in both cases, no bound from the optimal is guaranteed.

More recent research in coalition formation area has also begun to pay attention to dynamic environments, where agents may enter or leave the coalition formation process and many uncertainties are present (e.g. the coalition value is not fixed, but it is context-based [17]).

### 6.1 Virtual Power Producers

The aggregation of distributed generation plants gives place to the new concept of Virtual Power Producers (VPP). VPPs are multi-technology and multi-site heterogeneous entities, being relationships among aggregated producers and among VPPs and the remaining Electricity Market (EM) agents a key factor for their success. An aggregating strategy can enable owners of Distributed Generation to gain technical and commercial advantages, making profit of the specific advantages of a mix of several generation technologies and overcoming serious disadvantages of some technologies.

Any type of generation unit or load may be included: wind turbines, photovoltaic, mini turbines, micro-turbine, fuel cells, energy storage units, non-controllable loads, controllable loads etc. The typical size of single distributed energy resource units may range from a few kW to some MW.

In the scope of a VPP, aggregated producers (AP) can make sure their generators are optimally operated and that the power that is not consumed in their installation has good chances to be sold on the market. At the same time, VPPs will be able to commit to a more robust generation profile, raising the value of non-dispatchable generation technologies.

Under this context, VPPs can ensure secure, environmentally friendly generation and optimal management of heat, electricity and cold and optimal operation and maintenance of electrical equipment, including the sale of electricity to the EM. VPPs should adopt organization and management methodologies so that they can make distributed generation a really profitable activity.

VPPs must be flexible enough to use the advantages of its resources (e.g. market-based environmental value in the form of pollution and/or carbon credits, renewable energy credits) and overcoming their problems and limitations.

VPPs must identify the characteristics of each of the AP and try to optimize the selling activity so that each one delivers the biggest possible amount of energy. However, this is not simple due to uncertainty of generation associated with the technologies that depend from natural resources such as wind, sun, waves or water flows.

So, in order to have VPP able to coexist with other market agents, it is necessary that it gets profits and that has credibility in the EM. This context must be considered in VPPs organization and operation methodologies as their goal is to optimize their APs' profits in this market.

A successful achievement of VPPs' goals requires the use of a mix of adequate technologies for optimizing and supporting their activities. Under this scope agents and multi-agent systems are important technologies to adequately simulate EM behaviour and gather knowledge to provide decision-support to strategic behaviour. Taking into account the already described MASCEM characteristics, it can be a valuable framework to test VPP functioning under different market mechanisms and concerning different market strategic behaviour.

## 6. 2 VPP Coalition formation

From the point of view of the multi-agent system, VPP are seen as coalitions of agents, requiring specific procedures for coalition formation. Once a coalition is established, it can aggregate more agents or even discard some agents. This allows modelling all the decision making concerning VPP formation and also subsequent aggregation of more producers.

VPP needs to have an adequate knowledge of each potential aggregated producer characteristics. Some of the most important characteristics are:

- Nominal Power: the sum of nominal power installed in each producer;
- Available Power: the power a VPP can buy to the producer;
- Overload Power: some units may produce overload power for limited periods. The VPP may use this power in critical situations;
- Equipment characteristics: information concerning producers' equipment allows the VPP to know the power characteristic, reliability, maintenance periods, lifetime, relation with external factors, possible variations of the energy price in function of the cost of the primary resources, etc.
- Operating limits: for the units which are dependent from natural resources, it is possible that the primary resource must be below or above of equipment operating limits. This must be considered in risk analysis in the generation forecast. Usually when the resources forecast is near to the minimum machines operating limit the risk is small, but when they are near to the maximum limit the risk can be enormous;
- Grid connection characteristics: This is an important aspect if it is necessary to pay the losses in the lines; also the existence of two or more producers connected to the same electric substation should be considered; etc;
- Historical generation data: the availability of historic generation data can enable the VPP to get useful forecasting tools.

On one hand, each VPP classifies the producers according to several defined criteria. On the other hand, it establishes the goals of VPP formation or of VPP aggregation of more producers, according to its operating strategies and to its necessities at the moment. Aggregation proposals are then elaborated based on the resulting knowledge. Each producer decision of participating or not in VPP coalition is dependent on agents' market strategies and risk preference.

Once the VPP formation process finished, the VPP needs to coordinate its operation. The VPP must place bids in the market, considering the contracts with producers, the generation forecast, the reserves and its market strategy. According to its member generation capabilities and consumption needs, for a given period, the VPP agent will need to sell or buy electricity. VPP agents have the same market interface as Seller or Buyer agents.

However, as VPPs are themselves a set of other agents, there are some preliminary steps to define its bids. Firstly, all the capacity available from the different aggregated distributed energy resources must be gathered to establish the electricity amount to trade on the market. The different generation costs must be analyzed to define the interval for envisaged proposals. This means VPP agents have a utility function that aggregates all the involved units' characteristics. The analysis of the aggregated producers' proposals will be done according to each unit capabilities and costs.

After the market session, the VPP agent undertakes an internal dispatch, analyzing and adjusting its generation and reserve to maximize profits. VPP informs the aggregated producers about their dispatch. Finally, in function of the generation, the used and unused reserve of each producer and the established contracts of the VPP fulfilment, the VPP determines the producers' remuneration. The Introduction of VPP models in the simulator required to rethink the multi-agent architecture, namely in what concerns agent communication [18]. Figure 1 illustrates the simulator negotiation framework allowing for VPPs.

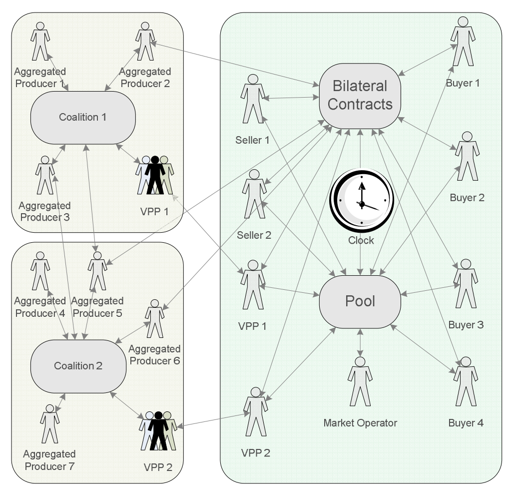


**Figure 1.** Negotiation framework regarding VPPs

Considering each VPP as a multi-agent system allows an interesting approach from both the performance and the conceptual point of view. In order to develop a computational implementation of this conceptual architecture, each VPP has to have its own facilitator. This means that each VPP has now its own facilitator that allows it to communicate with all the producers that are part of this coalition or intend to join it, independently from the rest of the simulation.

### 6. 3 VPP Case Study

In fact, our simulator is already being used to study several scenarios from real electricity markets, namely from real data obtained from OMEL, the Spanish Electricity Market. The data was analyzed by means of statistical and data mining techniques in order to have an illustrative scenario although with a limited number of agents. In this example we have a scenario with 7 buyers, 5 sellers, and 2 VPPs. Concerning VPPs, there is a VPP with 3 aggregated producers, all of which have wind farms; and another with 4 aggregated producers (1 photovoltaic plant, 1 wind farm, 1 co-generation and 1 mini-hydro). Figure 2 illustrates market transactions by agent.

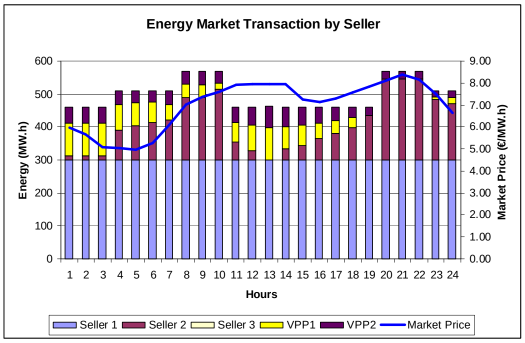


**Figure 2** Energy market transactions by seller

Market results showed a coherent behaviour which gives us confidence on using our simulator to test several VPP configurations under different scenarios in order to draw conclusions about VPP advantages and drawbacks for the involved agents and for the market.

## 7 CONCLUSIONS

In the near future, agent market strategies will be a common competitive manoeuvre for electronic markets. Market participant's strategic behaviour is very significant in the context of competition. In addition, the availability of new market knowledge obtained with data mining algorithms is vital for supporting marketing and sales. Also important is the development of agent-based tools that will help in understanding what kinds of electronic market strategies are appropriate. Agent coalitions is an important issue that can also be analysed by means of our tool. This is already being important to study and obtain conclusions about distributed generation penetration into electricity markets.

## ACKNOWLEDGMENTS

The authors would like to acknowledge FCT, FEDER, POCTI, POSI, POCI and POSC for their support to R&D Projects and GECAD Unit.

## REFERENCES

[1] A. Helleboogh, G. Vizzari, A. Uhrmacher, F. Michel. Modeling dynamic environments in multi-agent simulation. *JAAMAS*, vol. 14, nr. 1, pp. 87–116 (2007).

[2] M. J. Viamonte, C. Ramos, F. Rodrigues, and J. C. Cardoso. Simulating the Behaviour of Electronic MarketPlaces with an Agent-based Approach. *Procs of the International Conference on Web Intelligence – IEEE/WIC 2004*, Beijing, China, pp. 553-557 (2004),.

[3] I. Praça, C. Ramos, Z. Vale, M. Cordeiro. Intelligent Agents for Negotiation and Game-based Decision Support in Electricity Markets. *International Journal of Engineering Intelligent Systems, Special Issue "Intelligent Systems Application to Power Systems"*, vol. 13, nr. 2, pp. 147-154 (2005).

[4] M. J. Viamonte, C. Ramos, F. Rodrigues, and J. C. Cardoso. ISEM: A Multi-Agent Simulator For Testing Agent Market Strategies. *IEEE Transactions on Systems, Man and Cybernetics – Part C: Special Issue on Game-theoretic Analysis and Stochastic Simulation of Negotiation Agents*, vol. 36, nr. 1, pp. 107-113 (2006).

[5] I. Praça, C. Ramos, Z. Vale, M. Cordeiro. MASCEM: A Multiagent System that Simulates Competitive Electricity Markets. *IEEE Intelligent Systems*, vol. 18, nr. 6, Nov/Dec 2003, pp. 245-276 (2003).

[6] P. Faratin, C. Sierra, and N. Jennings. Negotiation Decision Functions for Autonomous Agents. *Int. J. Robotics and Autonomous System*, vol. 24 , nr. 3, pp. 159-182 (1998).

[7] T. Zhang, R. Ramakrishnon, and M. Livny. BIRCH: An Efficient Data Clustering Method for Very Large Databases. *Procs of the ACM SIGMOD Conference on Management of Data*, Montreal, Canada, pp. 103-114 (1996).

[8] J. Quinlan. *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, San Francisco, USA (1993).

[9] R. Agrawal, H. Manilla, R. Srikantand, H. Toivonen, and I. Verkamo. *Fast discovery of association rules*, AAAI Press (1996).

[10] J. Morris, A. Greenwald, and P. Maes. Learning Curve: A Simulation-based Approach to Dynamic Pricing. Electronic Commerce Research: Special Issue on Aspects of Internet Agent-based E-Business Systems. *Kluwer Academic Publishers*, vol. 3, nr. 3-4, pp. 245-276 (2003).

[11] S. Fatima, M. Wooldridge, and N. Jennings. An agenda-based framework for multi-issue negotiation. *Artificial In-telligence Journal* , vol. 152, nr. 1, pp. 1-45 (2004).

[12]A. Greenwald, J. Kephart. Shopbots and Pricebots. *Procs of the Sixteenth International Joint Conference on Artificial Intelligence - IJCAI*, Stockholm (1999).

[13] T. Sandholm, K. Larson, M. Andersson, O. Shehory and F. Tohme. Coalition structure generation with worst case guarantees. *Artificial Intelligence*, vol. 111, nr. 1-2,pp. 209– 238 (1999).

[14] V. D. Dang, N. Jennings. Generating coalition structures with finite bound from the optimal guarantees. *Procs 3rd International Conference on Autonomous Agents and Multi-Agent Systems*, Nova York, E. U. A., pp.564-571 (2004).

[15] O. Shehory, S. Kraus. A kernel-oriented model for coalition-formation in general environments: Implementation and results. *Procs of the Thirteenth National Conference on Artificial Intelligence*, pp. 134–140 (1996).

[16] S. P. Ketchpel. Forming coalitions in the face of uncertain rewards. *Procs of the Twelfth National Conference on Artificial Intelligence*, pp. 414–419 (1994).

[17] M. Klusch and A. Gerber. Dynamic coalition formation among rational agents. *IEEE Intelligent Systems*, vol. 17, nr. 3, pp. 42–47(2002).

[18]I. Praça, H. Morais, C. Ramos, Z. Vale, and H. Khodr. Multi-Agent Electricity Market Simulation with Dynamic Strategies & Virtual Power Producers. *Procs. Of the IEEE Power & Energy Society - 2008 PES General Meeting*, Pittsburgh, Pennsylvania, E.U.A. (2008).