

# Delegations Guided by Trust and Autonomy

Henry Hexmoor<sup>a</sup>, Shahram Rahimi<sup>a</sup>, Rachil Chandran<sup>b</sup>

<sup>a</sup>*Department of Computer Science, Southern Illinois University, Carbondale, IL. 62901*

<sup>b</sup>*Department of Computer Science and Computer Engineering, University of Arkansas, Fayetteville, AR 72701, USA  
Tel.:479.575.2420, fax: 479.575.5339; E-Mail:{hexmoor,chandra}@uark.edu*

**Abstract.** This paper explores delegation decisions predicated on models of trust and autonomy among agents. In socially rich environments, trust and autonomy of artificial agents are key attributes for rational delegation decisions. Social agents are affected by many social attributes such as benevolence, social exchanges, power, and norms. We present cognitively inspired working models of trust and autonomy for delegation among agents that accounts for these social attributes. These models are validated in a simulation of a group of agents who delegate tasks in order to locate the most socially well-suited individuals for performing that task. We define autonomy in such an objective manner that agents who have the highest autonomy with respect to a task will generally be the most socially well suited. Benevolence, social reciprocation, and norms most directly contribute to autonomy and proportionally to social suitability. Derived social power directly contributes to social suitability, whereas independently assigned social power ranking overrides the suitability. We show that trust and autonomy in an agent group changes under various social settings.

**Keywords:** Social norms, delegation theory, autonomy, trust, social influence, artificial societies, social simulation

## 1. Introduction

Social concepts such as benevolence, social exchanges, power, and norms are constructively used in modeling multi-agent system dynamics. Interpersonal trust as well as the issue of self-determination largely governs delegation among agents [58, 59]. The latter forms the basis of autonomy [15, 26, 51]. Some systems propose the use of auctioning or the use of a separate agent, i.e., a “broker,” to avoid the reasoning costs associated with delegation [21, 31, 55]. However, we focus on delegation and offer models of trust and autonomy that account for impacts of benevolence, social exchanges, power and norms [16, 29, 30]. Whereas the sense of being trusted contributes to an agent’s autonomy, the sense of trusting another agent contributes to an agent’s motivation to delegate [48]. Agent trust and autonomy are quantified using differing levels of benevolence, reciprocal social exchanges, power and norms. We believe that an understanding of the relationship between trust, autonomy, and related notions such as social power enables us to design more predictable artificial societies endowed with social cognition.

Broadly speaking, studies of autonomy can be categorized in three areas. The first is the need to adjust the independence level of agents in their interaction with humans. This is often called adjustable autonomy [7].

The second category is when an electronic virtual-society is under design with a heterogeneous set of agents and there is a need to design protocols and principles of interaction that produce and maintain harmonious dynamics. Agents in such environments may need to dynamically reason about teams or coalitions that require rich reasoning about autonomy [47]. The third category of exploring autonomy is in understanding group dynamics that include rich social interactions [28]. The focus on autonomy in this category is to understand that it is a supporting notion to primary social issues such as delegation. This article presents a study of autonomy from this perspective. More specifically, we consider personal autonomy for determining the nature of interaction with other agents [32]. An agent’s perceived degree of autonomy is used in deciding whether to delegate a task to others or to perform it by itself.

Endogenous variables are those whose values are determined inside the model from the values of exogenous variables [19]. In this paper we will limit endogenous autonomy sources to be the agent’s capability of performing a target task. In contrast, exogenous sources might be powers the agent perceives from other agents or permissions (i.e., deontologies) it experiences from others. We model the permissions in terms of a level of trust an agent perceives from peer agents.

An agent’s autonomy toward a task is affected by its capability and the sense of freedom it receives from other agents [28, 33]. This sense of freedom can be approximated by a combination of factors such as social ties, both the number of ties and strength of ties, and trust it perceives from others. The network of social exchange might have asymmetries. In addition to direct exchange, agents might experience indirect exchange via other

agents. An agent's trust in others generates benevolence and reciprocation of trust. Therefore, an agent who trusts others is likely to experience autonomy.

Agent trust is important in agent-mediated application services [61]. There are many ways to consider trust, including objective and subjective trust [36]. There have been many studies of trust, which have introduced many definitions of trust [9, 23, 24, 46, 49]. For instance, institutional trust is the trust that exists among individuals due to their participation in social norms and values of various institutions of which they are members, whereas system trust is an individual's trust in the reliability of social structures in its environment [39, 43, 45]. Trust between agents depends on many parameters including competency, histories, reciprocities, benevolence, culture, and reputation [12]. A survey of the various trust and reputation models prevalent today is found in [50].

We regard trust to be a prerequisite for delegation. We suggest that agent X's trust in agent Y with respect to task T, denoted by Trust (X, Y, T), is partly a function of agent X's perception of agent Y's benevolence towards it, and partly a function of agent X's perception of agent Y's capability toward task T, partly due to the balance of tit-for-tat [30], partly due to the power variations with others, and partly due to the variations in norms that each agent follows.

There are many other accounts of interpersonal trust [1]. Our conceptualization is closest to that of Castelfranchi's where he bases trust on competence and willingness of the trustee [12]. This approach to conceptualizing trust lends itself to formulating delegation between two individuals, which requires trust between delegator and delegee. Although we agree on the competence component of trust, we feel that Castelfranchi's definition could apply to complete strangers and does not take into account the interpersonal sense of trust that we wish to adopt. The "willingness" component does not capture the trustee's attitude towards the trusted. Interpersonal trust is also a function of familiarity and social ties. Social ties affect trust temporally. Interpersonal trust, our main focus, differs from system and institutional trust. Usually, trust levels accumulate and diminish gradually unless there are radical changes in agents' attitude toward each other, such as major changes in benevolence [1].

Another conceptualization is that trust is not a precursor to delegation but one between collaborating individuals who communicate [57]. Trust is in the degree of belief in the validity of messages. As the preference to the attribute changes in the measure of trust, the notion of trust changes. When benevolence is the major role, truth telling is taken into consideration. The agents' social interaction is taken into consideration when reciprocal social exchanges play a major role, and the agents' interaction with different power levels is considered when power is the major role.

In section 2 we will outline our proposed models for autonomy and trust. In section 3 we will describe an implemented simulation that creates a domain-neutral electronic society with recurring tasks. In section 4, we present salient results of dynamics set forth by our models. These results corroborate our intuitive understanding of trust and autonomy. In section 5 we offer concluding remarks and chart a course for further work.

## 2. Models of Trust and Autonomy

Our model of trust is a precondition to the formation of intentions to delegate a task, i.e., entrusting responsibility of task completion to another agent. An agent's assessment prior to delegation may include an analysis of risk and utilities, thus creating an intermediate notion of trusting value, prior to adoption of an intention. In most applications, such as those detailed in [40, 56, 60] trust has the consequence of reducing the need for the trusting agent to supervise or monitor the trusted agent. In this section we present a few models for trust and autonomy that increase in complexity. Each model introduces the impact of single social attribute and incrementally builds on a previously defined model.

The variety of trust definitions has added to the confusion and misconceptions of trust. In multiagent systems, trust has been related to models of other social notions such as autonomy, delegation, dependence, control, power and norms. These notions influence interactions between agents. We treat trust as a dyadic relation, i.e., the amount of trust each agent has in other agents.

**Trust:** Trusting value between agents A and B, and a target task t, denoted by Trust (A, B, t), is the amount of trust agent A has on agent B with respect to task t to be performed by agent B.

Trust can be computed for multiple agents or for individual agents. This value among the agents is subject to many factors such as benevolence of agents, general reputations, competencies, reciprocity among agents, histories or prior experiences, and environmental factors such as culture and organization.

**Autonomy:** Autonomy of an agent A with respect to a target task t to be performed by agent A, denoted by Autonomy (A, t), is the degree of self-determination an agent possesses to perform task t. Agent autonomy is often interpreted as freedom from human intervention, oversight, or control, but many definitions abound [8, 22, 35, 38]. Autonomy could be further construed as necessary for the flourishing of any agent, and thus any agent system [34]. Additionally, the behavior of autonomous agents is generally viewed as goal-directed. That is, autonomous agents act with the purpose of achieving their goals. Similar to Trust, autonomy of the

agents is subject to many factors including capability, being trusted, reciprocities, and power.

The first attribute we consider is *benevolence*. Differing benevolence levels contribute to different trends of trusting and trusted values, which leads to the decision to self-act versus to delegate. When agents are acquainted with one another, they may develop a balance of good-will and help toward one another. We model this in terms of an attitude of benevolence among agents [41]. Agents might have a general attitude of benevolence towards other agents, but here we treat benevolence as a dyadic relation. If an agent X has high benevolence toward agent Y, when agent Y asks agent X for a task, agent X is willing to accept the request. Furthermore, agents usually reciprocate with benevolence, i.e., benevolence usually begets more benevolence [52, 53]. However, benevolence is not necessarily symmetric. Agent X's benevolence toward agent Y is not necessarily the same as agent Y's benevolence toward agent X.

Benevolence does not subtract from an agent's autonomy. However, lack of benevolence toward other agents leads to a lack of being trusted, and that indirectly leads to lower autonomy. Benevolence might increase autonomy. Being benevolent contributes to being trusted by others, which might lead to being a popular delegate, i.e., being given many tasks. If such an agent is highly competent as well, the agent will accumulate successes. A combination of increased competence and being trusted will contribute to an agent's heightened sense of autonomy.

Benevolence might be considered as a positive social norm to adopt, although, in general, norm abiding does not directly affect trust or autonomy. Norm adoption does affect an agent's reputation, which may have effects on trust. Many fields of science such as biology, and philosophy have actually addressed the concept of benevolence. A number of researchers have defined, characterized and analyzed benevolent agents [4, 18, 53] Benevolence is defined as a concept of goodness, social duty, and utility function in [41]. Our concept of benevolence is summarized below:

An agent is benevolent if and only if:

1. The agent *voluntarily* helps other agents without being commanded to do so.
2. The agent's benevolent actions are intended to benefit the society to which the agent belongs.
3. The agent should not expect an *immediate* reward or benefit for its benevolent actions. If it did, then the agent is instrumental, not benevolent [10].
4. The agent's benevolent action is considered while the agent is pursuing one of its own goals in such a way that an agent should neither prevent nor help the agent accomplish its goal.

As defined earlier in our model, Trust is the degree of belief an agent has with respect to other agents and with

respect to a particular task. This value among agents is computed by equation 1.

$$\text{Trust (A, B, t)} = \frac{1}{2} [\text{capability (B, t)} + \text{benevolence (B, A, t)}] \quad (1)$$

In equation 1, A and B are agents, and t is the task to be performed by agent B. Capability (B, t) quantifies the agent B's known ability to perform task t. In general, capability requires opportunity. Here we are limiting capability to common knowledge of agents about one another's past performance and not their capacity. Benevolence (B, A, t) denotes the trustee agent B's level of well wishing towards truster agent A for performing a task t. We assume all agents objectively share information about the benevolence levels of all other agents. This value may be a result of many factors such as kinship (i.e., thick relationships), social ties (i.e., a variety of familiarity relationships), positive regard, alliance, coalition, team allegiance, predictability, shared values and norms, dependence, and even commitment. Arguably, this value is difficult to perceive. However, since we are building artificial agents we have control over this component and approximate this quality in biological organisms.

The autonomy of an agent is its ability to perform a task by itself; i.e., with no intervention from other entities and is computed by equation 2.

$$\text{Autonomy (A, t)} = \frac{1}{2} [\text{capability (A, t)} + \text{Average (T)}] \quad (2)$$

Average (T) is the average trust of all agents towards agent A and is measured by the following:

$$\frac{1}{n-1} \sum_{i=1}^{n-1} \text{Trust}(i, A, t)$$

Here Trust (i, A, t) is the trust values of agent i with regard to agent A and with respect to task t. We assume trust values of each agent towards other agents to be common knowledge. This simplifying assumption is instigated from the fact that we are attempting to model a community of agents that is closed, in that the members of the community do not change over time. For agents in a social network, the trusting values are limited to individuals with ties. The amount of trust an agent has in itself determines its competence for performing a task. Obviously, equation 1 affects equation 2. Let's take an extreme example of two agents Henry and Sue with Sue benevolent towards Henry with respect to all tasks, but Henry is not benevolent towards Sue on any task. Equation 1 will promote the trust of Henry in Sue. If Sue accepts delegated tasks and successfully executes them,

Sue increases her autonomy following equation 2. The benevolent agent, Sue, gains in autonomy. Henry's autonomy has little chance of increase since it is not trusted by Sue.

The social connections among the agents in a social network influence the willingness of an agent to delegate a task. Consider the cattle herding behavior of the East African Orma where cattle owners make sure a close relative is on the cattle camp [17]. Such trust of kin has motivated owners to adopt young men. This is a form of generalized reciprocity. With increased relationship, the owners keep loose account of services and payments. They may reward their loyal hired help by paying them bridewealth after many years of loyal service or marrying the hired herder to one of their own daughters [17]. Our model of trust combines tit-for-tat reciprocity; i.e., process-based trust with this generalized reciprocity. Although we agree on competence being important in forming trust, we feel that Castelfranchi and Falcone's definition could apply to complete strangers and does not properly account for the interpersonal sense of trust [12].

Social exchanges affect interpersonal trust based on the types of connections an agent has and the variation in the connections. Benevolence among agents is modeled when agents successfully accept and complete a delegated task. The change in benevolence is the model of alliance among agents. This is combined with balance of reciprocity. It hypothesizes that agents strengthen their ties as they interact while performing their assigned tasks, and delegate tasks to others who are benevolent. Strengthening social ties between agents increases their interpersonal trust.

Variations of trust and autonomy when agents are engaged in reciprocal exchanges are found in [41]. The objects of exchange are requests for action and free agreements to perform actions, i.e., favors. In reciprocal exchanges, agents initiate an exchange without knowing whether, when, or to what extent the other will reciprocate in the future. Thus, agents offer benefits unilaterally, and the exchange does not depend on negotiation. Any expectation of return is implicit and not reasoned. Exchange networks are states of interaction among a group of agents engaged in reciprocal exchange. These states change over time with each interaction. Exchange networks and proposed trust networks based on reciprocal exchanges are dynamic. Exchanges among agents may lead to familiarity and a social relation. However, general social ties capture a broader sense of closeness between partners, and exchanges may or may not contribute to social distance between individuals. For example, responding to a call for a charitable cause does not necessarily bring the giver and receiver any closer. Later in the discussion it is warranted that a social network must be derived from an exchange network under further assumption about social conditions.

A relevant point is that different levels of reciprocal exchange contribute to different trust and autonomy values, which are used as part of the decision to perform

a task or to delegate it. Another point is that average autonomy and trust among individuals in a group is highest when there are a certain number of social connections in the groups. Just as there is a curvilinear relationship between the number of ties in the organization's structure and their performance, too many or too few ties degrade trust and autonomy.

Assuming agents relate interpersonally, using values on exchange networks, we define a value that reflects harmony in delegation. Denoted by DelHarmony (A, B), this indicates the number of times agents A and B have agreed to delegation requests from one another after internally deliberating over all the necessary considerations, divided by the sum of the number of times agents A and B have made a delegation request to each other. Typically, DelHarmony is considered to increase with time, assuming that agent interactions increase with time. Delegating a task is an opportunity to increase delegation harmony. The DelHarmony value ranges from 0.0 to 1.0. When this value is 1.0 these agents are in perfect harmony, honoring each other's delegation request. The value of 0.0 is when they have never agreed on delegation or have never interacted. DelHarmony is computed only when agents interact at the interpersonal level. This condition is established when agents are considered to be in a social network explicitly. Many relationship types do not fit this condition such as when someone is making an anonymous purchase from a generic store.

The ranges of all capability, benevolence, trust, autonomy, and power are set from 0 to 10. In order to normalize trust and autonomy equations, coefficients are used in our models, i.e., equations 1-6. Equation 3 captures the revised trust computation value:

$$\text{Trusting value (A, B, t)} = (1/3) * [\text{capability (B, t)} + \text{benevolence (B, A, t)} + 10 * \text{DelHarmony (A, B)}] \quad (3)$$

If we do assume agents relate at an interpersonal level, using values on exchange networks, we define a value that reflects the balance of reciprocity they have with other agents. Balance of reciprocity(A) for an agent A is computed by adding two values and subtracting two values:

- Add the number of times delegated tasks by agent A have been agreed upon.
- Add the number of times agent A has made a delegation request regardless of accepting that request.
- Subtract the number of times agent A has agreed to a delegation request by another agent.
- Subtract the number of times agent A has been asked for delegation regardless of whether agent A has agreed to work on the task.

The sole consideration of an agent's ability for performing the tasks might not evaluate an agent's autonomy reasonably in an exchange network. A good rate of exchange among the agents will increase an agent's autonomy significantly. Hence, the two values in the Balance of reciprocity(A) that relate to an agent's ability were subtracted. Autonomy value is revised in equation 4 to account for social exchange. The number of agents, k, in the group are the agents with whom agent A engages in reciprocity. This value will count each agent once whether it is a delegatee or a delegator and it will not double count agents.

$$\text{Autonomy (A, t)} = (1/3) * [\text{capability (A, t)} + \text{Average (T)} + 10/k * \text{Balance of reciprocity(A)}] \quad (4)$$

Next, we consider the impact of norms as an interesting attribute in determining the behavior of agents. The principles of social norms have various impacts in modeling agents' trust and autonomy in a multi-agent system [54]. This paper explores variations of trust and autonomy when agents adhere to a set of pre-defined norms. A norm can be understood as a behavioral constraint on agents that interact with each other in a multi-agent society [42]. Norms are common in human society as well. Examples include norms of traffic, norms of business, norms of schools and many other environments. Norms can be represented as constraints, goals, or obligations that guide, control, or regulate proper and acceptable behavior among agents [27] and can even serve as filters that generate and select goals [33]. Lindbeck claimed that social norms may be regarded as giving rise to purposeful or "rational" behavior [37]. Castelfranchi has argued that a deliberative normative agent with norms impacts goal and plan generation and selection [11]. We define a notion of norm commonality we call Norm Affinity as the number of norms two agents share in common. Restated, norm affinity is the closeness between two agents over time as they share norms in common and is calculated using the following expression:

$$\text{Norm Affinity (A, B)} = \text{cn} / ((\text{Total number of norms that agent A follows} - \text{cn}) + (\text{Total number of norms that agent B follows} - \text{cn}) + \text{cn})$$

Here cn is the total number of norms that agents A and B have in common. Norm Affinity ranges from 0.0 to 1.0.

As stated earlier, trust is a dyadic relationship. Social ties are considered to be established among agents who are assumed to relate at interpersonal level [30]. These ties become stronger (if mutual benevolence prevails) with time. When these ties are combined with norms, interactions are controlled by norms that are collective rather than dyadic [24]. This is because all agents are guided by following the outlined norms. If any agent

does not follow the norms, then that particular agent will experience sanctions that might affect its autonomy level.

Cook and Hardin offer an alternative explanation for combined synergy between dyadic trust and social ties [17]. In this alternative, communal norms substitute the need for dyadic trust. Instead, the majority generally follow common norms and consequently this controls social order just as interpersonal relationships would. This is evident in the contrast between banking systems in the developed world where communal norms function instead of interpersonal trust and ties as in the financial institutions of the newly independent Russian Republic. Another analogy is the nature of evolution of communal laws. Communal laws are developed due to strong ties and norms of cooperativeness. This offers equivalence between a society governed by a network of dyadic trust and one that relies on a combination of social ties and norms. Trusting value is revised in equation 5 to include Norm Affinity.

$$\text{Trusting value (A, B, t)} = (1/4) * [\text{capability (B, t)} + \text{Benevolence (B, A, t)} + 10 * \text{Delegation Harmony(A, B)} + 10 * \text{Norm Affinity (A,B)}] \quad (5)$$

Only particular norms, when violated, produce sanctions. We consider a norm only when it is followed by at least half the number of agents. This assumption is based on observation of naturally occurring phenomenon. This fact also distinguishes norms from the notion of rules. While it is mandatory to follow rules, complying with norms is a choice. Hence we justify our assumption for considering only those norms that are followed by at least half the number of agents, by asserting that these norms are the more socially accepted ones in an agent community. The association of violation with sanctions is explained by the notion of *social control* as an incessant local (micro) activity of its units, aimed at restoring the regularities, prescribed by norms [6, 13]. The importance of punishment (i.e., sanctions) for the success of societies in evolutionary competition is argued in [31]. We define Fear of Norm in the following equation:

$$\text{Fear of Norm (A)} = (m/n) * (1/m) * \sum_{i=1}^m (S_i)$$

$$\text{i.e. Fear of Norm (A)} = (1/n) * \sum_{i=1}^m (S_i)$$

Here m/n is the ratio of number of norms that an agent has adopted to the total number of norms in the society. S<sub>i</sub> is sanction imposed by each norm and is summed over

all norms that an agent may follow. Agents use the trusting values to evaluate their own autonomy in deciding and performing the tasks. Equation 6 presents a revised autonomy value computation.

$$\text{Autonomy value (A, t)} = (1/4) * [\text{capability (A, t)} + \text{Average(T)} + 10/k * \text{Balance of reciprocity(A)} + 10 * \text{Fear of Norm (A)}] \quad (6)$$

Norms introduce a stabilizing effect on agent behavior [5]. We define Stability(A) of an agent's behavior as a ratio of the number of norms that an agent follows for a task to the total number of norms that govern that task in the society. Stability for an individual directly contributes to stability of a group of agents [44]. There is a lot of work reported on stability in multi-agent systems [2, 14]. Later in this paper, we draw attention to stability of different groups of agents that follow different number of norms in performing their tasks. Agents in each group perform tasks while under influence of assigned norms. When an agent selects a task, a test is performed to check if the agent is following the prescribed number of norms for the assigned task. For each agent in the group stability is calculated and later the average stability of the group is computed.

### 3. Simulation

The simulation environment we have implemented in Java consists of N agents and N tasks where each agent considers the same task repeatedly, i.e. each agent has a fresh version of the same task in each simulation cycle. However each agent does not have to perform the task assigned to it. Agents may perform tasks assigned to them or delegated by other agents. In this simulation it is assumed in general that as agents perform tasks, their trust, capabilities, and autonomy change.

Each agent compares its own autonomy to perform its predefined task to its assessment of trust in other agents with respect to the task. If other agents are perceived to have higher qualifications, the agent selects the agent whom it trusts the most for delegation. After a period of negotiated delegation, agents settle to perform tasks that need completion. In a social network the negotiation is limited to agents with direct social ties. For agents at differing power levels, power becomes an arbitration factor. For agents with equal autonomy their capability becomes a key factor. Agents tend to perform tasks for which their autonomy is highest. If every factor is equal, arbitration resorts to random selection.

We will describe our model of social power before presenting our algorithm. Power is an attribute that affects trust and autonomy. Power does not parametrically contribute to models of trust and

autonomy, but it affects reasoning about task selection. By this we mean that power affects trust and autonomy in an indirect fashion. Social power draws attention to the differences in influence between individuals and groups. In this paper, we consider power as it influences delegation decisions. It is assumed to be predefined and static. Social power in general can take on different forms such as economic power, knowledge as power, personal power, physical power, situational power etc. Our notion of power focuses on relative power. Power is considered a relative influence among agents when actions are delegated. A form of power as influence is given in [7]. Also models of power and trust in organizations are found in [25]. An agent's influence value is treated as uniform with respect to all other agents. This type of power is akin to ranks in an organization that creates hierarchy among agents. Power levels distinguish agents from one another.

In multi-agent systems, organizations are modeled with the existence of varying group sizes and relative power among individuals in mind. Studies of sociality in multi-agency strive to provide individual agent decision making with abilities to account for interaction and group awareness. Autonomy, trust, and power are among the social notions that are used to qualify interaction and awareness. Differing power levels and group sizes influence trust and autonomy among agents. The key point is that different levels of power contribute to different trust and autonomy values, which in turn are used as part of the delegation decision. Average autonomy and average trust in groups are compared with different power levels. Subgroups of agents with low and high power experience higher autonomy and trust than those agents with medium power in the group. The principle of social entropy given is extended to *power homogeneity* and we use it in our comparisons [3]. The power homogeneity of the agents were used to experiment with two conditions: commitment among agents to delegate a task with varying power homogeneity levels and testing whether autonomy can be used to predict the commitment trends.

We devise a metric that differentiates a group with agents at different power levels. Tucker Balch's social entropy differentiates the agents with different group size; however, it does not include power [3]. We extend the principle of social entropy to differentiate among groups having different power profiles. For simplicity, the power homogeneity is represented as PH and is computed by the following expression.

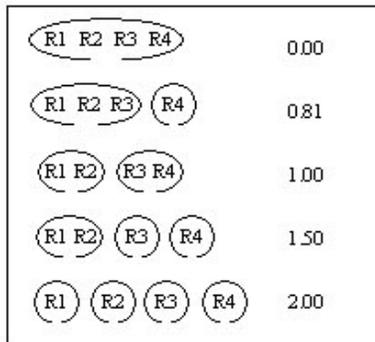
$$\text{PH} = \sum_{i=1}^m (p_i + w_i) \log_2 (p_i + w_i) \quad (7)$$

where  $m$  is the number of groups,  $p_i$  is defined as the ratio of the number of agents in a group  $i$  over the total number of agents.  $w_i$  is defined as the ratio of power level of the group over the maximum power level of all the groups.

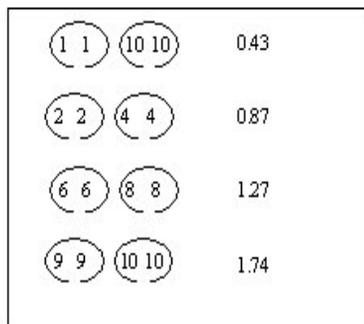
The following clarifies the power homogeneity. Figure 1 shows the social entropy of different groups where there are 4 agents. In power homogeneity no two groups share the same agent. R1, R2, R3, R4 are the agents. Figure 2 shows the power homogeneity of agents having social entropy of 1.00 but different power levels. There are two groups with two agents in each group. The numbers in ovals show their power levels.

There are two special cases to be considered for power homogeneity: First when there is a single group with all the agents having the same powers and second when all the agents are at different levels of power.

**Case 1:** When there is a single group, with all the agents having the same power, the social entropy of these agents will be 0.00 and the power homogeneity for such groups will be the same for the agents with different power levels. In order to differentiate among such groups at different power levels, we set the power homogeneity (PH) to the power level of the group.



**Figure 1.** Social entropy of different groups as discussed in [3]



**Figure 2.** Power homogeneity at different power levels with same entropy

**Case 2:** When powers of agents are different from one another, it is easy to contrast the power homogeneity of agents when the organization is very small. As the number of individuals in the group increases, comparison becomes harder. In the rest of this paper we will overlook such groups.

An example is considered to illustrate the power homogeneity expression among agents with different power levels but at the same social entropy level. Let there be four agents ( $A = \{A1, A2, A3, A4\}$ ) forming two groups (G1 and G2) with two agents in each group. The social entropy is 1.00. Let the power level of G1 be 1 and power level of G2 be 10. The power homogeneity as calculated from equation 7 is:

**i)** Power of G1 = 1, Power of G2 = 10

$$p_1 = 2/4 = 0.5, \quad w_1 = 1/10 = 0.1$$

$$p_2 = 2/4 = 0.5, \quad w_2 = 10/10 = 1$$

$$PH(A) = \frac{1}{2} \sum_{i=1}^2 (p_i + w_i) \log_2 (p_i + w_i)$$

$$= \frac{1}{2} ((p_1 + w_1) \log_2 (p_1 + w_1) + ((p_2 + w_2) \log_2 (p_2 + w_2)))$$

$$= \frac{1}{2} (((0.5 + 0.1) \log_2 (0.5 + 0.1)) + ((0.5 + 1) \log_2 (0.5 + 1)))$$

$$= 0.43$$

**ii)** Power of G1 = 2, power of G2 = 4

$$p_1 = 2/4 = 0.5, \quad w_1 = 2/4 = 0.5$$

$$p_2 = 2/4 = 0.5, \quad w_2 = 4/4 = 1$$

$$PH(A) = \frac{1}{2} \sum_{i=1}^2 (p_i + w_i) \log_2 (p_i + w_i)$$

$$= \frac{1}{2} ((p_1 + w_1) \log_2 (p_1 + w_1) + ((p_2 + w_2) \log_2 (p_2 + w_2)))$$

$$= \frac{1}{2} (((0.5 + 0.5) \log_2 (0.5 + 0.5)) + ((0.5 + 1) \log_2 (0.5 + 1)))$$

$$= 0.87$$

**iii)** Power of G1 = 9, power of G2 = 10

$$p_1 = 2/4 = 0.5, \quad w_1 = 9/10 = 0.9$$

$$p_2 = 2/4 = 0.5, \quad w_2 = 10/10 = 1$$

$$PH(A) = \frac{1}{2} \sum_{i=1}^2 (p_i + w_i) \log_2 (p_i + w_i)$$

$$= \frac{1}{2} ((p_1 + w_1) \log_2 (p_1 + w_1) + ((p_2 + w_2) \log_2 (p_2 + w_2)))$$

$$= \frac{1}{2} (((0.5 + 0.9) \log_2 (0.5 + 0.9)) + ((0.5 + 1) \log_2 (0.5 + 1)))$$

$$= 1.55$$

As shown above, the PH values increase with the increase in the power levels and decrease with decrease in the power levels of the groups. The values of the PH

Formatted: Polish

Formatted: Polish

Formatted: Polish

Formatted: Polish

do not have a range, as it varies with varying number of agents.

```

1. Initialize the values of capability matrix (C[]) to random values
   between 0 to 10.
2. Initialize the values of benevolence matrix (B[]) between 0.0 to
   10.0.
3. Initialize the values of tasknorms[]
4. Initialize the values of agentnorms[]
5. while (tasks remain) {
6.   for all agents and tasks {
7.     if (a == b) /* a, b stand for
agents*/
       A[a][t] = C[a][t] + average(T) + (10 * Balance of
reciprocity(a)) / k + 10 * Fear of Norm(a)
8.     else T[a][b][t] = C[a][t] + B[a][b][t] + 10 * DelHarmony(a, b) +
10 * Norm Affinity(a, b)
9.     for all agents and tasks {
10.      compare A[] with TV[] to find the most qualified agents for
tasks to delegate, delegate if not the most qualified }
11.    for all agents {
12.      compute the highest autonomy * power for each recurring
and delegated task
13.      execute the task with highest value }
14.    for all agents and tasks {
15.      compute the number of tasks being executed per iteration
and unsuccessful attempts.
16.      compute stability for each agent performing task
17.      update C[] with success and failures
18.      update B[] with success and failures }
19.    compute average trust and average autonomy for the group } }

```

This algorithm highlights application of trust and autonomy models.

- Average(T) is the average of trusting values of all agents with respect to self agent on particular task.
- B[] is the benevolence matrix of the agents.
- B[] initialized to 0.0 represents the lowest benevolence.
- B[] initialized to 10.0 represents the highest benevolence.
- TV[] is the matrix that holds the trusting values of agents with respect to tasks.
- S is the fear of norm
- n is the number of agents.
- tasknorms[] is the matrix that shows which norms govern a task.
- agentnorms[] is the matrix that shows which agents have adopted which norms.

**Figure 3.** Algorithm for computing average trust and average autonomy

An initial value of 0.0 is considered when we assume that there is no benevolence among the agents. Alternatively, an initial value of benevolence at level 10.0 is considered when a high benevolence among agents is assumed. The ranges of capability and trust are

between 0.0 and 10.0. Initially, capability values are randomly generated from 0 to 10. With unsuccessful attempts capability and benevolence of agents decrease whereas they increase with successful attempts. Power is pre-defined for all agents in the range from 0 to 10. A

task is performed in part based on the power of the agent that delegates the task relative to the power of the agent who is considering the delegated task.  $agentnorms[][]$  represent the binary agent-norm assignment matrix. It is binary in that it does not specify rates of adoption.  $tasknorms[][]$  is the task-norm binary matrix. The success or failure of an agent is determined by comparing the capability values of an agent with a randomly generated number ranging 0 to 10. If the random number has a value greater than the capability value of an agent, it is considered as a failure. If the number is lesser, then it is considered as a success. When an agent has a capability (say  $x$  units in performing a task  $t$ ), the agent can perform a task only when the task requires units less than or equal to its capability (i.e., greater than or equal to level  $x$ ). An agent may perform one task at any given time, and no two agents are allowed to do the same task in the same time unit. The success of a task is dependent only on the capability of the agent (as compared to the random number). The capability, trust, and relations among the agents are updated with the success or failure in performing the tasks. Average autonomy and trusting values of the agents are computed to observe possible relationships between the two notions with respect to the attributes that are considered.

The trusting value among the agents is computed with

$$T[a][b][t] = C[a][t] + B[a][b][t] + 10 * DelHarmony(a, b) + 10 * Norm Affinity(a, b).$$

The coefficients are used to normalize the equation. The autonomy of the agents is measured based on its capability in performing a task, average trust of the all other agents, balance of reciprocity with all other agents it has, and the fear of norm (i.e., the measure of fear when an agent does not follow the norm). Autonomy is calculated using the following expression

$$A[a][t] = C[a][t] + average(T) + (10 * Balance of reciprocity(a)) / k + 10 * Fear of Norm(a).$$

Capability is a random generation from 0 to 10 range and benevolence is set to zero initially, but varies with time. Autonomy and trusting values are compared to find a suitable task for an agent and the number for successful and unsuccessful attempts of the agents is computed. The capability and benevolence of agents change with the success or failure of the agent in performing the task.

At the core of our algorithm, agents select the tasks by comparing their  $A[][]$  and  $T[][]$  values. Unless an agent's autonomy value for a task is higher than all its trust values, agents delegate tasks to agents with the highest value of trust. Once agents receive delegations from all others, they compute their autonomy with respect to delegated tasks. For each delegated task, an agent has both a power level (of delegator agent) and autonomy value. For recurring tasks, the agent's own power level is

considered. If power is not a factor among agents, agents will select a task for which they have the highest autonomy value. When power and autonomy vary from task to task, autonomy value of a task is multiplied with the power associated with that task and then tasks are ranked according to this new value. The task with the highest rating is the one chosen for execution.

Finally, we consider commitment to repeated exchange as another parameter we quantify. This does not affect trust and autonomy. However, it parallels results in trust and we include it here for an interesting aside. An agent selects a task and decides to perform that task for another agent, taking power, autonomy and capabilities into consideration. Since a combination of factors is considered, an agent's decision might not always coincide with capability or trust. When exchange between agents coincides with trust and does not make sense in terms of capabilities then we say it is a case of *commitment*. Giving trust primacy over capability, is commitment. When the difference in capabilities between agents in a delegation exchange is greater than the difference in their trust then the exchange is said to be between trusted partners, which is motivated by commitment, otherwise it is considered to be a non-commitment or missed commitment.

Let  $C$  represent the capability,  $T$  represent the trusting value, and  $\Delta$  stand for difference in the values. Then:

$$\begin{aligned} \Delta T > \Delta C &\rightarrow \text{Commitment} \\ \Delta C > \Delta T &\rightarrow \text{non-Commitment} \end{aligned}$$

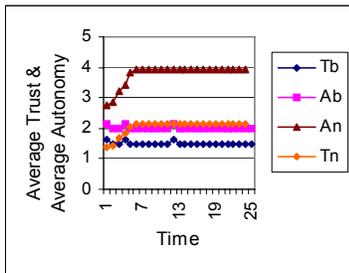
where  $\Delta C$  represents difference in the capabilities and  $\Delta T$  represents difference in the trust between the participating agents. Commitments among the agents are proportionally related to agent autonomy levels. Also, commitments are proportionally related to power homogeneity. Experimental results discussed in section 5 support these two hypotheses.

#### 4. Results and Discussions

This section presents results of trust and autonomy that are affected by the reciprocal social exchanges, power, and norms separately using our abstract simulation of agents and tasks. We performed adequate sets of experiments to achieve statistical validity. In the first set of experiments, effects of reciprocal social exchanges on trust and autonomy were observed, all at 25th time cycle. The average autonomy, average-trusting values, number of successful tasks and the number of unsuccessful attempts were recorded. In the first set of experiments, autonomy and trust of agents in two different social networks are compared.

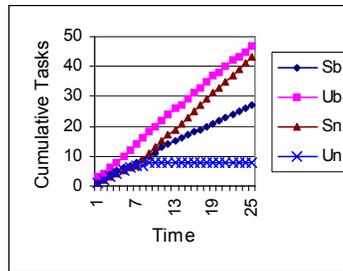
The average trust and average autonomy of the agents were computed with and without social exchange factors among the agents in a network, i.e., comparison of results between equations 1 and 2 versus equations 3 and 4. In the case where we considered social exchanges

among the agents, delegation harmony existed among the agents. Figure 4 shows results of agents with and without social exchanges. From Figure 4 we observe that the average trust and average autonomy values without any social exchanges using equations 1 and 2 (marked Tb and Ab in Figure 4) were very low. These values change slightly over time. There was a gradual increase in both curves that reached a constant level after a certain time. The average trust and average autonomy values where social exchanges take place were significantly greater using equations 3 and 4 (marked Tn and An in Figure 4). There was a gradual increase in both curves that might reach a constant level or vary over time. This is because of the exchanges and delegation harmony that agents have among them. The average autonomy climbed with time and reached almost a constant high level. The fluctuations in the curve may also be due to variations in capability of the agents. The magnitudes of the average trust and autonomy with social exchanges were consistently greater than those computed without any exchange among the agents (This observation is made by comparing Tn against Tb and An against Ab in figure 4). Benevolence value in both cases started at zero but changed as agents successfully delegated tasks to one another.



**Figure 4.** Average trust and average autonomy of two groups, one with social exchanges and the other group without social exchanges

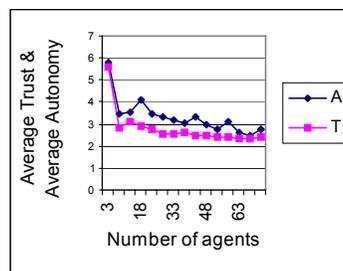
Figure 5 shows the results of cumulative task executions over time. It is observed that the number of successful tasks (Sn in Figure 5) where we considered social exchanges was higher than when agents have no exchange among them. The number of unsuccessful



**Figure 5.** Cumulative tasks with respect to time

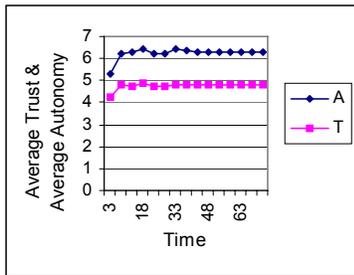
attempts (Un in Figure 5) was found to be very low when social exchanges were considered. Sb and Ub in Figure 5 are the number of successful and unsuccessful tasks obtained while social exchanges were not considered. These values show that when there were no social exchanges, it resulted in the lowest number of successful tasks being performed (Sb in Figure 5). At time  $t = 25$  the number of successful and failed tasks were 27 and 48 respectively. With the presence of exchange among the agents, the numbers improved to 43 successful and 8 failed tasks.

The second set of experiments were performed to observe the effects of social exchanges in varying sizes of agent community. In one experiment, we considered that there were no social exchanges among the agents, i.e., the agents have no delegation harmony. While in the other we considered there was the maximum level of delegation harmony among the agents. Average autonomy and trust values at the 25<sup>th</sup> unit of time were plotted for different population sizes. Figures 6 and 7 show autonomy and trust with various population sizes. Figure 6 shows the average trust and average autonomy of the agents in the absence of social exchange among the agents, i.e., using equations 1 and 2. It is observed that the average autonomy (shown as A in Figure 6) and average trust (shown as T in Figure 6) of the agents decreases with increase in agent population.



**Figure 6.** Average trust & average autonomy for each population size, at 25<sup>th</sup> time unit with no delegation harmony i.e., DelHarmony = 0.0.

In a network where the number of agents is three, the average trust and average autonomy were 5.61 and 5.81 respectively after 25 time units. With eight agents, the values were 2.83 and 3.5 respectively. These values clearly show that the average trust and average autonomy of the agents were lower in larger communities where there exist no exchanges among the agents in the network; there is no delegation harmony. Autonomy and trust is low and almost stable in populations of 23 or more.

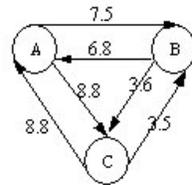


**Figure 7.** Average trust and average autonomy for each population size, at 25<sup>th</sup> time unit with a high Delegation harmony.

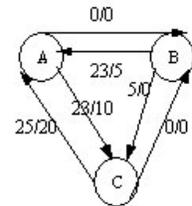
Figure 7 shows the results of the average trust and autonomy of the agents when there is high delegation harmony among the agents, using equations 3 and 4. It is observed from Figure 7 that both the average autonomy (shown as A in Figure 5) and the average trust (shown as T in Figure 7) of the agents increased with larger number of agents. Average autonomy and average trust of the agents attained an almost constant level with populations greater than 23. The average trust and average autonomy values were 4.25 and 5.25 respectively when the network had three agents. When the network consisted of eight agents, the values were 4.8 and 6.22 respectively.

Exchange network was discussed earlier. The term “exchange network” defines a network representation of agents where directed arcs denote the number of times an agent assigns a task to a particular agent and the number of times the assigned agent agree to the delegated task. It is observed that a trust network affects an exchange network. The exchange rate in the network is high among those agents that have high and equal trust for each other

and fewer among those agents that have low and equal trust for each other. Also, the chances of successful delegation will be high when an agent assigns a particular task to an agent that has more trust on it, i.e., if agent A has more trust on agent B than B on A, then the probability of agent A delegating the task to agent B will be more than agent B assigning a task to A. We describe how a social network can be derived from an exchange network. Figure 9 shows an exchange network of three agents A, B and C. From Figure 9 we see that the exchange rate between agent A and agent C is high. Figure 10 shows results from a social network where we have allowed all agents to be in personal relationship, i.e., the number of ties = 3.

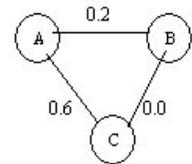


**Figure 8.** A trust network corresponding to Figure 4 at run = 25.



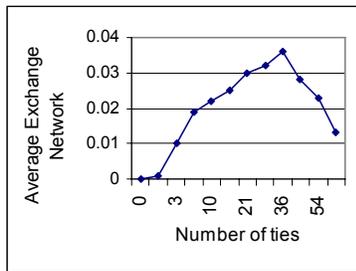
**Figure 9.** An exchange Network.

Figure 8 shows a graph that represents trust among agents. In Figure 9, just as the exchange rate between agent A and agent C is high, the trust between them is high and equal. Also, in the trust network, agent A has more trust on agent B and hence agent A delegates the tasks assigned by agent B. As trust between agent B and agent C is low and almost equal, exchange rate between agents B and C is zero.



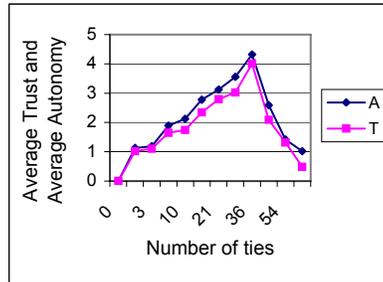
**Figure 10.** A social network corresponding to the exchange network in Figure 7

We performed experiments to observe the effect of varying social ties in the social network on trust, autonomy, exchanges, and performance of agents. Here, ties are considered to be established among agents who are assumed to relate at an interpersonal level. In this section DelHarmony (A, B) values are used to approximate strength of ties between agents A and B. As described in section 2, values on exchange networks were used to update harmony in delegation, i.e., DelHarmony (A, B) values. Figure 11 shows an average exchange network for a network of 12 agents. The maximum number of bi-directional ties in a network of  $n$  agents is  $(n*(n-1))/2$ . For  $n = 12$  the maximum number of ties is 66. The average exchanges among the agents in Figure 11 are average of all the DelHarmony values between the agents in the network. The values are recorded after the agents have had 25 time units of interaction.

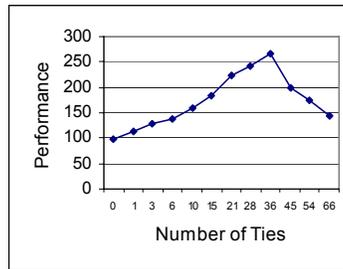


**Figure 11.** Average Exchange networks for 12 agents after 25 units of interaction at 0, 1, 3, 6, 10, 15, 21, 28, 36, 45, 54, 66 ties.

The average trust and average autonomy of the agents is low when only few ties exist among the agents. The trust and autonomy among agents is high when there are moderate ties among the agents within the network. The trust and autonomy levels increase gradually with increase in ties from low to moderate ties among the agents. I.e., it reaches a peak and then it drops. In a network with many ties among agents, the trust and autonomy of agents is once again on the average low. This pattern is mirrored in performance. The number of successful tasks performed by the agents will vary the same as trust and autonomy of the agents with varying ties in the network, i.e. the exchange is more common in a network of moderate level of ties coinciding with peaks in average levels of trust and autonomy.



**Figure 12.** Average trust and Average Autonomy



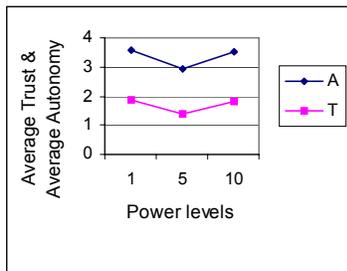
**Figure 13.** Number of tasks completed.

Figures 12 and 13 support our discussion. Figure 12 shows the average trust and average autonomy of the same network of 12 agents as in Figure 11. The variations of autonomy and trust against the number of ties we permit among the agents ranging from 0 to  $(n*(n-1))/2$  are shown. The average trust and average autonomy of the agents at the 25<sup>th</sup> time unit of each run is plotted for different number of ties in the network. The number of ties is the number of interpersonal relationships allowed to flourish. It is observed from the graph that the average trust and average autonomy curves gradually increase with increase in the ties from low range to medium range and then decrease as the ties increase further. From this it can be inferred that the average trust and average autonomy of the agents has a peak in the middle range of ties.

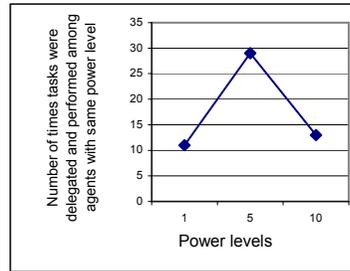
The performance curve of the agents with increasing number of ties among the agents in a network is shown in Figure 13. It is observed that the number of successful tasks performed by the agents is high when the agents are moderately connected when compared to the situation where the agents are either not connected or highly connected. This is a standard proposition well explored in sociology. It is shown that autonomy and trust values mirror that pattern. This is significant in that such peaks are not only beneficial for performance, but afford the highest trust and autonomy levels.

In a different set of experiments, results of average trust and average autonomy were observed with agents at different power levels. Two different experiments were performed here. In the first experiment, the results were obtained after 25 units of time. In each time unit, the average autonomy and average-trusting values were noted. Autonomy and trust of agents with three different power levels in the same group are considered and compared. The average trust and average autonomy of the agents were computed with group of agents that have equally low powers among them, group of agents with equally medium powers in second, and agents with equally high powers in the third group in the network. As the range of the power is set from 0 to 10, a value of 1 represents the lowest power, a value of 5 represents medium power and a value of 10 represents highest power. These powers are fixed among the agents in the groups over 25 simulation units. Figure 14 shows the results of different power levels in the same group that effect the agent's autonomy and trust.

From Figure 14 it is observed that the average trust and average autonomy values among agents with equally medium powers in the second group were lower than the first and third groups where power level among agents are lower and higher. This is because the agents in the group of medium power level tried to perform and assign most of the tasks to the agents at their own power level. The group of agents with low power level tried to perform more tasks for other agents, and that improved their ability in performing the tasks for other agents, thereby increasing their autonomy. As the agents in a group are improving in their autonomy levels, the trust among the agents within the group increases. The agents in the group of high power levels improved their autonomy and trust levels by convincing other power level agents to perform the tasks they assigned to them. Figures 15 and 16 support our discussion. Fluctuations in the curve are also due to variations in capability of the agents. Benevolence value started at zero but changed over time with the delegation of tasks by the agents.

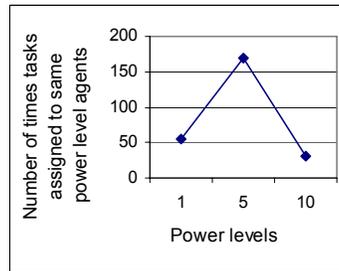


**Figure 14.** Average trust and Average autonomy of agents at different power levels.



**Figure 15.** Number of tasks that were delegated by, and performed by agents in the same power level.

Figure 15 shows the number of times an agent performed tasks for other agents in the same power level over 25 units of time. From Figure 15 it is observed that agents in the medium power level group tried to perform the most number of tasks for the agents with their own power level, while the other power group agents have performed fewer tasks determined by them. Figure 16 shows the number of times the agents assigned tasks for agents of same power level group over 25 units of time. From Figure 16 it is observed that agents in low and high power level groups assigned the tasks moderately fewer number of times to the agents that have the same power than the agents with medium power levels.

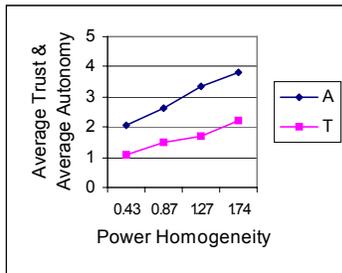


**Figure 16.** Number of tasks assigned for agents in the same power level group.

We say that autonomy and trust of agents that perform and assign tasks to the other agents with different power levels will improve faster as they gain the ability to perform different tasks for different power level agents and the ability to convince other agents to accept the assigned tasks.

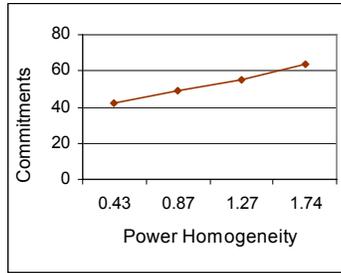
In the second set of experiments, average trust and average autonomy of the agents with different power homogeneity were observed. The results were observed at 25<sup>th</sup> time unit. Figure 17 shows how the average trust and average autonomy of the agents vary with different power homogeneity values but with the same social

entropy. The experiments were performed with 4 agents where there are two groups and each group has two agents. The social entropy for this set is 1.00. The two groups have different power levels. It is observed from Figure 17 that the average trust (T as shown in Figure 17) and average autonomy (A as shown in Figure 17) of the agents with different power homogeneity increases with increase in the power homogeneity. This is because as the level of power increases, agents becomes more autonomous in performing or getting tasks done by other agents. When the difference in the power levels of the groups is less, the agents are more feasible in understanding one another, and this increases their trust level when working with agents that have less difference in their power levels. The same principle holds for any group levels with same social entropy and different power homogeneity.



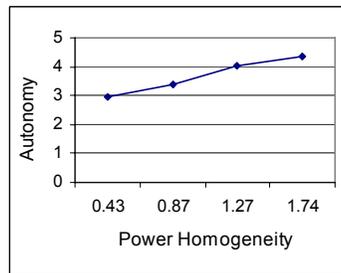
**Figure 17.** Average trust and Average autonomy of agents with power homogeneity.

In the third set of experiments, two tests were performed to observe how commitment and autonomy varies among the group of agents with different power homogeneity (PH). The experiments were performed with 4 agents in the group, and the results were observed at the 25<sup>th</sup> time unit. Figure 18 shows the results of a test that computed the level of commitment among the agents at different power levels. It is observed from Figure 18 that the level of commitment increases with increase in the power homogeneity. The level of commitment was 42 when power homogeneity was 0.43. When the power homogeneity was 4.02 the level of commitment was 55 and the commitment was 64 when the power homogeneity of the group of agents is 4.36.



**Figure 18.** Commitments at different power Homogeneity

Another test was performed to observe the relationship between autonomy and power homogeneity. Figure 19 shows average autonomy at different power homogeneity levels. It is observed that the average autonomy increased gradually with increase in power homogeneity among agents.



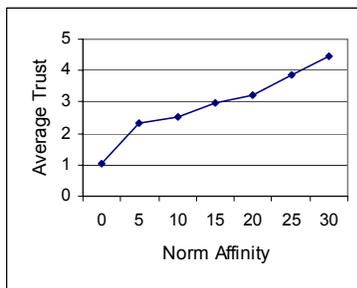
**Figure 19.** Autonomy at different power Homogeneity

The average autonomy at PH = 0.43 was 2.95. At PH = 0.87 the average autonomy was 3.41, and at PH = 1.74 the average autonomy was 4.36 respectively. Changes in autonomy predict changes in commitment among the agents at different homogeneity levels. In other words, autonomy is a predictor of commitments.

The remainder of this section discusses the results of average trust and average autonomy variations when norms are considered. Experiments were performed to observe variations in average trust and average autonomy of groups of agents with various influences under norms. In our simulation, the number of norms that an agent follows is randomly selected. Each norm is assigned to at least half the agents. As mentioned earlier we are not modeling norms with low adoption rates. Each norm is modeled to carry a sanction with a value that is randomly selected between 0 and -1. In the first experiment, the results were observed for 50 units of time where average autonomy of the agents is computed. In the second experiment the results were observed for 50 units of time

where average trusting values of the agents are recorded. As said earlier, each norm is followed by at least half the agents. Agents share norms in common when they are in a multi-agent system. Sharing of norms is termed *norm commonality*. Agents develop certain trust relations with other agents based on considerations of the norms that agents share in common. This is termed Norm Affinity. Average trust of agents was observed with varying norm commonalities and norm affinities among group of agents. Here again, Benevolence values started at zero but changed over time with the delegation of tasks by the agents. Power among the agents is fixed and is a randomly generated.

Figure 20 shows variations of average trusting value among a group of agents with varying norm affinity. The average trust among a group of agents increases with increasing norm affinity among the agents. This is because the more the agents interact while following the same norms, the more the agents are aware of one another, and this consequently increases agent trust levels. When the norm affinity among the group of agents is zero (i.e., Norm Affinity = 0) the average trusting value among the agents was 1.04. When the norm affinities among the group of agents were 10 and 15 the average trusting values among the agents were 2.54 and 2.96 respectively. When the norm affinities were 25 and 30 the average trusting values among the group of agents were 3.85 and 4.44 respectively. As the norm affinity among the agents increased the average trusting value among the group of agents increased.

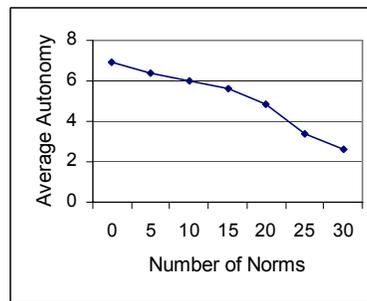


**Figure 20.** Average trusting value of different groups of agents that had different norm affinities were recorded at 50<sup>th</sup> cycle of simulation.

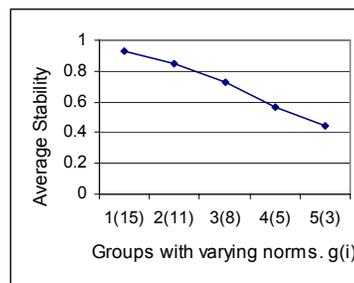
The greater the number of norms that an agent follows, the more the number of sanctions that exist. Each norm has its own sanction that when violated applies to the agent that transgressed that norm. The fear of norm has an impact on an agent's autonomy. As the number of norms the agents follows increases the average autonomy of the agents decreases. Figure 21 shows the average autonomy of a group of agents with varying norms. We performed experiments where norms were equally distributed among the agents. When agents have no

norms to follow (i.e., number of norms = 0) average autonomy of the agents was relatively high (i.e., average autonomy = 6.89). Average autonomy of the agents was 5.59 when agents followed the most number of norms (i.e., number of norms = 15). When the number of norms was 30, the average autonomy of the agents was 2.65. Hence, as the number of norms the agents followed increased, the average autonomy declined.

Another set of experiments was performed to observe the variations in *stability* of agents as the agents performed the tasks under various numbers of norms. Each assigned task is guided by a certain number of norms. The experiment recorded *stability* averaged over all members of agents in a given group when the group has differing patterns of norm adoption. The experiment was performed considering 5 different groups with different numbers of norms to follow how many agents in each group. The values of each group in the experiment were noted at the 25<sup>th</sup> cycle unit. Figure 22 shows the variations in *stability* across different groups. This graph illustrates that *stability* declines as the number of norms the groups of agents follow in performing a task decreases. The maximum number of



**Figure 21.** Average Autonomy of different group of agents at 50<sup>th</sup> cycle unit.



**Figure 22.** Stability curves of two different experiments where results were observed at 25<sup>th</sup> cycle unit for each group

norms prescribed in our experiment to perform a task was 15. The group of agents that followed all the norms attained an average stability of 0.93 and the group of agents that followed only 3 norms out of 15 norms exhibited an average stability of 0.44. 'g(i)' in the graph represents the number of norms (i) the agents in group g are following.

## 5. Conclusions

Model of autonomy and trust are offered that rely on several social attributes: benevolence between agents, a social exchange network where there exists delegation harmony among the agents, power levels among agents, and norms that agents followed. The models were intentionally kept simple to illustrate and emphasize the role of these social attributes individually.

With social ties, an agent would trust another agent with a task, only if the latter agent is capable of performing the task and the amount of relationship between the first and second agent is sufficient. An agent experiences autonomy with respect to a task if it is capable of performing it and it is trusted by other agents with regard to that task, along with the balance of exchange it has with other agents. It was observed from the experiments that when social exchanges among the agents in a social network are weak, their autonomy and trust are lower than those of agents with high exchanges among the other agents. Also, when there is a good amount of exchange harmony in a social network, the trust and autonomy of the agents is higher as compared with that of the trust and autonomy of agents among whom there are fewer exchanges. With diminished harmony in a social network, fewer tasks are completed and vice versa. The latter makes sense in terms of fewer agents considering task delegation.

The role of relative power among agents on autonomy and trust was explored. An agent experiences higher autonomy with respect to a task if it is capable of convincing other agents to perform tasks on its behalf. From our experiments, we observed that both unusually high and unusually low power levels tend to produce higher trust and autonomy levels than at medium power levels. The notion of power homogeneity was shown to affect trust and autonomy. When the power homogeneity is high, the trust and autonomy levels are high as well. It was observed that the rate of commitment to interact with familiar partners mirrors changes in autonomy in groups of different power homogeneity.

The role of norm adoption with respect to the agents' tasks was explored with respect to autonomy and trust. Norms largely accentuate average trust among and attenuate average autonomy of a group of agents. *Stability*; i.e., predictability of agent behavior, is shown to decline as the number of norms socially prescribed is minimally adhered. We plan to explore the domain-neutral simulation described in this paper for other explorations for engineering agents that are endowed

with awareness and manipulation of their social reality. This will pave the way to build computational simulation tools in order to aid study of social cognition as a new discipline. Crafting and using social agents for our complex worlds is inevitable and we believe our paper guides us in that path.

## References

- [1] A. Abdul-Rahman and S. Hailes, Supporting Trust in Virtual Communities, In Proceedings Hawaii International Conference on System Sciences 33, Maui, Hawaii, 2000.
- [2] H. Balakrishnan and S. Srinivasan, Analyzing stability in wide-area network performance, In ACM SIGMETRICS Conference on Measurement and Modeling of Computer Systems, 1997.
- [3] T. Balch, Social Entropy: a New Metric for Learning Multi-robot Teams, In Proceedings Flairs, Daytona, 1997.
- [4] A. L. C. Bazzan, R. H. Bordini, and J. A. Campbell, Moral Sentiments in Multi-Agent Systems, In Proceedings of the 5<sup>th</sup> International Workshop on Agent Theories, Architectures, and Languages, 1998.
- [5] G. Boella and L. Lesmo, Deliberate normative agents, In Social Order in Multi Agent Systems, R. Conte and C. Dellarocas, eds., Springer Verlag, Berlin, 2001.
- [6] R. Boyd, H. Gintis, S. Bowles, and P. J. Richerson, The evolution of altruistic punishment, In Proceedings of the National Academy of Sciences(USA) 100, pp 3531-3535, 2003.
- [7] S. Brainov and H. Hexmoor, Quantifying Autonomy, In Agent Autonomy, Kluwer Academic Publishers, Dordrecht, pp 43-56, 2003.
- [8] S. Brainov and T. Sandholm, Power, Dependence and Stability in Multiagent Plans In American National Conference on Artificial Intelligence, 1999, Orlando, Florida, pp. 11-16.
- [9] C. Castelfranchi and Y. Tan, Eds., Trust in Virtual Societies, Kluwer Academic Publishers, Dordrecht, 2001.
- [10] C. Castelfranchi, Engineering social order, In Proceedings of Engineering Societies in Agents' World(ESAW00), Berlin, 2000.
- [11] C. Castelfranchi, F. Dignum, Catholijn M. Jonker, J. Treur Deliberate Normative Agents: Principles and Architecture, presented at Agent Theories, Architectures and Languages (ATAL-99), Orlando, Florida, 1999, pp.364-378.
- [12] C. Castelfranchi and R. Falcone, Principles of trust for MAS: Cognitive anatomy, social importance, and quantification, In Proceedings of the Third International Conference on Multi-Agent Systems, 1998, pp 72-79.
- [13] C. Castelfranchi and R. Falcone, Trust and Control: A Dialectic Link, In Applied Artificial Intelligence Journal, vol. 14, no. 8, 2000, pp 799-823
- [14] M. Chli, D. Wildez, P. Goossenaerts, J. Abramov, V. Szirbik, N. Correia, L. Mariano, and P. Ribeiro, Stability of Multi-Agent Systems, In IEEE Systems, Man and Cybernetics, 2003.
- [15] J. Christman and J. Anderson, Eds. Autonomy and the Challenges to Liberalism: New Essays, 2003.
- [16] R. Conte, R. Falcone, and G. Sartor, Agents and Norms: How to fill the gap, AI and Law, special issue on Agents and Norms, Kluwer Academic Publisher, 1999, pp 1-15

- [17] K. Cook and R. Hardin, Trust and Trustworthiness, In Trust in society, K. S. Cook, Ed., Russell Sage Foundation., New York, 2000.
- [18] C. Darwin, The decent of man and selection in relation to sex, John Murray, London, 1871.
- [19] M. J. Druzdzal, ESP: A mixed initiative decision-theoretic decision modeling system, In Working Notes of the AAAI-99 Workshop on Mixed-initiative Intelligence, 1999, pp. 99-106.
- [20] J. Engminger, Reputation, Trust, and the Principal Agent Problem, in: Trust in society, K. S. Cook, Ed., Russell Sage Foundation, 2002.
- [21] P. Fankhauser and T. Tesch, Agents, a broker, and lies, in: Proc. of Ninth Intl. Workshop on Research Issues on Data Engineering: Information Technology of Virtual Enterprises, 1999, pp.56-63.
- [22] B. Friedman and H. Nissenbaum, Software Agents and User Autonomy, in: Proceedings of the First International Conference on Autonomous Agents, 1997, pp 466-469
- [23] D. Gambetta, Can We Trust Trust ?, in: Trust: Making and Breaking Cooperative Relations, D. Gambetta, ed., Oxford Press, Oxford, 2000.
- [24] R. Hardin, Trust and trustworthiness, Russell Sage Foundation, New York, 2002.
- [25] C. Hardy, N. Phillips, and T. Lawrence, Distinguishing Trust and Power in Interorganizational Relations: Forms and Facades of Trust, In Trust Within and Between Organizations, C. Lane and R. Bachmann, eds., Oxford University Press, pp 64-87, 1998.
- [26] C. Heckman and O. J. Wobbrock, Liability for Autonomous Agent Design, In Autonomous Agents, 1998.
- [27] V. Henricus, Norm Autonomous Agents, PhD thesis, Department of Computer and Systems Sciences, Stockholm University, Sweden, 2000.
- [28] H. Hexmoor and G. Beavers, In Search of Simple and Responsible Agents, in: Proceedings of GSF Workshop on Radical Agents, 2002.
- [29] H. Hexmoor and P. Poli, Benevolence, Trust and Autonomy, In Proceedings of the International Conference on Artificial Intelligence, 2003.
- [30] H. Hexmoor and P. Poli, Effects of Reciprocal Social Exchanges on Trust and Autonomy, in: Proceedings of Third International Conference on Intelligent Systems Design and Applications, 2003.
- [31] H. Hexmoor and S. Battula, Discovering Human Desired Norms of Interaction Improves Social Adeptness in Agents, in: AAAI Spring Symposium on Human Interactions with Autonomous Systems in Complex Environments, 2003.
- [32] H. Hexmoor and J. Vaughn, Computational Adjustable Autonomy for NASA Personal Satellite Assistants, in: ACM Symposium of Applied Computing (SAC-02), 2002.
- [33] H. Hexmoor, Stages of Autonomy Determination, In IEEE Transactions on Man, Machine & Cybernetics, Part C Vol. 31, No. 4, 2001, pp. 509-517.
- [34] T. E. Hill, Jr., Autonomy and Self Respect, Cambridge University Press, 2001.
- [35] S. R. H. Joseph and T. Kawamura, Why Autonomy makes the Agent, in: Agent Engineering, J. Liu, ed. World Scientific Publishing, 2001, pp.7-28.
- [36] K. Krishna and A. bin Maarof, Hybrid Trust Management Model for MAS Based Trading Society, in: International Arab Journal of Information Technology(IAJIT), Vol I, No. 0, 2003, pp 60-90.
- [37] A. Lindbeck, Incentives and Social Norms in Household Behavior, in: American Economic Review, vol. 87 (2), 1997, pp. 370-377.
- [38] M. Luck and M. d'Inverno, A formal Framework for Agency and Autonomy, in: Proceedings of the First International Conference on Multi-Agent Systems, 254-260, AAAI Press / MIT Press, 1995
- [39] B. Magnus, Norms in artificial decision-making, Artificial Intelligence and Law, vol 7, no 1, 1999, pp. 17-35.
- [40] S. Marsh, Formalising Trust as a Computational Concept, Phd thesis, University of Stirling, Department of Computer Science and Mathematics, 1994.
- [41] A. M. Mohamed and M. N. Huhns, Multi-agent Benevolence as a Societal Norm, in: Proc. of the First Workshop on Norms and Institutions in MAS at Autonomous Agents, Barcelona, 2000.
- [42] L. D. Molm, Coercive Power in Social Exchange, Cambridge University Press, Cambridge, 1997.
- [43] J. Nicole, H. Saam, and H. Andreas, Simulating Norms, Social Inequality, and Functional Change in Artificial Societies, presented at Journal of Artificial Societies and Social Simulation (JASSS), Vol. 2, No. 1, 1999.
- [44] S. Olafsson, On the Stability of Strategies in Competitive Systems, International Journal of Systems Science, vol. 26, no 6, 1995, pp. 1289-1312.
- [45] P. A. Pavlou, Y. Tan, and D. Gefen, The Transitional Role of Institutional Trust in Online Interorganizational Relationships, in: Proceedings of the 36th HICSS Conference, 2003.
- [46] M. Prietula and K. Carley, Boundedly Rational and Emotional Agents Cooperation, In Trust and Deception in Virtual Societies, C. Castelfranchi and T. Yao-Hua, eds., 2001, pp. 169-193
- [47] D. V. Pynadath and M. Tambe, Automated teamwork among heterogeneous software agents and Humans, Journal of Autonomous Agents and Multi-Agent Systems, vol. 7, 2003, pp. 71-100
- [48] S. D. Ramchurn, C. Sierra, L. Godo, and N. R. Jennings, A Computational Trust Model for Multi-Agent Interactions Based on Confidence and Reputation, In Proceedings of 6th International Workshop of Deception, Fraud and Trust in Agent Societies, 2003, pp. 69-75.
- [49] S. D. Ramchurn, D. Huynh, and N. R. Jennings, Trust in multiagent systems, The Knowledge Engineering Review, 19(1), 2004, pp 1-25
- [50] J. Sabater, C. Sierra, Review on Computational Trust and Reputation Models, Artificial Intelligence Review, 24(1), 2005, pp. 33-60
- [51] J. B. Schneewind, The Invention of Autonomy: A History of Modern Moral Philosophy, Cambridge University Press, Cambridge, 1997.
- [52] S. Sen, Reciprocity: a Foundational Principle for Promoting Cooperative Behavior Among Self-interested Agents, In Proceedings of The Second International Conference on Multi-Agent Systems (ICMAS), 1996, pp. 322-329.
- [53] R. L. Trivers, The Evolution of Reciprocal Altruism, In Quarterly review of Biology, vol. 46, 1971, pp. 35-57.
- [54] H. Verhagen,

- On the learning of norms,  
In Proceedings of MAAMAW99, 1999.
- [55] L. Wagner,  
Dealing with Curious Players in Secure Networks,[online paper],  
April 2004, [September 2004], available at:  
<http://arxiv.org/ftp/cs/papers/0404/0404004.pdf>
- [56] W. Winsborough, K. Seamons, and V. Jones,  
Automated Trust Negotiation,  
In DARPA Information Survivability Conference and  
Exposition, vol. I, 2000, pp. 88-102.
- [57] H. Wong and K. Sycara,  
Adding Security and Trust to Multi-Agent Systems,  
In Proceedings of Autonomous Agents '99 Workshop on  
Deception, Fraud, and Trust in Agent Societies, 1999.
- [58] M. Wooldridge,  
Reasoning about Rational Agents: The Intelligent Robotics and  
Autonomous Agents Series,  
MIT Press Cambridge, Cambridge, Massachusetts, 2000.
- [59] M. Wooldridge and N. R. Jennings,  
Intelligent agents: Theory and practice,  
The Knowledge Engineering Review, 10(2), 1995, pp. 115–152.
- [60] B. Yu and M. P. Singh,  
A Social Mechanism of Reputation Management in Electronic  
Communities,  
in: Proceedings of Fourth International Workshop on Cooperative  
Information Agents, 2000.
- [61] H. Yuh-Jong,  
Some Thoughts on Agent Trust and Delegation,  
In Proceedings of fifth international conference on  
Autonomous agents, 2001.