

Organizations in Artificial Social Systems

Guido Boella¹ and Leendert van der Torre²

¹ Dipartimento di Informatica - Università di Torino- Italy. E-mail: guido@di.unito.it

² University of Luxembourg. E-mail: leendert@vandertorre.com

Abstract. In this paper we introduce organizations and roles in Shoham and Tennenholtz' artificial social systems, using a normative system. We model how real agents determine the behavior of organizations by playing roles in the organization, and how the organization controls the behavior of agents playing a role in it. We consider the design of an organization in terms of roles and the assignment of agents to roles, and the evolution of organizations. We do not present a complete formalization of the computational problems, but we illustrate our approach by examples.

1 Introduction

The basic idea of the artificial social systems approach of Shoham and Tennenholtz [12, 13] is to add a mechanism, called a social law, that will minimize the need for both centralized control and on-line resolution of conflicts. A social law is defined as a set of *restrictions* on the agents' activities which allow them enough freedom on the one hand, but at the same time constrain them so that they will not interfere with each other. Several variants have been introduced to reason about the design and emergence of social laws. Shoham and Tennenholtz use game theoretic approach and inherit the advantages and drawbacks of game theory. On the one hand they work in a well understood framework, agent interactions can be defined precisely and computational problems can be defined in a precise way, but on the other hand due to the used abstractions it is difficult to differentiate agents and to simulate complex systems.

Several extensions have been proposed to this game-theoretic approach to artificial social systems. Shoham and Tennenholtz [12] introduce off-line design of *useful* social laws for artificial agent societies, and Fitoussi and Tennenholtz [7] distinguish two criteria to choose social laws called *minimal* and *simple* social laws. Shoham and Tennenholtz [13] study the emergence of *rational* social laws in repeated games instead of their off-line design. Briggs and Cook [6] introduce so-called *flexible* social laws that can be violated if an agent cannot obey the law [1]. Moreover, Tennenholtz [14] introduces *stable* social laws as a kind of qualitative equilibrium, in the sense that agents can deviate from the law, but they do not want to do so when the other agents follow it. This approach deals with bridging social laws with conflict resolution. Brafman and Tennenholtz [5] study efficient learning equilibria in repeated games. Boella and van der Torre [3] introduce *enforceable* social laws by extending artificial social systems with a control system – called a normative system – that represents the (consequences of) social laws.

There is a significant body of literature on the role of organizations in multi-agent systems going back to the beginnings of the field in the late 70's [9]. Despite the popularity of the game theoretic approach to artificial social systems as initiated by Shoham and Tennenholtz, as far as we know organizational concepts have not yet been introduced in them. However, the use of organizations could be useful to explain the interaction of social laws and other social concepts such as roles and norms. Moreover, organizations allow to describe the system at different levels of abstraction. Finally, the introduction of organizational concepts leads to new interesting computational problems which can be defined and studied using the game-theoretic framework, such as an organizational design problem (decompose the organization into a set of roles such that the organizational goals are achieved if the roles' goals are achieved), a role assignment problem (assign real agents to roles such that goals of roles and thus goals of organization are achieved), etc. We are therefore interested in the following research questions in this paper:

1. How can organizations and roles be defined in artificial social systems?
2. How do role playing agents determine the behavior of an organization?
3. How does the organization use real agents playing roles in organization to enforce social laws (so-called defender agents)?
4. How can we define an organizational design problem in artificial social systems?
5. How can we define a role assignment problem in artificial social systems?
6. How can we model the evolution of organizations in artificial social systems?

To answer these questions, we use a model of artificial social systems and enforceable social laws developed in [3, 4] as an extension of Tennenholtz' stable social laws. The normative system is represented by a socially constructed agent. Roughly, a social law is in force when it can be extended to a stable social law. Design of social laws can be formalized as updating the utility function of the normative system. In this paper the organization is modeled as the normative system, i.e., a control system, and we extend the model of artificial social systems with another class of socially constructed agents called roles.

This extension with a normative system of the game-theoretic approach to artificial social systems builds on work in normative multiagent systems. This work formalizes norms as a kind of soft constraint, just like other approaches, but it also considers what happens when a norm is violated (and how this can be repaired), permissions and rights and their relations to obligations, how norms change in time, the negotiation of new norms, decision making in normative systems, the interaction among normative systems, norms as a coordination mechanism, etc. Moreover, as Searle [10] argues, a distinction can be made between two types of rules, a distinction which also holds for formal rules like those composing normative systems:

“Some rules regulate antecedently existing forms of behaviour. For example, the rules of polite table behaviour regulate eating, but eating exists independently of these rules. Some rules, on the other hand, do not merely regulate an antecedently existing activity called playing chess; they, as it were, create the possibility of or define that activity. The activity of playing chess is constituted

by action in accordance with these rules. The institutions of marriage, money, and promising are like the institutions of baseball and chess in that they are systems of such constitutive rules or conventions” ([10], p. 131).

For Searle, institutional facts like marriage, money and private property emerge from an independent ontology of “brute” physical facts through constitutive rules of the form “such and such an X counts as Y in context C” where X is any object satisfying certain conditions and Y is a label that qualifies X as being something of an entirely new sort. E.g., “X counts as a presiding official in a wedding ceremony”, “this bit of paper counts as a five euro bill” and “this piece of land counts as somebody’s private property”.

In this paper we consider how the behavior of an organization is determined by the behavior of role playing agents. We therefore consider how real agents enforce control in artificial social systems. To be able to reason about roles when no agent has been assigned to it yet, we distinguish between the possible behaviors of the role and the possible behaviors of the real agent. In a fully specified organization, the behavior of the organization is determined by the behavior of the roles, and the behavior of the role is determined by the behavior of the real agents. We formalize how role playing agents determine the behavior of a organization using the constitutive norms of the normative systems. The relations between the behaviors is represented by a counts-as relation among strategies, such that a strategy of a real agent can count as a strategy of a role, and a strategy of a role can count as a strategy of the organization. These counts-as rules are abstractions of constitutive norms defining the organization.

As an illustrating example of the kind of analysis which can be done in our extension of the game-theoretic approach to artificial social systems, consider some real agents and an organization controlling their behavior. The behavior of the socially constructed organization is determined by the real agents in the system, for example by a policeman. Moreover, the behavior of the policeman is also controlled by the normative system, so to get system stability, we may have an infinite set of agents each controlling one another. So there is a police-policeman controlling the policeman, a police-police-policeman controlling the police-policeman, etc. In reality there is no infinite set of agents. Instead it is assumed that at some point the agents are trusted, due to the bounded reasoning of agents, or we have two policemen, each controlling the other and thus breaking the infinite sequence. For example, assume that we have four agents $\{a_1, a_2, p_1, p_2\}$ and the usual kind of prisoner’s dilemma for $\{a_1, a_2\}$. Then we have two policemen $\{p_1, p_2\}$, where p_1 can punish a_1 if he defects, and p_2 can punish a_2 if he defects. Moreover, p_1 can punish p_2 if a_2 defects but p_2 does not punish him, and p_2 can punish p_1 if a_1 defects but p_1 does not punish him. Our game theory can be used to analyze under which conditions this solution works.

The layout of this paper is as follows. In Section 2 we discuss artificial social systems and stable social laws as introduced by Tennenholtz and colleagues, and we introduce our extension with explicit normative system and enforceable social laws. In Section 3 we discuss how role playing agents determine the behavior of the organization, and how the organization controls agents playing a role in it. In Section 4 we discuss the (top-down) design of organizations, in Section 5 the assignment of roles, and in Section 6 we discuss the evolution of organizations in artificial social systems.

2 Artificial social systems and social laws

Shoham and Tennenholtz [12] introduce social laws in a setting without utilities. They define also *rational* social laws [13] as social laws that improve a social game variable. A game or multi-agent encounter is a set of agents with for each agent a set of strategies and a utility function defined on each possible combination of strategies. We extend artificial social systems with a control system, called a normative system, to model enforceable social laws. Following Boella and Lesmo [2], the normative system is represented by a socially constructed agent called the normative agent or agent 0. In [3], the normative system is represented by the set of control strategies of agent 0, but not by a utility function.

Definition 1. A *normative game* (or a normative multi-agent encounter) is a tuple $\langle N, R, S, T, U_1, U_2 \rangle$, where $N = \{0, 1, 2\}$ is a set of agents, R , S and T are the sets of strategies available to agents 0, 1 and 2 respectively, and $U_1 : R \times S \times T \rightarrow \mathbb{R}$ and $U_2 : R \times S \times T \rightarrow \mathbb{R}$ are real-valued utility functions for agents 1 and 2, respectively.

We use here as game variable the maximin value, following Tennenholtz [14]. This represents safety level decisions, see Tennenholtz' paper for a motivation.

Definition 2. Let R , S and T be the sets of strategies available to agent 0, 1 and 2, respectively, and let U_i be the utility function of agent i . Define $U_1(R, s, T) = \min_{r \in R, t \in T} U_1(r, s, t)$ for $s \in S$, and $U_2(R, S, t) = \min_{r \in R, s \in S} U_2(r, s, t)$ for $t \in T$. The maximin value for agent 1 (respectively 2) is defined by $\max_{s \in S} U_1(R, s, T)$ (respectively $\max_{t \in T} U_2(R, S, t)$). A strategy of agent i leading to the corresponding maximin value is called a maximin strategy for agent i .

A social law is useful with respect to an efficiency parameter q if each agent can choose a strategy that guarantees it a payoff of at least q .

Definition 3. Given a normative game $g = \langle N, R, S, T, U_1, U_2 \rangle$ and an efficiency parameter q , we define a social law to be a restriction of S to $\bar{S} \subseteq S$, and of T to $\bar{T} \subseteq T$. The social law is useful if the following holds: there exists $s \in \bar{S}$ such that $U_1(R, s, \bar{T}) \geq q$, and there exists $t \in \bar{T}$ such that $U_2(R, \bar{S}, t) \geq q$.

A social law is quasi-stable if an agent does not profit from violating the law, as long as the other agent conforms to the social law (i.e., selects strategies allowed by the law).

Definition 4. Given a normative game $g = \langle N, R, S, T, U_1, U_2 \rangle$, and an efficiency parameter q , a quasi-stable social law is a useful social law (with respect to q) which restricts S to \bar{S} and T to \bar{T} , and satisfies the following: there is no $s' \in S \setminus \bar{S}$ which satisfies $U_1(R, s', \bar{T}) > \max_{s \in \bar{S}} U_1(R, s, \bar{T})$, and there is no $t' \in T \setminus \bar{T}$ which satisfies $U_2(R, \bar{S}, t') > \max_{t \in \bar{T}} U_2(R, \bar{S}, t)$.

When the set of strategies R of agent 0 is a singleton, then our definitions reduce to those of Tennenholtz [14]. With the extension of agent 0 representing the control system we define enforceable social laws as quasi-stable social laws in normative games where the strategies of agent 0 may have been restricted [3].

Definition 5. Given a normative game $g = \langle N, R, S, T, U_1, U_2 \rangle$, and an efficiency parameter e , a social law (i.e., a restriction of S to $\bar{S} \subseteq S$, and of T to $\bar{T} \subseteq T$) is enforceable if there is a restriction of R to $\bar{R} \subseteq R$ such that \bar{S}, \bar{T} is quasi-stable in the normative game $g = \langle N, \bar{R}, S, T, U_1, U_2 \rangle$.

In [4] we extend normative games with a utility function of agent 0, to represent the enforced norms. Since agent 0 is a socially constructed agent, in the sense of Searle [11], its utility function can be updated. In particular, the enforcement of a social law by $\bar{R} \subseteq R$ is represented by giving \bar{R} strategies a high utility, and $R \setminus \bar{R}$ strategies a low utility. Moreover, we vary the utility of agent 0 depending on the strategies played by the other agents, and by considering incremental updates of the utility function to represent the evolution of artificial social systems. Formally, we extend a normative game with a utility function $U_0 : R \times S \times T \Rightarrow \mathbb{R}$, we define $U_0(r, S, T) = \min_{s \in S, t \in T} U_0(r, s, t)$ for $r \in R$, and we define useful and quasi-stable social laws in the obvious way. Enforced social laws are defined as follows.

Definition 6. Given a normative game $g = \langle N, R, S, T, U_0, U_1, U_2 \rangle$, and an efficiency parameter e , a social law (i.e., a restriction of S to $\bar{S} \subseteq S$, and of T to $\bar{T} \subseteq T$) is enforced if there is a unique restriction of R to $\bar{R} \subseteq R$ such that $\bar{R}, \bar{S}, \bar{T}$ is quasi-stable.

The game in Table 1 illustrates that the computational problem to find quasi-stable laws corresponds in extended normative games to the identification of enforced social laws. The table should be read as follows. Strategies are represented by literals, i.e., atomic propositions or their negations. Each table represents the sub-game given a strategy of agent 0, represented by $\neg n$ and n , respectively. Agent 1 is playing columns and agent 2 is playing rows. The values in the tables represent the utilities of agent 0 (in italics), 1 and 2.

$\neg n$	p	$\neg p$	n	p	$\neg p$
q	<i>3,3,3</i>	<i>0,4,1</i>	q	<i>3,3,3</i>	<i>1,2,1</i>
$\neg q$	<i>0,1,4</i>	<i>1,2,2</i>	$\neg q$	<i>1,1,2</i>	<i>0,2,2</i>

Table 1. What is the enforced social law?

Agent 0 can play strategy $\neg n$ or n , agent 1 can play strategy p or $\neg p$, agent 2 can play strategy q or $\neg q$. When the normative system plays $\neg n$, the sub-game of agent 1 and 2 is a classical prisoner's dilemma. Intuitively, the strategy $\neg n$ corresponds to the state before the social law is introduced, and n corresponds to the introduction of a control system that sanctions an agent for deviating from p, q . For example, the utility of agent 1 in $\neg p, q, n$ (2) is lower than its utility in $\neg p, q, \neg n$ (4) due to this sanction.

When the normative system plays n , the agents are always worse off compared to the normative agent playing $\neg n$, all else being equal. Nevertheless, due to the dynamics of the game, the overall outcome is better for both agents. For example, in the sub-game defined by strategy $\neg n$, the only Nash equilibrium is 2, 2. Now suppose we set the efficiency parameter to 3, which means that all agents will be better off. If the normative system plays n , then the sub-game has a Nash equilibrium which is the (Pareto optimal) 3, 3. This explains why the agents accept the possibility to be sanctioned.

3 Role playing agents enforcing social laws

The behavior of the organization is determined by real agents. We formalize this intuition in our game theoretic setting such that in a fully specified organization, the organization can play only one strategy. Likewise, if a role is assigned to an agent, then the role can play only the unique strategy determined by the agent playing the role. However, if the organization has not been fully specified yet, or the role has not yet been assigned to agents, then the behavior of the socially constructed agents is not deterministic.

3.1 Organized games

To define that role playing agents determine the behavior of an organization, we assume that there is a set of counts as conditionals defined on strategies of agents, which decrease the set of strategies socially constructed agents can play. At first sight, it may seem strange that we relate the behaviors of agents with counts-as conditionals. Normally, at a much more detailed level of analysis, counts-as conditionals are used to say that a piece of paper counts as money, or that going through a particular kind of ceremony counts as marrying two people. In general, counts-as conditionals are used to create institutional acts, such as money, marriage, property, liability, etc. However, at our level of abstraction, this implies that the behavior of the organization is defined by counts-as conditionals. We do not claim that our notion of counts-as conditionals covers all possible notions of counts-as conditionals, but it is sufficient for our purposes.

The basic idea of an organizational structure with counts-as conditionals is as follows. First, an organizational structure is a relation $Org \subseteq 2^N \times N$ that relates a set of agents to an agent, such that $(A, i) \in Org$ means that set of agents A directly determines the behavior of agent i . The relation $\{(a, b) \mid \exists(A, b) \in Org, a \in A\}$ reflects a hierarchy on agents and is therefore anti-reflexive, anti-symmetric and anti-transitive. Moreover, for each $(A, i) \in Org$, a counts-as conditional from a set of agents A to agent i is a function from the set of strategies played by the agents A to a function on the strategies of agent i : $counts-as : \times_{k \in A} S_k \rightarrow (S_i \rightarrow S_i)$. Applying the counts-as conditional to a game results in a new game, defined by $U_j(s_1, \dots, s_i, \dots, s_n) = U_j(s_1, \dots, counts-as(s_A)(s_i), \dots, s_n)$. Applying all counts-as conditionals is formalized by applying all these rules from the bottom of the organizational structure to the top. The counts-as conditionals have to satisfy the following properties, which intuitively represents that the agents can no longer distinguish between the strategies within an equivalence class of agent i .

1. The strategies of the agents A determine an equivalence relation on the set of strategies of agent i . For all $(A, i) \in Org$ with associated $counts-as$, and for any set of strategies S_1 and S_2 of the agents A , the reflexive, symmetric and transitive closures on $counts-as(S_1)$ and on $counts-as(S_2)$ are the same equivalence relation. If A determines the behavior of agent i , then it is a universal relation.
2. For each equivalence class, all strategies of the equivalence class are mapped onto the same strategy. For all $(A, i) \in Org$ with associated $counts-as$, and all set of strategies S of the agents A , if $(s_1, s_2), (s_2, s_3) \in counts-as(S)$, then $s_2 = s_3$.

3. If there are two sets of agents that together determine the behavior of agent i , then the order of applying the counts-as rules to a game is irrelevant (to simplify the procedure). For all $(A, i), (B, i) \in Org$ with associated $counts-as_1$ and $counts-as_2$, we have $counts-as_1(S_A) \circ counts-as(S_B) = counts-as_1(S_B) \circ counts-as(S_A)$.

Definition 7. An organized game (or an organizational multi-agent encounter) is a tuple $\langle N, S_0, \dots, U_0, \dots, Org, counts-as \rangle$, where $N = \{0, 1, 2, \dots\}$ is a set of n agents, S_i is the set of strategies available to agent i , and $U_i : S_0 \times \dots \times S_{n-1} \rightarrow \mathbb{R}$ is the utility function for agent i , and Org and $counts-as$ are as defined above.

An organized game can be reduced to a normal game using the procedure described above. Thus, from the bottom of the organizational structure to the top, we apply the counts-as rules. We call the bottom of the organizational structure the real agents, and all the other agents socially constructed agents. Moreover, if the top of the organizational structure is unique, then we refer to it as the organization. If there are more socially constructed agents in between the real agents and the organization, then we call the socially constructed agents just above the real agents the roles, and the other socially constructed agents we call functional areas.

3.2 Illustration of organized games

We illustrate the counts-as conditionals first by extending the running example with a *defender* or police agent enforcing the control system. Assume a police agent 3 playing either the strategy to work w or not to work $\neg w$. We assume he is lazy and gives utility 10 to not working and utility 0 to working. The organizational structure is such that the behavior of the policeman determines the behavior of the organization, $Org(3, 0)$. There are no roles or functional areas. Moreover, assume that the associated counts-as conditional is that w counts as n , and $\neg w$ counts as $\neg n$. Thus, we have $counts-as(w)(x) = n$ and $counts-as(\neg w)(x) = \neg n$ for $x \in \{n, \neg n\}$.

If we reduce the game using the counts-as rules, the game in Table 2 results. The cells of the table have been extended with the utility of agent 3. The left table contains the utilities when agent 3 does not work and the right table represents the utilities when agent 3 does work. Since the utilities of agent 0 for all its strategies have become the same due to the counts-as rules, we only represent one of its strategies.

$\neg w$	p	$\neg p$	w	p	$\neg p$
q	3,3,3,10	0,4,1,10	q	3,3,3,0	1,2,1,0
$\neg q$	0,1,4,10	1,2,2,10	$\neg q$	1,1,2,0	0,2,2,0

Table 2. p, q is not an enforced social law

In this game the norm will not be enforced, because the policeman is lazy and will play $\neg w$ whatever the other agents play. From a role assignment perspective, assigning agent 3 to enforce the social law of the organization was not a smart choice. Moreover, the example illustrates that the organization should also contain mechanisms to motivate the agents playing a role in it. There two issues are discusses in further sections.

We now illustrate how the mechanism used to define the behavior of the organization can be used to define the behavior of roles and role playing agents. We assume that there is an additional socially constructed agent for the police role, as an abstraction from the real agent playing the police role. Assume that there is a role 4 which can play strategy police in role r and no police in role $\neg r$ respectively. We define this socially constructed role by giving utility 10 to r and utility 0 to $\neg r$. The organizational structure is such that the behavior of the police role determines the behavior of the organization, $Org(4, 0)$, and we no longer assume $Org(3, 0)$. Moreover, assume that the associated counts-as conditional is that r counts as n , and $\neg r$ counts as $\neg n$. Thus, we have $counts-as(r)(x) = n$ and $counts-as(\neg r)(x) = \neg n$ for $x \in \{n, \neg n\}$.

If we reduce the game using the counts-as rules, the game in Table 3 results. This figure can be read as follows. The cells of the table have been extended with the utility of role 4 (in italics, since it is a socially constructed agent). The left table contains the utilities when role 4 is not being played and the right table represents the utilities when agent 4 is being played. Agent 3 does not influence the game so we have not represented its utilities. In the case of an ideal police role, p, q is an enforced social law.

$\neg r$	p	$\neg p$
q	$3,3,3,?,0$	$0,4,1,?,0$
$\neg q$	$0,1,4,?,0$	$1,2,2,?,0$

r	p	$\neg p$
q	$3,3,3,?,10$	$1,2,1,?,10$
$\neg q$	$1,1,2,?,10$	$0,2,2,?,10$

Table 3. p, q is an enforced social law (enforced by r)

Moreover, assume that agent 3 is assigned to the role 4. The organizational structure is such that the behavior of the police role determines the behavior of the organization as before, $Org(4, 0)$, and now we also have that the behavior of the police agent determines the behavior of the police role, $Org(3, 4)$. Moreover, assume that the associated counts-as conditional is that as before r counts as n , and $\neg r$ counts as $\neg n$, but now also w counts as r , and $\neg w$ counts as $\neg r$. Thus, we have as before $counts-as(r)(x) = n$ and $counts-as(\neg r)(x) = \neg n$ for $x \in \{n, \neg n\}$, and now also $counts-as(w)(x) = r$ and $counts-as(\neg w)(x) = \neg r$ for $x \in \{n, \neg n\}$. If we reduce the game using the counts-as rules, the game in Table 4 results.

$\neg w$	p	$\neg p$
q	$3,3,3,10,0$	$0,4,1,10,0$
$\neg q$	$0,1,4,10,0$	$1,2,2,10,0$

w	p	$\neg p$
q	$3,3,3,0,10$	$1,2,1,0,10$
$\neg q$	$1,1,2,0,10$	$0,2,2,0,10$

Table 4. p, q is not an enforced social law

This figure can be read as follows. The cells of the table have been extended with the utility of role 4. The left table contains the utilities when role 4 does not being played and the right table represents the utilities when agent 4 is being played. As before in Table 2, with a lazy policeman, p, q is not an enforced social law.

4 Organizational design problem

We now consider the construction of socially constructed agents. The design problem consists of various sub-problems: we have to design the utility function of the organization (or normative system), the socially constructed agents such as functional areas and roles with their utility functions, and the organizational structure with the counts-as norms. While doing this, we have to ensure that the organizational goals are achieved. In the previous examples, for example, when we assigned the ideal police role the social law was enforced, but when we assigned a lazy police agent to the police role, the social law was no longer enforced. Likewise, if we would assign another role such as a secretary, probably the social social law would not be enforced either.

We now consider the example in the introduction, where, to get system stability, we may have an infinite set of agents each controlling one another. So there may be a police-policeman controlling the policeman, a police-police-policeman controlling the police-policeman, etc. In real life there is no infinite set of agents, for example because at each step of recursion control become easier (it is easier to check a policeman or a judge than a mafia boss), sanctions are easier to apply (just remove earning), and the controller is more motivated to stick to the rules since its power depends on the reputation of the institution he works in (no fun to be a policeman of a corrupted police: no one listen to you). Here we consider the possibility that police agents control each other.

So we consider an organization, agent 0, with four real agents, 1, 2, 3, 4. For simplicity we do not consider roles or functional areas. Agent 1 and 2 play the same prisoner's dilemma as before. The behavior of the socially constructed organization is determined by policemen 3 and 4. Agent 3 can play w_1 or $\neg w_1$, and agent 4 can play w_2 or $\neg w_2$. Moreover, the behavior of the policeman is also controlled by the normative system, in the sense that the two policemen are controlling each other. Consider the utility functions in Table 5, where we assume that the behavior of the organization is deterministic (i.e., determined by the two policemen).

$\neg w_1, \neg w_2$	p	$\neg p$	w_1, w_2	p	$\neg p$
q	3,3,3,5,5	0,4,1,5,5	q	3,3,3,3,3	1,2,1,3,3
$\neg q$	0,1,4,5,5	1,2,2,5,5	$\neg q$	1,1,2,3,3	0,2,2,3,3
				w_1	$\neg w_1$
			w_2	?,?,?,3,3	?,?,?,0,10
			$\neg w_2$?,?,?,10,0	?,?,?,5,5

Table 5. p, q is an enforced social law

This table should be read as follows. The first two tables are the same as before, besides the fact that they do not depend on w but on w_1 and w_2 , that is, whether both the agent 3 and agent 4 work. Both agents prefer not to work over working; we did not represent the case in which only one of them works in these tables. The third table details the utilities of the two policemen. When one of them works but the other does not, then the working agent gets a high utility and the one not working a low one.

This represents that the one not working is sanctioned by the other one. To keep things simple, the utilities of the policemen do not depend on the strategies of the two other agents.

If we consider only the last table, then we see that this is again a prisoner's dilemma. The two policemen would prefer not to work, but the only stable outcome is that they work. The reason is that if they do not work, they may be punished by the other police agent. This is independent of the strategies of the other agents. Consequently, we have w_1 and w_2 , and therefore the prisoner's dilemma for the first two agents is evaded too. Summarizing, p, q is again an enforceable social law.

It is instructing to consider the case in which agent 1 and 2 play $\neg p$ and $\neg q$, but agent 3 and 4 do not punish them, and they do not punish each other. Before our formal analysis, intuition might tell us that this should also be an equilibrium. However, it is not the case, because agent 3 and agent 4 cannot cooperate. Remarkably, the prisoner's dilemma for agent 3 and 4 has led to this solution, which shows that although the prisoner's dilemma may indicate a social problem in general, in some particular cases like the one under consideration, the prisoner's dilemma may be used to solve a social problem.

One may object to our analysis that agent 1 and 2 have increased their utilities by introducing a social law, so why don't agent 3 and 4 also create a social law to increase their utilities? The answer is, of course, that an artificial social system must be designed such that defender agents like our policemen cannot change the normative system. Agent 1 and 2 would suffer from such a new social law, and they should have the power to block it.

We have a huge freedom in designing organizations in this framework, and it is therefore at this point not clear how to define the organizational design problem as a computational problem. We leave this for further research.

5 Role assignment problem

The role assignment problem is to find a set of real agents such that the real agents determine the behavior of the roles, and enforce the social law. In other words, in the role assignment function a game is given, and we are looking for an organizational structure with counts-as conditionals.

At this moment, it may be useful to take a further look into our roles: what are they precisely? When there is no agent playing the role, then the role describes the ideal behavior of a role playing agent. If there is an agent playing the role, then the role describes the actual role playing agent.

From this interpretation follows that there can be only one agent playing a particular role. If there are for example two policemen, as in our running example, then we have to introduce two roles. This illustrates that roles in our setting are what is sometimes called a role instance.

6 Evolution of organizations

We may further extend the example by introducing another role and another agent who has the power to change the normative system. This is well known from political science, in particular from the separation of powers in the Trias Politica.

The social law design problem in this setting is, given a normative game, to define a new utility function for the normative system [4]. The principle that we like to maintain as much as possible from the existing social laws can be represented by the use of the principle of minimal change. Table 6 represents the evolution of an artificial social system by an incremental increase of the utility of agent 0 to the efficiency parameter of the new social law.

$\neg n_1, \neg n_2$	p_1	p_2	$\neg p_1, \neg p_2$	n_1	p_1	p_2	$\neg p_1, \neg p_2$
q_1	0,3,3	0,4,1	0,6,0	q_1	1,3,3	1,4,1	1,0,0
q_2	0,1,4	0,2,2	0,0,0	q_2	1,1,4	1,2,2	1,0,0
$\neg q_1, \neg q_2$	0,0,6	0,0,0	0,0,0	$\neg q_1, \neg q_2$	1,0,0	1,0,0	1,0,0
				n_2	p_1	p_2	$\neg p_1, \neg p_2$
				q_1	3,3,3	3,1,1	3,0,0
				q_2	3,1,1	3,0,0	3,0,0
				$\neg q_1, \neg q_2$	3,0,0	3,0,0	3,0,0

Table 6. Iterated design

The first table represents that the normative system does not impose a control system, the second table represents that there is a sanction for playing $\neg p_1, \neg p_2$ or $\neg q_1, \neg q_2$, and the third table represents that there is an additional sanction for playing something else than p_1 and q_1 . The first social law is $\bar{S} = \{p_1, p_2\}, \bar{T} = \{q_1, q_2\}$ based on control system $\bar{R} = \{n_1, n_2\}$, and the second social law is $\bar{S} = \{p_1\}, \bar{T} = \{q_1\}$ based on control system $\bar{R} = \{n_2\}$.

7 Related research

There is a lot of related research in organizational science, in multiagent systems, in normative multiagent systems, and in game theory. For example, evolutionary game theory can be used as an inspiration to make the ideas in this paper more precise. However, as far as we know, these game theories have not been reduced to artificial social systems setting we have considered in this paper.

There are also many systems where the issues discussed in this paper can be used to study the interaction among agents. For example, it can be used in the development of electronic institutions where self-interest agents (from different organizations) form virtual organizations (coalitions) to solve a users problem [8].

8 Concluding remarks

Enforceable social laws are a bridge between two important theories of social systems. On the one hand artificial social systems based on social laws, and on the other hand normative multiagent systems based on norms and deontic logic. In this paper we illustrate how organizations and roles, which were already defined in normative multiagent systems, can be defined in artificial social systems. Moreover, we illustrated how the game theory used in artificial social systems can be used to analyze the interaction among role playing agents in an organization. The use of organizations could be useful to explain the interaction of social laws and other social concepts such as roles and norms. Moreover, organizations allow to describe the system at different levels of abstraction. Finally, the introduction of organizational concepts leads to new interesting computational problems which can be defined and studied using the game-theoretic framework, such as an organizational design problem (decompose the organization into a set of roles such that the organizational goals are achieved if the roles' goals are achieved), a role assignment problem (assign real agents to roles such that goals of roles and thus goals of organization are achieved), etc.

Organizations, functional areas and roles are defined as agents in artificial social systems. The advantage is that we do not extend the ontology of the games we play, and we can use the same game theory as used in other artificial social systems. Role playing agents determine the behavior of an organization using counts-as rules, which transform a game into another game. The resulting game is again a standard game which can be used to determine quasi-stable and enforceable social laws. An organized game with organizational hierarchy and counts-as conditionals can be used to define new computational problems. The organization uses real agents playing a role in the organization by controlling their behavior too. We can define the control of such defender agents by using the game dynamics. In particular we have illustrated that we can use a prisoner's dilemma to ensure that agents do not cooperate to evade the control of the normative system.

Two kinds of organizational design problems are given. The first organizational design problem is given a game including the roles, find counts-as rules to associate roles to the organization such that its behavior is determined by the roles, and there is stable social law such that organizational goals are achieved. The second organizational problem is to create roles (with appropriate utilities) such that the first problem can be solved. The role assignment problem can simply be defined as associating the roles with the agents, such that the agents determine the behavior of the roles, and the organization goals are achieved. The evolution of organizations can be modeled by updating the utilities of the organization. Here additional principles can be accepted, such as minimal change criteria.

In future work, we intend to further formalize the various notions introduced in this paper, study the complexity of the decision problems, and test the model on some examples. We are also interested in further extensions of our model, for example for other organizational structures, or systems in which all powers of the Trias Politica are formalized.

References

1. E. Alonso. Rights and argumentation in open multi-agent systems. *Artificial Intelligence Review*, 21(1):3–24, 2004.
2. G. Boella and L. Lesmo. A game theoretic approach to norms. *Cognitive Science Quarterly*, 2(3-4):492–512, 2002.
3. G. Boella and L. van der Torre. Enforceable social laws. In *Procs. of AAMAS'05*, pages 682–689, New York (NJ), 2005. ACM Press.
4. G. Boella and L. van der Torre. The evolution of artificial social systems. In *Procs. of IJCAI'05*, pages 1655–1556. Professional Book Center, 2005.
5. R. Brafman and M. Tennenholtz. Efficient learning equilibrium. *Artificial Intelligence*, 159(1-2):27–47, 2004.
6. W. Briggs and D. Cook. Flexible social laws. In *Procs. of IJCAI'95*, pages 688–693, 1995.
7. D. Fitoussi and M. Tennenholtz. Choosing social laws for multi-agent systems: Minimality and simplicity. *Artificial Intelligence*, 119(1-2):61–101, 2000.
8. T. J. Norman, A. Preece, S. Chalmers, N. R. Jennings, M. Luck, V. D. Dang, T. D. Nguyen, V. Deora, J. Shao, A. Gray, and N. Fiddian. Conoise: Agent-based formation of virtual organisations. In *Procs. of 23rd SGAI International Conference on Innovative Techniques and Applications of AI*, pages 353–366, Cambridge, UK, 2003.
9. W.G. Ouchi. A conceptual framework for the design of organizational control mechanisms. *Management Science*, 25(9):833–848, 1979.
10. J.R. Searle. *Speech Acts: an Essay in the Philosophy of Language*. Cambridge University Press, Cambridge (UK), 1969.
11. J.R. Searle. *The Construction of Social Reality*. The Free Press, New York, 1995.
12. Y. Shoham and M. Tennenholtz. On social laws for artificial agent societies: off-line design. *Artificial Intelligence*, 73(12):231–252, 1995.
13. Y. Shoham and M. Tennenholtz. On the emergence of social conventions: Modeling, analysis and simulations. *Artificial Intelligence*, 94(1-2):139–166, 1997.
14. M. Tennenholtz. On stable social laws and qualitative equilibria. *Artificial Intelligence*, 102(1):1–20, 1998.