Role-based Rights in Artificial Social Systems

Guido Boella
Dipartimento di Informatica
Università di Torino
Italy
guido@di.unito.it

Leendert van der Torre
CWI Amsterdam
and Delft University of Technology
The Netherlands
torre@cwi.nl

Abstract

In this paper we use normative systems to introduce roles and rights in the game-theoretic artificial social systems developed by Shoham and Tennenholtz. We model normative systems as socially constructed agents whose behavior is determined by a set of role playing agents. Roles are again modeled as socially constructed agents, and the roles' behavior is the ideal behavior of agents playing the roles. In our approach, the strategies of the role correspond to the rights that can be exercised by the role. In other words, rights are powers extending the set of strategies of an agent – not constraining them! – due to the new opportunities to exercise rights. We consider the role assignment problem of how to assign agents to roles such that the role playing agent is expected to behave like the ideal behavior of the role. We also consider how the normative system controls the behavior of agents playing a role in it.

1. Introduction

The basic idea of the artificial social systems approach of Shoham and Tennenholtz [7, 8] is to extend classical game theory with 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. In this paper we raise the following questions:

1. How can roles and rights be defined in Shoham and Tennenholtz' game-theoretic artificial social systems, such that role-based rights are powers increasing the agent's possible strategies?
2. How can we define a role assignment problem in such artificial social systems?

Despite the popularity of Shoham and Tennenholtz' game-theoretic artificial social systems, as far as we know organizational concepts have not yet been introduced in them. A problem is that classical game theory seems too abstract to represent involved notions like roles and rights, because the theory only contains agents, strategies, and utility functions. However, we have already shown how to represent a normative system – by representing the normative system as a special kind of agent [3, 4]. We call it a socially constructed agent. We use the same approach to define roles. Roles are again modeled as socially constructed agents, and the roles' behavior is the ideal behavior of agents playing the roles.

Rights have been addressed in this game-theoretic setting by Alonso [1], in the sense that "a right is considered as a set of restrictions on the agent's activities which allow them enough freedom, but at the same time constrain them." He then continues to distinguish rights from social laws, and illustrates his notion of rights by an example from traffic law, where rights prevent two cars driving into each other on a crossroads. However, we believe that this example should be modeled with obligations or prohibitions instead of rights, and in general that the characteristic property of a right is its power to *increase* the set of possible agent strategies. Moreover, in our approach role assignment means that an agent can decide the strategy the role is playing, i.e., can decide which rights are exercised. The set of strategies and thus the agent's freedom and autonomy may be decreased by the role's responsibilities and obligations, but that is another story.

The layout of this paper is as follows. In Section 2 we discuss our extension of artificial social systems with an explicit normative system and enforceable social laws, as introduced in earlier work. In Section 3 we introduce roles and rights, and in Section 4 we discuss the role assignment problem.

2. Artificial social systems and social laws

In this paper we use a model of artificial social systems and enforceable social laws developed in [3, 4] as an extension of Tennenholtz' stable social laws [9], which we briefly repeat in this section. Shoham and Tennenholtz [7] introduce social laws in a setting without utilities. They define also *rational* social laws [8] 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 [3, 4]. 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. As observed by Tennenholtz [9, footnote 4], the extension of a two player game to n player game is straightforward, and due to space limitations we do not give the details; we write R for the set of strategies of the normative system.

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 $U_1 : R \times S \times T \rightarrow I\!R$ and $U_2 : R \times S \times T \rightarrow I\!R$ are real-valued utility functions for agents 1 and 2, respectively.*

We use here as game variable the maximin value, following Tennenholtz [9]. 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 $U_2(R, S, t) = \min_{r \in S, s \in S} U_2(r, s, t)$ 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 $\overline{S} \subseteq S$, and of T to $\overline{T} \subseteq T$. The social law is* useful *if the following holds: there exists $s \in \overline{S}$ such that $U_1(R, s, \overline{T}) \geq q$, and there exists $t \in \overline{T}$ such that $U_2(R, \overline{S}, t) \geq q$.*

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 $\overline{S}$ and T to $\overline{T}$, and satisfies the following: there is no $s' \in S \setminus \overline{S}$ which satisfies $U_1(R, s', \overline{T}) > \max_{s \in \overline{S}} U_1(R, s, \overline{T})$, and there is no $t' \in T \setminus \overline{T}$ which satisfies $U_2(R, \overline{S}, t') > \max_{t \in \overline{T}} U_2(R, \overline{S}, t)$.*

When the set of strategies R of agent 0 is a singleton, then our definitions reduce to those of Tennenholtz [9]. 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 q, a social law (i.e., a restriction of S to $\overline{S} \subseteq S$, and of $\overline{T} \subseteq T$) is* enforceable *if there is a restriction of R to $\overline{R} \subseteq R$ such that $\overline{S}, \overline{T}$ is quasi-stable in the normative game $g = \langle N, \overline{R}, S, T, U_1, U_2 \rangle$.*

In [4] we extend normative games with a utility function of agent 0, to represent the enforced social laws. Since agent 0 is a socially constructed agent, using Searle's terminology [6], its utility function can be updated. In particular, the enforcement of a social law by $\overline{R} \subseteq R$ is represented by giving $\overline{R}$ strategies a high utility, and $R \setminus \overline{R}$ strategies a low utility. Moreover, we go beyond the framework of enforceable social laws by varying 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 $U_0 : R \times S \times T \Rightarrow I\!R$, 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 q, a social law (i.e., a restriction of S to $\overline{S} \subseteq S$, and of $\overline{T} \subseteq T$) is* enforced *if there is a unique restriction of R to $\overline{R} \subseteq R$ such that $\overline{R}, \overline{S}, \overline{T}$ is quasi-stable.*

Summarizing, the normative system is represented by a socially constructed agent. Design of social laws can be formalized as updating the utility function of the normative system. Computational problems can be defined to find enforceable social laws (with respect to an efficiency parameter). For further motivation and discussion, consult the above mentioned papers.

3. Role-based rights

Despite the fact that we model the normative system as an autonomous agent, the behavior of the normative system is determined by agents playing a role in it. In this section we also model this aspect of normative systems. We extend the model of artificial social systems with another class of socially constructed agents – called roles – determining the behavior of the normative system. In particular, we replace the set of strategies of the normative system R by a set of set of strategies of the roles in the normative system, written as $R_1, \ldots, R_n$. A strategy of the normative system corresponds to a strategy for each of its roles, which is represented by $R = R_1 \times \ldots \times R_n$. The strategies of the role are its possible behaviors associated with exercising the rights or powers of the role.

Definition 7 *Let $R_1, \ldots, R_n$, S and T be the sets of strategies available to agent 0.1, …, 0.n, 1 and 2, respectively. The set of strategies of agent 0 is determined by the strategies of the roles, $R = R_1 \times \ldots R_n$.*

Moreover, we consider not only a utility function of the normative system, but also utility functions of the roles. These utility functions represent the roles' responsibilities.

Definition 8 *A role-based normative* game *(or a* role-based normative multi-agent encounter*) is a tuple $\langle N, R_1, \ldots, R_n, S, T, U_0, U_{0.1}, \ldots, U_{0.n}, U_1, U_2 \rangle$, where $N = \{0.1, \ldots, 0.n, 1, 2\}$ is a set of agents, R_i, S and T are the sets of strategies available to agents $0.i$ for $1 \leq i \leq n$, 1 and 2 respectively, and $U_0, U_{0.1}, \ldots, U_{0.n}, U_1, U_2 : R \times S \times T \rightarrow I\!R$ are real-valued utility functions for the agents.*

Quasi-stable social laws and enforced social laws are defined in the obvious way. Since the normative system cannot play any strategies, it is not used to determine which social laws are enforced. However, it may be used for the assignment of agents in the following section.

Definition 9 *Given a role-based normative game $\langle N, R_1, \ldots, R_n, S, T, U_0, U_{0.1}, \ldots, U_{0.n}, U_1, U_2 \rangle$, and an efficiency parameter q, a social law (i.e., a restriction of S to $\overline{S} \subseteq S$, and of $\overline{T} \subseteq T$) is* enforced *if there is a unique set of restriction of R_1 to $\overline{R}_1 \subseteq R_1$, …, R_n to $\overline{R}_n \subseteq R_n$ such that $\overline{R}_1, \ldots, \overline{R}_n, \overline{S}, \overline{T}$ is quasi-stable.*

The motivation for introducing multiple roles is that a police agent playing the police role has himself to be controlled by the normative system. Table 1 visualizes the possibility that police roles control each other. This table should be read as follows. Agent 1 can play p or $\neg p$ (columns), and agent 2 can play q or $\neg q$ (rows). The last two numbers of each sequence represent the utilities of agent 1 and 2, and the first table thus represents that agent 1 and 2 play a classical prisoner's dilemma.

$\neg w_1, \neg w_2$	p	$\neg p$
q	*4,5,5*,3,3	*1,5,5*,4,1
$\neg q$	*1,5,5*,1,4	*1,5,5*,2,2

w_1, w_2	p	$\neg p$
q	*3,3,3*,3,3	*0,3,3*,2,1
$\neg q$	*0,3,3*,1,2	*0,3,3*,2,2

	w_1	$\neg w_1$
w_2	*?,3,3*,?,?	*?,0,10*,?,?
$\neg w_2$	*?,10,0*,?,?	*?,5,5*,?,?

Table 1. p, q, w_1, w_2 **is an enforced social law**

Moreover, police role 0.1 can play either working w_1 or not working $\neg w_1$, and police role 0.2 can play either w_2 or $\neg w_2$. The behavior of the normative system is determined by police roles 0.1 and 0.2. The first table represents that the police roles are not controlling the agents, and the second table represents the case in which they do. Both role playing 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 police roles. When one of them works but the other does not, then the working role 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 the exposition manageable, the utilities of the police roles do not depend on the strategies of the two other agents.

The last table is again a prisoner's dilemma. The two police roles would prefer not to work, but the only stable outcome is that they work. The reason is that if they do not work, then they may be punished by the other police role. 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 enforced social law.

It is instructing to consider the case in which agent 1 and 2 play $\neg p$ and $\neg q$, but agent 0.1 and 0.2 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 0.1 and agent 0.2 cannot cooperate.

The example also illustrates a lesson for the construction of the socially constructed roles. Agents playing defender roles like our police roles should not be able to change the normative system to increase their utility. If the police roles would have this power, then they could change their prisoner's dilemma in the same way as the original prisoner's dilemma of agent 1 and 2 has been dealt with. Agent 1 and 2 would suffer from such a transformation, as due to the game dynamics it would result in $\neg w_1, \neg w_2$ and thus $\neg p, \neg q$, and they should have the power to block it.

4. Role assignment

A role assignment relates roles with agents.

Definition 10 *Given a role-based normative game* $\langle N, R_1, \ldots, R_n, S_1, \ldots, S_m, U_0, \ldots, U_{0.n}, U_1, \ldots, U_m \rangle$, *and an efficiency parameter* q, *a role assignment is a (possibly partial) function* $A : \{1 \ldots n\} \rightarrow \{1, \ldots, m\}$. *When* A *is a complete function, we call* A *a complete role assignment.*

When an agent is assigned to a role, it can determine which strategy the role plays. We model the effect of a role assignment as a game transformation, such that an agent determines the strategies of a role. For example, when an agent with strategies S plays a role with strategies R_i, then the new strategies of the agent are $S \times R_i$. Assigning a role to an agent increases the set of strategies of the agent, because it can now exercise the role's rights.

Definition 11 *Given a role-based normative game* $\langle N, R_1, \ldots, R_n, S_1, \ldots, S_m, U_0, \ldots, U_{0.n}, U_1, \ldots, U_m \rangle$, *an efficiency parameter* q, *and a role assignment* $A : \{1 \ldots n\} \rightarrow \{1, \ldots, m\}$. *The effect of the role assignment is a new role-based normative game in which for each agent* a *and role* r *with* $a = A(r)$ *we replace* S_a *by* $S_a \times R_r$ *and we remove* R_r. *The utility functions are updated to account for the new order of the strategies in the obvious way.*

For example, consider an extension of our example with two agents 3 and 4, where the utility for agent 3 and 4 corresponds precisely to the utilities of the police roles. This represents that the two agents are perfect for the job, and the game transformation leads to the same game as before. A way to interpret the existence of ideal agents is that the roles have been designed for these agents.

Moreover, a new class of computational problems can be defined, not studied thus far in the area of artificial social systems, which have to do with finding a good match between the quasi-stable social laws before the role assignment, defined in terms of the ideal roles, and the quasi-stable social laws after the role assignment. We propose the following computational role assignment problem.

Definition 12 *Given a role-based normative game* $\langle N, R_1, \ldots, R_n, S_1, \ldots, S_m, U_0, \ldots, U_{0.n}, U_1, \ldots, U_m \rangle$, *and an efficiency parameter* q. *Find a complete role assignment* $A : \{1 \ldots n\} \rightarrow \{1, \ldots, m\}$ *such that the quasi stable social laws of the role-based normative game are identical to the quasi-stable social laws of the normative game that results from the role assignment.*

If such a role assignment cannot be found, the utility function of the normative system may be a guide to select the best role assignment.

5. Concluding remarks

In this paper we illustrate how game theory can be used to analyze the interaction among role playing agents exercising rights in an organization, without extending the ontology of games. The normative system and its roles are defined as socially constructed agents, and role-based rights are implicitly represented by the roles' strategies exercising the roles' rights. Moreover, we show how game dynamics can be used to control not only the behavior of ordinary agents, but also of defender agents determining the behavior of the normative system. We introduce the role assignment problem, defined as finding a role assignment such that the quasi-stable laws do not change.

A topic for further research is an integration of our rights-as-powers with the formalization of rights advocated by Alonso [1], as well as with other formalizations of rights proposed in deontic logic such as [10]. An alternative formalization of organizations and roles in artificial social systems using constitutive norms is pioneered in [5]. Moreover, we are interested in further extensions of our model, for example in which all powers of the Trias Politica are formalized. E.g., we may further extend the example by introducing another role with the power to change the normative system, and define new computational problems for this extension.

References

[1] E. Alonso. Rights and argumentation in open multi-agent systems. *Artificial Intelligence Review*, 21:3–24, 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*. ACM Press, 2005.

[4] G. Boella and L. van der Torre. The evolution of artificial social systems. In *Procs. of IJCAI'05*, 2005.

[5] G. Boella and L. van der Torre. Organizations in artificial social systems. In *Proceedings of AAMAS-05 workshop From Organizations to Organization Oriented Programming (OOOP05)*, 2005.

[6] J. Searle. *The Construction of Social Reality*. The Free Press, New York, 1995.

[7] Y. Shoham and M. Tennenholtz. On social laws for artificial agent societies: off-line design. *Artificial Intelligence*, 73 (1-2):231 – 252, 1995.

[8] Y. Shoham and M. Tennenholtz. On the emergence of social conventions: modeling, analysis, and simulations. *Artificial Intelligence*, 94 (1-2):139 – 166, 1997.

[9] M. Tennenholtz. On stable social laws and qualitative equilibria. *Artificial Intelligence*, 102 (1):1–20, 2000.

[10] L. van der Torre and Y. Tan. Rights, duties and commitments between agents. In T. Dean, editor, *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI'99)*, pages 1239–1244, 1999.