Designing Organized Multiagent Systems through MDPs

Moser Fagundes, Roberto Centeno, Holger Billhardt, and Sascha Ossowski
{moser.fagundes, roberto.centeno, holger.billhardt, sascha.ossowski}@urjc.es

Centre for Intelligent Information Technologies (CETINIA)
University Rey Juan Carlos
Madrid, Spain

Abstract. In this paper we present an approach to design an Organized Multiagent Systems (OMAS) for teamwork. We use a general formal model for OMAS that employs the notion of organizational mechanisms. The purpose of such mechanisms is influencing the behaviour of the agents towards more effectiveness with regard to some objectives. To achieve our goal we use Markov Decision Processes (MDPs) as a framework to design the organizational mechanisms. In order to illustrate our approach we use the medical emergencies domain where ambulances have to be selected in order to assist and transport patients to the hospitals.

1 Introduction

Ongoing research goals on Multiagent Systems (MAS) include the development of autonomous agents capable of reasoning and acting in open systems. Heterogeneous agents with different designs may join these open systems, and there is no guarantee that they will follow any behavior pattern, neither that they understand other entities. In this context the organizational concept can be of great value to provide support and regulate these MAS, ensuring the well-functioning of the whole system.

The paper [2] claims that the concept of organization is not restricted to the existence of some entity with a global purpose. Organizational structures may also exist (or emerge) as a mean to aid agents in their decision making processes in an uncertain environment. That work proposes a general formal framework for organizations, founded on the idea of *organizational mechanisms* that can be classified as *informative* or *regulative*. Informative mechanisms are well-suited for MAS where the agents have incomplete (and possibly inaccurate) knowledge about the environment, including other agents. These agents have to estimate and evaluate the expected utility of each possible action course (or of the actions it is aware of) to decide what actions to take next. From a micro level perspective any additional information may improve the agent's decision. From a macro perspective the informative mechanisms can be used to influence the agents' behavior without any adjustment on their autonomy. On the other hand the regulative mechanisms adjust the agents' autonomy by changing their capabilities and actions' outcome. This hard enforcement aims to regulate the agents' behavior when particular environmental states take place.

In this paper we explore *organizational mechanisms* for efficient teamwork in an Organized Multiagent System (OMAS). To achieve our goal we design the organizational mechanisms with Markov Decision Processes (MDPs)[1]. Such Markovian models are well-known and have already been applied successfully in multiagent teaming [5]. To illustrate our approach we use the medical emergencies context where people request medical assistance for emergencies. The employment of advances on MAS to improve this type of medical service [4, 3] is an area of significant potential since agent systems provide tools to simulate and evaluate different organizational models before applying changes on real-world systems.

This paper is organized as follows: in Section 2 we define Organized Multiagent Systems and detail the organizational mechanisms to be explored along the paper; Section 3 presents our design for an OMAS for medical emergencies done with MDPs; finally in Section 4 we draw the conclusion and future work.

2 Organized Multiagent Systems

According to [2] an OMAS is a tuple $\langle RA, A, \chi, \phi, x_0, \varphi, OM \rangle$ where: RA is a set of rational agents; A is a possibly infinite action space that includes all possible actions that can be performed in the system; χ is the environmental state space; ϕ:$\chi \times A^{|RA|} \times \chi \rightarrow [0..1]$ is the transition probability distribution which describes how the environment evolves as a result of agents' joint actions; x_0 stands for the initial environmental state of the system; φ:$RA \times \chi \times A \rightarrow \{0, 1\}$ is the agents' capability function describing the actions the agents are able to perform in a given environmental state $\varphi(a,x,ac)$=1 ($\varphi(a,x,ac)$=0) means that an agent is able (not able) to perform action ac in the state x; OM is a non-empty set of organizational mechanisms.

The organizational mechanisms OM can be divided in two kinds: *informative* and *regulative*. An informative organizational mechanism is a function that given a partial description of an internal state of an agent and taking into account the partial view that the mechanism has of the current environmental state, provides information:

$$\Gamma : S' \times \chi' \rightarrow I$$

where:

- S' represents the set of possible partial descriptions of agents' internal states;
- χ' is the set of partial views of environmental states;
- I represents an information space.

The information provided may consist of a set of actions an agent can take but it is possibly not aware of, a recommendation of a particular action, or information about the consequences that a given action may have.

Regulative mechanisms produce modifications in the environment with the aim to improve the system's behavior from a macro level perspective. They can be divided in two types:

- An *incentive* mechanism Υ_{inc} for an OMAS is a function that given a description of the environmental state of OMAS produces changes in the transition probability distribution of OMAS:

$$\Upsilon_{inc}:\chi\text{'}\rightarrow[\chi\times A^{|RA|}\times\chi\rightarrow[0..1]]$$

– A *coercive* mechanism Υ_{coe} for an OMAS is a function that given a description of the environmental state of OMAS produces changes in the agent's capability function of OMAS:

$$\Upsilon_{coe}:\chi\text{'}\rightarrow[RA\times\chi\times A\rightarrow\{0,1\}]$$

3 An OMAS Approach for Medical Emergencies

Different emergency centres may have different ways for handling their assistances, however they have common roles and procedures: centres receive help requests and send ambulances to assist the patients; if the ambulance crew is not able to provide the adequate treatment in situ, then the patient is transported to a hospital.

3.1 OMAS Specification

In our medical emergencies OMAS the rational agents can play one of the following roles: *ambulance*, *patient* or *hospital*. It is assumed that the system has *p* ambulances, *q* patients and *r* hospitals.

$$RA = \{ambulance_1, ..., ambulance_p, patient_1, ..., patient_q,\\ hospital_1, ..., hospital_r\}$$

The action space A consists of: *move* for moving to an adjacent position; $assist_j$ for establishing a commitment where an agent is going to assist the j^{th} patient; $pick_j$ for picking the j^{th} patient; $release_j$ for releasing the j^{th} patient; *wait* for waiting for medical assistance; and *skip* for doing nothing.

$$A = \{move, assist_j, pick_j, release_j, wait, skip\}$$

An environmental state space consists of a set of features χ=$\Xi_1\times...\times\Xi_n$, where each Ξ_i, $1\leq i\leq n$, corresponds to a single feature. Each feature Ξ_i can have a single value to be selected from a vector of possible values. The set of features to be specified in our MAS are decomposed in following subsets:

$$\chi = \chi_{ambulances}\times\chi_{patients}\times\chi_{hospitals}$$

The $\chi_{ambulances}$ is composed by the ambulances' state and position. Their state $\Xi_{as[i]}$ can assume the values *available*, what means the ambulance is not transporting a patient and consequently it is available to be assigned for an assistance, or $patient_j$, what corresponds to $ambulance_i$ transporting $patient_j$. The ambulances' position $\Xi_{ap[i]}$ can assume the value that corresponds to the area the ambulance is located in. Within $\chi_{patients}$ the feature $\Xi_{pp[j]}$ corresponds to the $patient_j$'s position, while the feature $\Xi_{ps[j]}$ stands for $patient_j$'s state. The patient's initial state is *healthy*, but when she gets *sick* she requests help and changes her state to *waiting*. When the ambulance begins the assistance the $patient_j$'s state $\Xi_{ps[j]}$ assumes the value *ambulance*. When the patient is hospitalized her state assumes the value *hospital* until she becomes *healed*. Finally, the $hospital_k$'s location is represented in the feature $\Xi_{hp[k]}$.

$\chi_{ambulances} = \Xi_{as[1]} \times ... \times \Xi_{as[p]} \times \Xi_{ap[1]} \times ... \times \Xi_{ap[p]}$
$\chi_{patients} = \Xi_{ps[1]} \times ... \times \Xi_{ps[q]} \times \Xi_{pp[1]} \times ... \times \Xi_{pp[q]}$
$\chi_{hospitals} = \Xi_{hp[1]} \times ... \times \Xi_{hp[r]}$

$\Xi_{as[i]} = \{available, patient_1, ..., patient_q\}$ where $1 \leq i \leq p$
$\Xi_{ap[i]} = \{area_1, ..., area_n\}$ where $1 \leq i \leq p$
$\Xi_{ps[j]} = \{healthy, sick, waiting, ambulance, hospital, healed\}$ where $1 \leq j \leq q$
$\Xi_{pp[j]} = \{area_1, ..., area_n\}$ where $1 \leq j \leq q$
$\Xi_{hp[k]} = \{area_1, ..., area_n\}$ where $1 \leq k \leq r$

The set of available actions for a rational agent in a particular environmental state is given by a capability function $\varphi: RA \times \chi \times A \rightarrow \{0,1\}$. Table 1 summarizes the set of available actions (column *A*) for the agents (column *RA*) in particular environmental states (column χ). In example, the action *move* is available for ambulances in all states of the world. The action $pick_j$ is available for an ambulances only in environmental states where the ambulance and $patient_j$ occupy the same physical position, the ambulance is available and patient is waiting for medical assistance.

Table 1. Capabilities for agents involved in the medical emergencies domain.

RA	*A*	χ
$ambulance_i$	*move*	ALL
	$assist_j$	$(\Xi_{as[i]}=available) \wedge (\Xi_{ps[j]}=waiting)$
	$pick_j$	$(\Xi_{ap[i]}=\Xi_{pp[j]}) \wedge (\Xi_{as[i]}=patient_j) \wedge (\Xi_{ps[j]}=waiting)$
	$release_j$	$(\Xi_{as[i]}=patient_j) \wedge (\Xi_{ps[j]}=ambulance)$
$hospital_k$	$pick_j$	$(\Xi_{hp[k]}=\Xi_{pp[j]}) \wedge (\Xi_{ps[j]}=sick)$
	$release_j$	$\Xi_{ps[j]}=healed$
$patient_j$	*move*	$\Xi_{ps[j]}=healthy$
	wait	$\Xi_{ps[j]}=sick$
ALL	*skip*	ALL

The function φ defines the agents' capabilities for particular states of the world, but it does not provide information regarding how the environment evolves when an agent executes an action. Such environmental modifications are driven by the transition function $\phi: \chi \times A^{|RA|} \times \chi \rightarrow [0..1]$, which gives the probability distribution over states of the world that can take place through the execution of actions on the current state.

To specify the function ϕ for each $x \in \chi$ can be exhausting for a χ with a high number of features. In order to facilitate the specification of ϕ (qualitative and quantitatively) we use the following rules[1]:

[1] The rules were specified taking into account that the agents are capable of executing the given actions in the current state of the world - in other words, actions were already filtered by the capability function φ (e.g. in the rule (r2) we assume the condition $((\Xi_{as[i]}=available) \wedge (\Xi_{ps[j]}=waiting))$ holds).

(r1) if *ambulance*$_i$ executes *move* then she will move to the adjacent position; if $\Xi_{as[i]}$=*patient*$_j$ and $\Xi_{ps[j]}$=*ambulance* then $\Xi_{pp[j]}$ assumes the same value of $\Xi_{ap[i]}$ (if *ambulance*$_i$ is transporting *patient*$_j$ then their position have to be the same);
(r2) if *ambulance*$_i$ executes *assist*$_j$ then $\Xi_{as[i]}$=*patient*$_j$;
(r3) if n ambulances execute the *pick*$_j$ (pick the same patient) at the same time then the chances of success for each one is 1/n; if *ambulance*$_i$ is successful then $\Xi_{as[i]}$=*patient*$_j$ and $\Xi_{ps[j]}$=*ambulance*, otherwise $\Xi_{as[i]}$=*available*;
(r4) if *ambulance*$_i$ executes *release*$_j$ then $\Xi_{as[i]}$=*available* and $\Xi_{ps[j]}$=*sick*;
(r5) if *hospital*$_k$ executes *pick*$_j$ then $\Xi_{ps[j]}$=*hospital*;
(r6) if *hospital*$_k$ executes *release*$_j$ then $\Xi_{ps[j]}$=*healthy*;
(r7) if *patient*$_j$ executes *move* then she will move to the adjacent position;
(r8) if *patient*$_j$ executes *wait* then $\Xi_{ps[j]}$=*waiting*.

In exception of (r3) the remaining rules are deterministic. However these rules can be easily modified in order to model the uncertainty on the agent's actions. In example in (r1) the *move* action can be adjusted to have a probability p of success and (1-p) of remaining in the same position.

3.2 Organizational Mechanisms

As previously stated organizational mechanisms consist of processes for regulating the behavior of agents. To model such processes we create a set of organizational actions A_{OM} to represent the mechanisms and specify a reward function R_{OM} that motivates the OMAS to behave like a MDP[2]. There is no necessity for extending the transition function ϕ since the organizational mechanisms do not cause changes on the environment features.

In our application the organizational action space A_{OM} contains informative mechanisms (aom1–2) and a regulative mechanism (aom3):

(aom1) inform *ambulance*$_i$ that *patient*$_j$ at *area*$_x$ is waiting for assistance;
(aom2) inform *ambulance*$_i$ that *patient*$_j$ at *area*$_x$ is already attended;
(aom3) change probabilities for ambulance$_i$'s *move* in order to cover an *area*$_x$.

The informative mechanism represented in (aom1) benefits the ambulances placed nearest the patient. The second mechanism (aom2) aims to avoid situations where multiple ambulances go for the same patient. Finally (aom3) intends to keep the city without undercovered areas where no ambulance is available. The reward function R_{OM} defines rewards for executing these organizational actions in particular situations:

(rwom1) R_{OM}(($\Xi_{as[i]}$=*available*) $\wedge$ ($\Xi_{ps[j]}$=*waiting*), *aom1*)=*r1* and *r1* is normalized in the interval [0..1], where *r1*=1 for the nearest ambulance and *r1*=0 for farthest ambulance;

[2] A MDP can be described as a tuple $\langle S,A,T,R\rangle$, where: S is a finite set of states of the world; A is a finite set of actions; $T{:}S\times A\rightarrow\pi(S)$ is a state-transition function, giving for each state and action, a probability distribution over states of the world ($T(s_i,a,s_j)$ for the probability of executing a at s_i, and ending at s_j); $R{:}S\times A\rightarrow\mathbb{R}$ is a reward function that gives the expected immediate reward gained by the agent for taking each action in each state of the world ($R(s_i,a)$ for the reward executing a at s_i).

(rwom2) $R_{OM}((\Xi_{as[i]}=patient_j) \wedge (\Xi_{ps[j]}=waiting), aom2)=r2$ and $r2$=-1 for the nearest ambulance and $r2$=1 for the remaining ambulances on the way to assist $patient_j$;

(rwom3) $R_{OM}((\Xi_{ap[i]} \neq area_x) \wedge (\Xi_{as[i]}=available)$, aom3)=$r3$, where $r3$=1 for the available ambulance that is nearest $area_x$, and $r3$=-1 for the remaining ambulances.

4 Conclusion and Future Work

This paper presents an OMAS design for effective multiagent teamwork in the medical emergency domain. Our approach aims to bridge the gap between the abstract organizational model and the OMAS development by exploring MDPs. In this paper we focused on the organizational mechanisms, leaving the agents' design outside of our scope. We assume that agents are rational and, thus, tend to maximize their rewards.

By employing MDPs to build organizational mechanisms we highlight how they can be thought. Since we are dealing with autonomous agents, there is no guarantee that the usage of such mechanisms will bring the desired outcomes. Through *informative* mechanisms agents can be persuaded by receiving relevant information, but there is a chance that these agents do not take that information into account. By using *incentive* mechanisms it is possible to change actions outcomes, however there is no guarantee the agents will perceive such changes. Additionally, these incentives may not have an influence on the pursuit of the agents' intentions remember agents may not be aware of others' goals. Combining *incentive* with *informative* mechanisms, we assure that agents are aware of the possible outcomes their actions will have and can act accordingly.

Future work on Organized MAS is twofold. The first research direction consists of applying learning techniques for MDPs to design dynamic organizations. The second one aims the investigation of medical emergency scenarios with complex interactions, competition and establishment of commitment between involved parties (e.g. ambulance companies have to compete to assist a patient, and independent of the circumstance at least one ambulance has to be assigned to help that patient).

References

1. R. Bellman. A Markovian Decision Process. *Journal of Mathematics and Mechanics*, 6:679 – 684, 1957.
2. Roberto Centeno, Holger Billhardt, Ramon Hermoso, and Sascha Ossowski. Organising MAS: A Formal Model Based on Organisational Mechanisms. In *24th Annual ACM Symposium on Applied Computing (SAC2009), Hawaii, USA, March 8-12*, page to appear, 2009.
3. Roberto Centeno, Moser Fagundes, Holger Billhardt, and Sascha Ossowski. Supporting Medical Emergencies by MAS. In *Distributed Systems and Artificial Intelligence Applications (KES AMSTA 2009)*, page to appear. Springer-Verlag, 2009.
4. Beatriz López, Bianca Innocenti, and Dídac Busquets. A Multiagent System to Support Ambulance Coordination of Urgent Medical Transportation. *IEEE Intelligent Systems*, 23(5):50 – 57, 2008.
5. Ranjit Nair and Milind Tambe. Hybrid BDI-POMDP Framework for Multiagent Teaming. *Journal of Artificial Intelligence Research (JAIR)*, 23:367–420, 2005.