Influence-Based Autonomy Levels in Agent Decision-Making

Bob van der Vecht[1,2], André P. Meyer[1], Martijn Neef[1], Frank Dignum[2], and John-Jules Ch. Meyer[2]

[1] TNO Defence, Security and Safety, The Hague,
bob.vandervecht, andre.meyer, martijn.neef@tno.nl
[2] Department of Information and Computing Sciences, Universiteit Utrecht, Utrecht
dignum, jj@cs.uu.nl

Abstract. Autonomy is a crucial and powerful feature of agents and it is the subject of much research in the agent field. Controlling the autonomy of agents is a way to coordinate the behavior of groups of agents. Our approach is to look at it as a design problem for agents. We analyze the autonomy of an agent as a gradual property that is related to the degree of intervention of other agents in the decision process. We define different levels of autonomy in terms of inter-agent influences and we present a BDI-based deliberation process in which different levels of autonomy can be implemented.

1 Introduction

This research is motivated by a perspective on automation of distributed systems. As systems are becoming more capable of performing complex tasks, a number of new applications can be thought of where different actors collaborate to reach joint goals. Collaboration can be achieved in several ways and coordination of action always plays an important role. In *mixed-initiative systems* several types of collaboration occur in a dynamic manner. Mixed-initiative means that the initiative for actions of the system comes from multiple actors.

If we look at the engineering of mixed-initiative systems, an agent-based approach seems logical and appropriate. Concepts that are used in the agent community like *situatedness* and *proactiveness* [1] are recurring themes when dealing with the design of such systems. We believe that the concept of *autonomy* of the actors plays a crucial role as well. Actors in mixed-initiative systems need to perform tasks on multiple levels of autonomy during the collaboration.

Mixed-initiative systems can consist of human and artificial actors. Human beings can collaborate with agents, or several types of agents with each other. In this paper we will analyze some engineering issues of agents for mixed-initiative systems and more specifically the concept of autonomy in agents. We propose a decision model with different levels of autonomy. We deal with the problem of autonomy in agent design by introducing inter-agent influence types in the reasoning process. In Sect. 2.1 and 2.2, we describe the function of autonomy in mixed-inititative system, and define the concept of autonomy itself. Section 2.3 addresses related work on agent autonomy. In Sect. 3

we present a deliberation loop for autonomy-aware BDI agents. We show how different levels of autonomy can be defined and how they relate to the agent decision model. Section 4 explains our practical approach and illustrates it with an experiment. We give some conclusions in Sect. 5.

2 Autonomy

The first part of this section describes the function of autonomy of actors in mixed-initiative systems in a general way. In the second part, we focus on autonomy in artificial agents, and we present the definition of autonomy as used in this research.

2.1 Mixed-Initiative Systems

A property of mixed-initiative systems is that the system makes use of different types of collaboration among the actors. We have tried to identify some system requirements of mixed-initiative systems and we try to meet these requirements using an agent-based approach. Figure 1 shows some of the concepts that actors in mixed-initiative systems should be aware of in order to handle different coordination types dynamically. In this paper we focus on the issue of autonomy, although neither autonomy, nor any of the other requirements can be studied in isolation. The various concepts interact and influence each other and add to the complexity of this research topic.

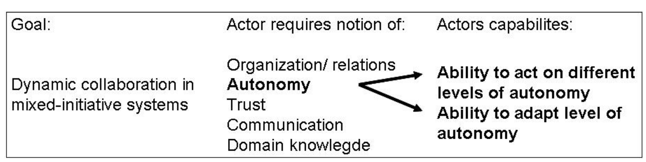


Fig. 1. Important issues concerning dynamic collaboration in mixed-initiative systems

The concept of autonomy is related to the ability of taking initiative. In human cooperation, people readily change their level of dependence with respect to others. This can be regarded as adaptation of their autonomy level. This ability is also essential for agents in mixed-initiative systems, and is currently still lacking. In this paper we will introduce a solution for implementing different levels of autonomy in agents.

2.2 Agent Autonomy

The term autonomy is used in many definitions of agents and is regarded as one of the key features of an agent [1]. Being autonomous means that the agent has control over both its internal state and its behavior. By assuming that agents are autonomous entities, we expect to know something about the way they are internally constructed.

However, the fact that an agent is autonomous does not imply that it has to make all its decisions by itself. In the context of this paper, we look at autonomy as a relational property. We consider the levels of agent autonomy with respect to a certain goal and with respect to other agents. Therefore the degree of autonomy can be defined as the degree of intervention by other agents on the decision making process of how to reach that goal. Barber defines autonomy in a similar manner [2]: *An agent's degree of autonomy, with respect to some goal that it actively uses its capabilities to pursue, is the degree to which the decision-making process, used to determine how that goal should be pursued, is free from intervention by any other agent*.

We have adopted Barber's definition of autonomy for the purpose of this research. It states that an agent can have complete internal autonomy, but deliberately restrict or limit its autonomy in the decision-making process when it is pursuing a certain goal. This means that it allows influences of other agents in its decision-making. In the reasoning process of an agent, the level of autonomy is relevant at every point where the agent actually makes a choice. We consider only agents that are more or less deliberate, since they are aware of their choices. We will present examples of different levels of autonomy in the decision-making process and describe how autonomy can be included in the decision model of agents.

It is widely recognized that the ability of an agent to make decisions autonomously is a strong feature [3]. In some settings it is inevitable to allow autonomous decisions, for example if communication with others fails, or if there is no time to negotiate actions. On the other hand, some tasks require coordination, and then it is necessary to be able to predict an agent's actions, in order to avoid mistakes or conflicts. By limiting its autonomy the agent becomes (partly) dependent of others. At the same time the agent uses capabilities of others. The level of autonomy of agents in the decision-making process with respect to a certain goal is important for cooperation in groups of agents. Controlling and adjusting the agent autonomy can be used in coordination principles.

One perspective on autonomy is that it is an internal feature of an agent, such that it controls its own internal state and its behavior. Another perspective is to look at autonomy as a relational property and considere an agent's autonomy with respect to a certain goal and with respect to other agents. Then autonomy becomes a gradual property of the decicion-making process of the agent. We would like to argue here that a truly internally autonomous agent should be able to reason about its level of autonomy in the decision-making process and should be able to adjust its autonomy with respect to a certain goal and to other agents. In the following sections, we first evaluate the use of agent autonomy in other research, and then propose our reasoning model.

2.3 Related Work on Agent Autonomy

Although agent autonomy has been subject to a lot of research, there is no agreement on one definition. Reason for this could be that autonomy is often seen as a property of agents, but it is possible to look at it from different perspectives. Carabelea et al. [4] have given an overview of those perspectives and have tried to classify them. They call the property of an agent being autonomous *self-autonomy*. They distinguish three main types of autonomy in the relation between an agent and its surronding: *user-autonomy*, *social-autonomy* and *norm-autonomy*. Our approach of agent autonomy is from a relational perspective, and fits in their definition of *social-autonomy*.

Controlling the autonomy of an agent is a way to coordinate the behavior of groups of agents. This coordination can be achieved by explicitly implementing the relations between agents inside their behaviors, e.g. by predefined protocols. We want to argue that this undermines the *self-autonomy* of an agent; the agent does not control its internal state anymore. Agent organizations are an approach to improve coordination without touching the *self-autonomy* as feature of agents. Several researchers propose ways for defining organizational relations, for example by using norms. The Opera model proposes an expressive way for defining organizations in terms of an organizational model, a social model and an interaction model [5]. It uses norms in the description of roles and interaction schemes to define obligations and permissions of an agent. In its approach it explicitly distinguishes between the organizational model and the agents who will act in it. Another example is Moise+, which provides an organizational middleware for agents, which checks whether actions of agents are allowed or not according to the governing organizational rules [6]. Both approaches separate the organizational model from the agent model. This choice is very legitimate for their goal of developing organizational models. We believe that we can make the concept of agent organizations more powerful by designing agents that can handle different levels of autonomy in the decision-making process. The freedom an agent should get in its decision making can be described in organizational rules and norms.

Work on *adjustable autonomy* of agents is done by Scerri [7] and Barber [8]. Their work is motivated by issues on development of human-agent collaborative systems. Scerri's work includes an implementation of a classification task, where humans and agents work together in a dynamic manner. The agent reasons about when the human or when the agent should perform the task. This is a kind of reasoning about autonomy by using *transfer-of-control* strategies. However, the strategies that Scerri uses, are specifically developed for this classification domain. The system shows mixed-initiative behavior, but only in this domain. The general concept of autonomy is not included in the reasoning process of the agents. In the approach of Barber [8] different levels of autonomy are related to styles of decision making. The focus of their work is on decision strategies for organizations and on the interaction that comes from the choices of autonomy levels. Their aim is to make the agents select the best organizational structure autonomously as a group. Our view on autonomy comes close to theirs, but, in contrast, we use the notion of autonomy for the design of a decision model for single agents and allow them to reason about it individually.

3 Decision Making on Different Levels of Autonomy

In this section we will introduce the concept of autonomy in the reasoning model of an agent. First we will briefly explain the deliberation process of the agent. Then we analyze how we can distinguish different levels of autonomy in the decision-making process and we will integrate autonomy in the agent deliberation.

3.1 Agent Deliberation: Introducing the OODA-Loop

In the agent deliberation process we distinguish four sub-processes: 1. do observations and receive messages, 2. process the observations and messages, and determine their semantics 3. decide on the next action and 4. perform the selected action. These processes can be recognized in the four phases of agent deliberation: Observe, Orient, Decide and Act (OODA). We borrowed these terms from the OODA-loop as it is used in decision-making processes in the military command and control domain [9]. In its generalized form, the OODA-loop can be seen as a cycle for all sorts of decision-making processes. It is comparable to Perceive-Reason-Act cycles that are used in the agent reasoning domain. The reasoning phase has been split into two phases, one for information processing and the other for deciding on actions.

The four phases in the OODA-loop can be implemented in a sequential or in a parallel way (cf. Fig. 2). When parallel they can be seen as separate processes sharing resources, but each with its own frequency.

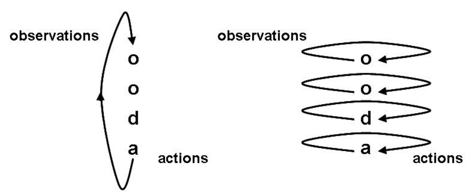


Fig. 2. Sequential and parallel OODA-processes

The Orient phase prepares a world model for the agent to reason with. It transfers raw observation data to data structures that are used in the Decide phase. The designer has some freedom in defining how much information processing takes place in the Orient phase. For example, if certain higher-level information processing is seen as an optional action, the choice for higher-level information processing should deliberately be made in the Decide phase. In the Decide phase, the agent reasons about the actions to take.

We believe that reasoning mechanisms using cognitive notions like beliefs, desires and intentions (BDI) are a good approach for agent reasoning. Several BDI reasoning models have been proposed. For example, 3APL [10], [11], provides the designer with a formalized programming language which is designed for BDI-agent programming. In order to reach its goals, the agent reasons about its beliefs and plans. We use 3APL reasoning in the Decide phase of our deliberation model.

Since we use a BDI-reasoning model in the Decide phase, we will translate both the observations and the content of the messages to beliefs and goals in the Orient phase. These are the data with which the agent can reason properly. Summarizing the four OODA-phases as we use them:

- **Observe** In the Observe phase, the agent observes the environment with its sensors and receives incoming messages. Observations and messages contain the *presented world state*, i.e. the world as presented by the sensors. There is no connection yet with the agent's beliefs.
- **Orient** In the Orient phase, the observations and messages are processed. The beliefs of the agent are updated with those observations and messages. Result of the orient phase is an *interpreted world state*; the world as the agent interprets it. The belief base may include knowledge that the agent derived and that is not observable by the sensors. In the Orient phase, the Decide phase will be prepared; the interpreted world state is the world state with which the agent will continue its reasoning. All concepts that are necessary for the BDI-reasoning process need to be defined. Figure 3(a) shows that the belief base and the goal base of the agent are updated.
- **Decide** The Decide phase is the actual reasoning phase as proposed by 3APL and other BDI programming-languages. The agent reasons with its beliefs and goals and decides upon the next action. Figure 3(b) shows the constructs that are used; Beliefs, Goals, Basic Actions and Practical Reasoning Rules [11].
- **Act** In this phase the action that has been selected will be executed. Actions can be internal actions of the agent (i.e. the capabilities in the 3APL program) or external actions that take place in the environment.

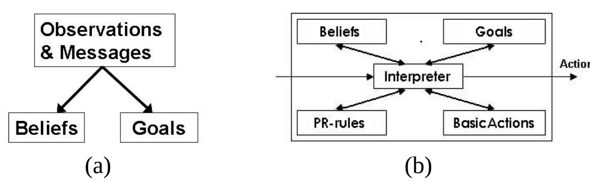


Fig. 3. Two phases of the OODA-loop: a) the Orient phase, and b) the Decide phase

We use the OODA loop to illustrate how different kinds of influences take part in the agent deliberation. In our definition, the agent's level of autonomy is related to the degree of influence of other agents in the reasoning process of the agent. In the next section we will define different levels of autonomy by relating inter-agent influences to the different phases of the reasoning process.

3.2 Inter-Agent Influences

There are three types of influence between agents to be distinguished, [3], [8]: influence by environmental modification, influence by belief alteration and influence on goal/task determination.

All three types of influence can come together and all types of influence can change the decision on the actions selected. We will explain the types of influence and show

how they can be integrated in the decision-making process of an agent. The OODA-cycle as we previously introduced is used to demonstrate this.

Influence by Environmental Modification Influence by environmental modification is achieved by modifying the agents' environment. It influences the agent via its observations. This type of influence affects the Observe phase of the agent deliberation. What is done with those observations and how they are processed is up to the agent itself.

Influence by Belief Alteration Influence on beliefs between agents occurs when one agents informs another by sending a message. Belief influence implies that the agent receives a message and processes the information, i.e. integrates the content with its beliefs. The contents of a message can contain knowledge about the environment or an opinion of the best action to take. Belief influence is based on communication, and therefore it implies that the agents have a shared ontology about the concepts they communicate about. Belief influence reaches upto and including the Orient phase of the OODA-loop. The decision on actions is completely up to the agent itself.

Influence on Goal/Task Determination Influence on goal/task determination occurs when an agent determines the tasks or goals for another agent. The selection process of goals and tasks takes place in the Decide phase of the OODA-cycle and therefore this type of influence reaches to this phase. If agent A has influence on the goal/task determination of agent B it implies that agent B considers suggestions for next actions proposed by agent A, or even stronger, agent B just follows commands of agent A without any doubt. In order to receive a command from another agent or to get someone's opinion about the best next action, agents need to communicate about *goals* and *plans* and need to be able to send *commands* and *opinions* to each other. Therefore a certain level of belief influence between the agents is required.

Barber [2] has identified a spectrum of decision-making styles on goals and tasks as shown in Fig. 4. The spectrum ranges from completely autonomous to completely command driven. In between there are several types of joint decision making, with *true consensus* as the ultimate form of cooperation.

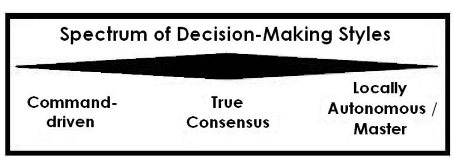


Fig. 4. Spectrum of Decision Making Styles by Barber,[2]

With completely autonomous decision making, there is no influence on goal/task determination. The agent will determine its goals and its actions all by itself, no possible solutions coming from other agents will be considered in the decision-making process. In a fully command-driven decision-making style, there is a full influence on goal/task determination. The agent is dependent on the commands of its partner in the hierarchical relation. With joint decision making styles different types of influences of other agents are possible, for example, an agent can collect opinions of other agents and use them for its decision making.

3.3 Influence-Based Levels of Autonomy

Table 1 summarizes how the influences are linked to the phases in the OODA cycle. We can implement the OODA cycle in such a way that only the required types of influences from other agents take effect.

Table 1. The inter-agent influences in the deliberation process

Type of influence	OODA-phase	Corresp. function
Environmental Modification	Observe	observe()
Belief Influence	Observe Orient	receive_messages() process_messages()
Goal- Task determination	Observe Orient Decide	receive_messages() process_message() commit to commands

We will relate inter-agent influences to levels of autonomous decision making. We can design reasoning profiles for agents in an agent organization in terms of the influence types. Some basic examples of those profiles:

- *Solipsistic*: An agent with a solipsistic personality does not care about other agents. Messages from others are ignored. Also the goal/task determination is free from influences. The agent creates and selects its own goals and plans. Influence via environmental modification is still possible. By manipulating an agent's environment, it is possible to influence the agent's behavior. In a solipsistic agent the direct influence of other agents affects the Observe phase of the reasoning process.
- *Trusting* or *naive*: A trusting or naive agent will process messages from others and belief the content. It is under belief influence of others. The decision of which action to perform next is made by the agent itself. The influence of other agents reaches to the Orient phase of the reasoning process.
- *Obedient*: If Agent A is obedient with respect to agent B, it will do what agent B says without considering other opinions. Its tasks and goals are determined by agent B. Agent B can send a command message to agent A. This message is processed in the Orient phase, and in the Decide phase agent A commits to the command. In the Orient phase, there is influence on beliefs of agent A: it now believes it received a command from agent B. And in the Decide phase there is influence on goal/task determination. This can be done via pre-defined plans in the deliberation

cycle of agent A, that demand that if there is an Obedience relation with agent B and a command from agent B then the agent has to commit to the command. In hierarchical decision-making the direct influence of other agents reaches to the Decide phase.

These are a few examples that use extreme forms of some inter-agent influences. Of course more complex profiles are possible as well. All the reasoning profiles can be implemented using our model of influenced-based autonomy levels. In some sense the profiles provide a basic interaction protocol embedded internally in the agent's deliberation process.

3.4 Towards Adjustable Autonomy

Mixed-initiative systems have the property of using different types of collabotation among the actors. The ability of switching dynamically between those types of collaboration is still lacking in agents. We have introduced levels of autonomy in agent reasoning and a next step would be to make the agent reason about its autonomy and allow it to adjust its autonomy level. An agent can not fully control its own level of autonomy. An agent's autonomy is *bilateral adjustable* [3], which means that the level of autonomy can be adjusted by the agent itself as well as by other agents. For example, an agent asking for help instructions chooses to consider options generated by others and therefore it becomes dependent of other agents and lowers its degree of autonomy while pursuing its goal. In a hierarchical relation the master can tell the assistent to solve a problem on its own. Then the assistent becomes autonomous in solving the problem.

We want to look at agents reasoning about their own autonomy. Question is where in the agent-reasoning model the decision about the desired autonomy level should be made. Dastani et al. [12] have analyzed autonomy in the deliberation of BDI agents. Several choices on the deliberation level of an agent influence the agent's autonomy in its decision for new goals and tasks. They propose a meta-language for agent deliberation, which allows the construction of different deliberation cycles. Switching between autonomy levels then could be done in the deliberation cycle itself.

In order to make agents adjust their own autonomy, we need to find rules for switching between autonomy levels. In the next section we will describe experiments that show us some properties of agents collaborating at certain autonomy levels. On basis of the performances in different situations we want to find rules about which autonomy level would be desired in which situation.

4 Experiment

In this section we introduce an experiment for illustrating the concept of autonomy levels in agent decision-making. We have defined an agent organization, in which the agents can operate on different levels of autonomy. We want to observe properties of the organization in the different compositions and compare the performances.

4.1 Organizational Description

The general setting is a fire brigade organization. There is a world with fires, firefighters and a coordinator. The aim of the organization is to extinguish the fires as fast as possible. In the agent organization two roles have to be fulfilled: Coordinator and Firefighter.

- *Coordinator*: Plan which fire is to be extinguished by which fireman, and send commands to the firefighters.
- *firefighter*: Move around randomly to look for a fire, select a fire and extinguish the selected fire.

The agents playing the roles have been implemented following the decision model described above. The phases of the OODA-loop have been implemented sequentially and the 3APL-reasoning mechanism is used in the Decide phase. Goals, beliefs, plans, and basic actions have been made explicit.

In our simulation, firefighters are situated in an environment, where fires can pop up. A firefighter can move to the fire and extinguish it. Fires are growing gradually in time, except for when they are being extinguished, then the fire size decreases and they will disappear. The firefighters have a limited view. The coordinator agent has a global view, it can see all fires. The only action the coordinator can take is sending commands to the firefighters, telling them which fire they should extinguish. The coordinator has one handicap, which is that he can send only one message per time interval.

We have equipped the firefighters with three different profiles: *solipsistic*, *trusting* and *obedient*. All required influence types are represented in a single OODA-loop. We can create a profile by activating or de-activating the functions corresponding to the influence types as shown in Table 1. As variable we consider only influence on beliefs and on goal/task determination. In all profiles influence via environmental modification is possible, so the `observe()`-function is always active.

- *Solipsistic*: Solipsistic firefighters observe the fires and select the fire they want to extinguish all by themselves. Influence by environmental modification is possible, for example when a firefighter observes that a fire is getting smaller, because another firefighter is extinguishing it. The agents do not process any message from each other or from the coordinator agent.
- *Trusting*: Trusting firefighters are communicating with the other agents. If they see a fire while they are busy extinguishing another fire, they send a message to the other agents to inform them that there is a fire that needs to be extinguished. The receiving agent processes the message. The information of the particular fire is added to its beliefs. There is still influence by environmental modification by other agents extinguishing fires. There is belief influence by sending messages to each other informing them about fires. The `receive_message()`-function is active, as well as the `process_message()`-function.
- *Obedient*: Obedient firefighters are commanded by another agent, who tells them which fire they have to extinguish. They do not take initiative by themselves. There is still influence by environmental modification by other agents extinguishing fires. There is belief-influence by the other agent commands and possibly informing about unknown fires. Therefore both functions `receive_message()` and the

process_message()-function are both active. There is influence on goal/task determination by following the 3APL plans constructed for the obedient relation. A 3APL rule that makes a firefighter agent to follow a command to extinguish a certain fire, could be:

```
<- (obedient(Boss) AND
     command(Boss, fightFire, FireX))
  | fightFire(FireX)
```

It can be read such that the agent no matter what goal it has, if it has an obedient relation with *Boss* and has received a command from *Boss* to extinguish *FireX*, it will adopt the plan *fightFire(FireX)*. *Boss* and *FireX* are variables. The command *fightFire* is written as a constant. This assures that the agent only adopts commands that it knows.

In our experiment we used an organization consisting of two firefighters and one coordinator. We have varied the organizational composition by varying the reasoning profiles of the firefighters with respect to the other agents. We have created three different organizations, as shown in Table 2. In the *solipsistic* organization the firefighter agents ignore all other agents, their profile with respect all other agents is *solipsistic*. In the *trusting* organization the firefighters trust each other, but ignore the coordinator agent. In the *obedient* organization the firefighters are obedient to the coordinator and ignore each other. Note that the difference between the organizations has been created purely by constructing different levels of autonomy for a firefighter agent with respect to other agents in the organization.

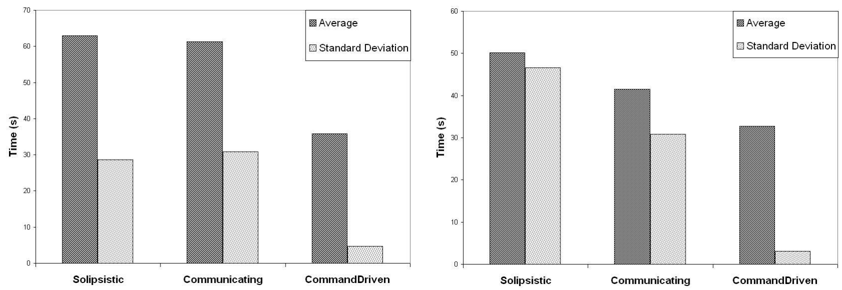
Table 2. Reasoning profiles of firefighters in their relation with the other agents. Three different organizations.

Organizational characteristic	**Profile firefighter regarding firefighters**	**Profile firefighter regarding coordinator**
Solipsistic	Solipsistic	Solipsistic
Trusting	Trusting	Solipsistic
Obedient	Solipsistic	Obedient

Furthermore we used environments with different characteristics. The firefighters started in a world containing five fires. In on situation the fires were spread randomly over the field. In the second situation the five fires were clustered in a group. The performance of the organizations was evaluated by measuring the time it took to extinguish all the fires.

4.2 Results and Discussion

What we wanted to show is that the organizations perform differently in environments with other characteristics. We have run several simulations of the different implementations of the firefighter organization. In our results we show the results of 100 runs in six situations (three organizations in two environments). Figures 5(b) and 5(a) show the average extinguish times and the corresponding standard deviation.



(a) (b)

Fig. 5. Extinguish times over 100 runs; average and standard deviation. a) Random fire distribution and b) Clustered fire distribution

Comparing the results of the clustered and random environments, we see the biggest difference in performance for the *trusting* firefighters. They perform worse in the situation of randomly spread fires. This can be explained by the fact that they will only view the fires one by one. They are not able to take advantage of their communication, since tn their behaviour it was specified that they would inform each other in the situation when they saw more then one fire. They have to search for each fire individually, which explains the growing standard deviation. As a result they perform comparable with the *solipsistic* firefighters.

If we consider the only clustered environment, the results in Fig. 5(b) show a difference in performance of all three organizations. The *obedient* firefighters perform best. They get orders from the coordinator agent, which has a global view, so all fires are known from the start and the coordinator just sends the firefighters to the right places. In the *solipsistic* and *trusting* organizations the firefighters do not process messages from the coordinator agent. They have only a local view, so they first need to look for the fires. When one firefighter has found a fire, he will see the other fires as well in the clustered situation. The *trusting* firefighters exploit this knowledge by telling each other about the fires, so the second firefighter immediately joins to help. *Solipsistic* firefighters do not have this ability of information sharing. As an illustrative example we show the results of a typical run in the clustered situation in Fig. 6. On the x-axis the time is given and on the y-axis the total fire size in the environment for all three organizational types. It is visible that once the *trusting* firefighters have found the first fire, the fire size decreases faster than for the others. In the *obedient* organization the fire size decreases most constantly.

In the above presented simulations, the *obedient* organization outperformed the other two on average. The organization uses the global view of the coordinator agent. The firefighters follow the orders of the coordinator and it does not really matter how

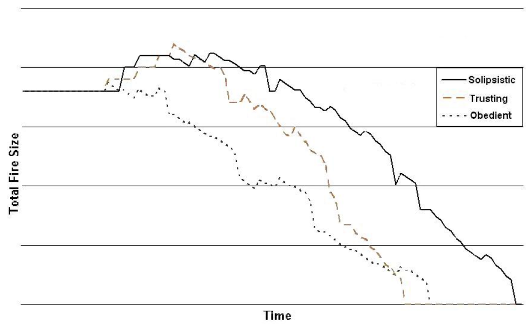


Fig. 6. Total fire size over time, fires are clustered.

the fires are distributed. However, this type of organization has limitations as well. The organization works by a centralized approach and is very sensitive to the performance of the coordinator agent. The restriction on the number of messages that the coordinator agent can send, is not a big issue in the small organization we use here, but it will be in larger organizations. Another problem is that the observations of the coordinator agent play a very important role, since he determines which fire is to be extinguished by which firefighter. Failure of the observations of the coordinator has big consequences. We have run the same test with randomly distributed fires and the *obedient* organization, but with the restriction that the coordinator could only see two third of the field. The average extinguish times of succesful run were comparable of the results in Fig. 5(a), but of the 100 runs the organization failed in 87 cases, because the coordinator missed at least one of the fires. This test shows the necessity of at least some autonomy of the actors in a distributed system.

All results of our experiment are explained by analyzing the information flow in the agent organization. We want to point out that the goal of this experiment was mainly a proof of concept. We have defined different organizations by making the agents reason on different levels of autonomy with respect to the other agents. In our experiment the organization was still static. By defining the autonomy levels internally in the agents in one reasoning model we also create possibilities for switching between autonomy levels at runtime, and therewith allow dynamic organizations. We feel that we need some more experiments to define rules for the agents for switching between autonomy levels.

5 Conclusion and Future Work

This research is motivated by the belief that mixed-initiative and adjustable autonomy are important aspects of future distributed systems and require specific attention. Autonomy of actors is one of the key features when we talk about *mixed-initiative*. The actors should be able to handle several levels of autonomy with respect to a certain goal and with respect to others. We believe that an agent-based approach is a promising way for developing such systems. We have analyzed the concept of autonomy in the decision-making process of agents. Our aim here is to add the notion of autonomy to the reasoning model of an agent.

The level of autonomy is related to the degree of intervention of other agents in the decision making process. We have proposed four phases in the agent deliberation: *Observe*, *Orient*, *Decide* and *Act* and we have described three types of influence between agents: environmental influence, influence on beliefs and influence on goal or task determination. We have linked the different influence types to the first three phases of the agent deliberation. Autonomy levels have been defined in terms of inter-agent influences and we have shown how they can be implemented in the agent's reasoning. In our experiment using a fire-brigade organization, we have created three reasoning profiles for the agents based on the autonomy levels. The three organizational types have performed differently in environments with other characteristics.

As future work, we want to formalize the concepts we used in our reasoning model. We want to construct a mechanism to allow agents to reason about their autonomy. By extending our experiments we want to find rules for the agent to decide on switching between automomy levels. Furhtermore, we are interested in including human interaction in our simulation environment in order to conduct experiments concerning adjustable autonomy in the human-agent interaction domain.

Acknowledgement. The research reported here is part of the Interactive Collaborative Information Systems (ICIS) project, supported by the Dutch Ministry of Economic Affairs, grant nr: BSIK03024.

References

1. Jennings, N.R.: On agent-based software engineering. Artif. Intell. **117**(2) (2000) 277–296
2. Barber, K.S., Martin, C.E.: Agent autonomy: Specification, measurement, and dynamic adjustment. In: Autonomy Control Software Workshop, Autonomous Agents 99. (1999) 8–15
3. Falcone, R., Castelfranchi, C.: The human in the loop of a delegated agent: the theory of adjustable social autonomy. IEEE Transactions on Systems, Man, and Cybernetics, Part A **31**(5) (2001) 406–418
4. Carabelea, C., Boissier, O., Florea, A.: Autonomy in multi-agent systems: A classification attempt. In: Agents and Computational Autonomy. (2003) 103–113
5. Dignum, V.: A Model for Organizational Interaction: based on Agents,founded in Logic. Utrecht University, PhD Thesis (2004)
6. Hübner, J.F., Sichman, J.S., Boissier, O.: A model for the structural, functional, and deontic specification of organizations in multiagent systems. In: SBIA. (2002) 118–128
7. Scerri, P., Sycara, K., Tambe, M.: Adjustable autonomy in the context of coordination. In: AIAA 3rd
8. Barber, K.S., Han, D.C., Liu, T.H.: Strategy selection-based meta-level reasoning for multi-agent problem-solving. In: AOSE. (2000) 269–283
9. Boyd, J.: A discourse on winning and losing. Maxwell Air Force Base, AL: Air University Library Document No. M-U 43947 (unpublished collection of briefing slides and essays) (1987)
10. Hindriks, K.V., de Boer, F.S., van der Hoek, W., Meyer, J.J.C.: Agent programming in 3apl. Autonomous Agents and Multi-Agent Systems **2**(4) (1999) 357–401

11. Dastani, M., van Riemsdijk, B., Dignum, F., Meyer, J.J.C.: A programming language for cognitive agents goal directed 3apl. In: PROMAS. (2003) 111–130
12. Dastani, M., Dignum, F., Meyer, J.J.C.: Autonomy and agent deliberation. In: Agents and Computational Autonomy. (2003) 114–127