Theories about architecture and performance of multi-agent systems

Gazendam, Henk W.M.; Jorna, René J.

Theories about architecture and performance of multi-agent systems

Henk W.M. Gazendam and René J. Jorna[1][2][3]

SOM theme A: Multi-level Interactions Within Firms

Abstract

Multi-agent systems are promising as models of organization because they are based on the idea that most work in human organizations is done based on intelligence, communication, cooperation, and massive parallel processing. They offer an alternative for system theories of organization, which are rather abstract of nature and do not pay attention to the agent level. In contrast, classical organization theories offer a rather rich source of inspiration for developing multi-agent models because of their focus on the agent level. This paper studies the plausibility of theoretical choices in the construction of multi-agent systems. Multi-agent systems have to be plausible from a philosophical, psychological, and organizational point of view. For each of these points of view, alternative theories exist. *Philosophically*, the organization can be seen from the viewpoints of realism and constructivism. *Psychologically*, several agent types can be distinguished. A main problem in the construction of psychologically plausible computer agents is the integration of response function systems with representational systems. *Organizationally*, we study aspects of the architecture of multi-agent systems, namely topology, system function decomposition, coordination and synchronization of agent processes, and distribution of knowledge and language characteristics among agents. For each of these aspects, several theoretical perspectives exist.

---

[1] Prof.dr. Henk W.M. Gazendam is professor of Information Systems in the Public Sector at the Faculty of Public Administration at Twente University and associate professor of Information Strategy at the Faculty of Management and Organization at Groningen University (P.O.Box 800, NL-9700-AV Groningen, The Netherlands, tel +31-50-3637078, email h.w.m.gazendam@bdk.rug.nl); Dr. René J. Jorna is associate professor of Decision Support Systems at the Faculty of Management and Organization at Groningen University (tel +31-50-3633864, email r.j.j.m.jorna@bdk.rug.nl).

[2] This article has been presented at the III European Congress of Psychology, Tampere, Finland, July 4-9, 1993. This research report is a slightly revised edition of October 18, 1996.

[3] The authors thank Gerard Gaalman for his valuable comments.

1. Introduction

Multi-agent systems are promising as models of organization because they are based on the idea that most work in human organizations is done based on intelligence, communication, cooperation, and massive parallel processing. They offer an alternative for system theories of organization, which are rather abstract of nature and do not pay attention to the agent level. In contrast, classical organization theories offer a rather rich source of inspiration for developing multi-agent models because of their focus on the agent level (Paragraph 2).

This paper discusses the question "*which theoretical choices are the most plausible in the construction of multi-agent systems from a philosophical, psychological, and organizational point of view?*". In this discussion, three concepts can be seen as central: the concept of agent, the concept of architecture, and the concept of (emergent) organizational characteristics, or, for short, organizational performance. It turns out that, in defining these concepts, we touch upon all sorts of problems and approaches that are described in philosophy, in cognitive science or artificial intelligence and in organization and management theory. In analyzing these problems and approaches, it turns out that the philosophical point of view one takes is the most fundamental one. Take, for example, the concept of architecture. What is an architecture? Is there a theory about the emergence and evolution of organizational architecture? Modern organization and management theory says that the answer to this question depends on the point of view taken. According to the machine metaphor of organization, architecture is the way the organization is built out of its basic components: agents, material objects, work constellations, communication paths, and their spatio-temporal ordering. The machine metaphor approach can be seen as being based on the philosophical point of view of ontological, objective realism. According to the organism metaphor of organizations, the architecture of an organization is the way the organization is composed of suborganizations fulfilling a specific function; these suborganizations can be decomposed into smaller suborganizations, and so on. In many cases, the organism metaphor uses the language of general systems theory. This variant of the organism metaphor approach can be seen as being based on the philosophical point of view of Platonic realism, especially if one considers the reality assigned to abstract mathematical system concepts. According to the ecology metaphor of organization, architecture is the natural and agent-made environment in which agents are wandering around in order to fulfill their needs. The ecology metaphor approach can be subsumed under objective realism favoured by a kind of natural constructivism. According to the mind metaphor of organizations, architecture is the collection of symbol structures, including the rules that govern the use and production of those symbol structures, that is present in the organization. Culture, language and knowledge are connected to the use and production of symbol structures (or signs, in terms of Peirce). The mind metaphor approach can be seen as being based on the philosophical pragmatism or semiotic constructivism.

Let us return to our question about the theoretical choices that have to be done when constructing multi-agent systems. Choices are necessary at a philosophical level, at a psychological level, and at an organizational or architecture level. For each of these choices, alternative theories exist. *Philosophically*, the organization can be seen from the viewpoints of realism and constructivism. The concept of *architecture* is necessary to understand organization as a principle of arrangement. However, an organization also can be viewed as an entity. The status of this entity depends on the philosophical point of view taken (paragraph 3). *Psychologically*, several sorts of *agents* can be distinguished, based on the concepts of response function and representation. A main problem in the construction of psychologically plausible computer agents is the integration of response function systems with representational systems (paragraph 4). *Organizationally*, we study aspects of the architecture of multi-agent systems, namely topology, system function decomposition, coordination and synchronization of agent processes, and distribution of knowledge and language characteristics among agents. For each of these aspects, several theoretical perspectives exist (paragraph 5). A research program that aims at testing organization theories based on the construction of multi-agent systems is explained (paragraph 6). In paragraph 7, we draw some conclusions.

2. Multi-agent systems as a new perspective in organization theory

In this paragraph, we review several organizational theories with respect to assumptions on agents, architectures, and the resulting performance.

2.1. Classical organization theories

Classical organization theories are, surprisingly, relatively rich because of their focus at the agent level. They are related to the machine metaphor and the topology of objects and agents architecture. The classical organization theories of Taylor, Fayol, and Weber, subsumed by Morgan (1986) under the machine metaphor, see an organization as a whole consisting of agents performing tasks. There is a fixed structure of agent tasks and agent communication. Virtually no attention is paid to symbol structures. Fayol's (1916) management principles can be applied in multi-agent systems. Interesting principles concern specialization related to learning and the communication speed resulting from communication topologies.

2.2. System theories of organization

System theories of organization are related to the organism metaphor and the system function architecture. The organistic organization theories, for instance contingency theory (Burns and Stalker, 1961: Woodward, 1965: Lawrence and Lorsch, 1967), adopt a system-theoretical approach to organizations. In this approach, the organization as a whole is seen as the basic object. This object can be decomposed in subsystems, each of which has a function in the system as a whole. These subsystems can be decomposed further, and so on. The way of thinking of the system theoretical approach is top-down, and opposes the bottom-up way classical organization theories see organizations.

System-theoretical organization theories often do not descend to the concrete level of human agents. Therefore, they lack the richness stemming from a psychologically plausible model of human behavior. Systems theory sees an organization as a black box. The behavior of the system as a whole is visible on the outside of the black box. The behavior of the human agents is invisible in the darkness of the inside of the black box. A consequence of this black box nature of systems theory is that the operationalization of theoretical concepts leading to measurable indicators is a rather difficult step. Because of this rather abstract character of system theoretical organization theories, their content is relatively poor, as has been shown by an analysis of their conceptual framework (Gazendam, 1993). The system-theoretical approach to organizations, as far as it stays at the level of the system as a whole, tends to be an unfruitful Lakatosian research program (Kieser and Kubicek, 1983).

The strong point of systems theory is that it offers a set of abstract principles that stimulate your fantasy in interpreting organizational situations. The most important of these principles is the Law of Requisite Variety. *Ashby's Law of Requisite Variety* states that the variety of control measures must match the variety of disturbances. "Only variety in R [the regulator's actions] can force down the variety due to D [the disturbant's actions]; only variety can destroy variety." (Ashby, 1956/1970: 110)4. In other words, only variety can control variety.

---

[4] In another formulation of his Law of Requisite Variety, Ashby states that the capacity of the channels of communication to be used for perceiving the disturbances and for transmitting the control measures must match the capacity of the disturbance generator. "R's capacity as a regulator cannot exceed R's capacity as a channel of communication." (Ashby, 1956/1979: 115). This means that the capacities of the channels of communication used limit the variety of control measures that effectively can be applied.

The use of Ashby's Law of Requisite Variety may be criticized because of its mechanistic and oversimplified vision on organizations as systems. The use of a more generalized form of Ashby's law takes the edge off this possible criticism (Gazendam, 1993). In a more generalized form, Ashby's law may be seen as governing the response function of an organization on changing (internal or external) circumstances. This interpretation is consistent with (1) Ashby's original game-theoretical notion of the role of variety, and (2) the system-theoretical principle that organizations can be investigated at various levels of abstraction with various time scales.

Ashby's law is an example of a rather abstract system-theoretical theory that can be interpreted in many ways. By a suitable interpretation, it is also applicable to multi-agent systems and can predict a better performance of a multi-agent system than a single-agent system because the variety of agents and agent processes in a multi-agent system generally will lead to a more optimal variety. In a similar way, it is possible to use several general system-theoretical principles or theories for predicting that multi-agent systems are better than single-agent systems. Examples of such principles or theories are:

- stability (stress localization; Simon's principle of near decomposability)
- variety (Ashby's law of requisite variety);
- information processing capacity (Galbraith's analysis; De Leeuw's law of planning and control effectiveness);
- complexity reduction (Simon's principle of satisficing and bounded rationality; Pylyshyn's complexity profiles of architectures);
- learning based on response functions (Cohen's law of specialization and cooperation);
- growth (optimal organization module size; Kastelein and BSO's cellular design); inheritance and evolution as related principles.

In models of computational ecologies (Huberman and Hogg, 1988; Kephart, Hogg, and Huberman, 1989), entropy and chaos emerge as characteristics at the system level.

This use of these abstract system-theoretical notions is not without danger. The operationalization of the abstract concepts used is a difficult path that requires additional, secondary, theories. One is never sure whether a slightly different operationalization of the abstract principle exists that predicts totally different effects.

2.3. Knowledge-based multi-agent theories

An *agent* is an autonomous and intelligent being. Examples are a human being, or a simulator of a human being in the form of a more or less autonomous and intelligent entity realized by software running on a computer system. The latter agent type will be called a computer agent. A *multi-agent* system is a system consisting of agents that communicate and cooperate. Human agents as well as computer agents can participate in multi-agent systems. Multi-agent systems that consist purely of human agents have been studied by organization and management theorists during at least a century. The autonomy and intelligence aspects of the agent concept, and the resulting questions about communication and coordination in multi-agent systems, however, form a new perspective in organization and management theory. Multi-agent systems that consist purely of computer agents are mainly used as simulation models of human organizations, but can be used to perform certain tasks also. Multi-agent systems that consist of human agents as well as computer agents have been studied as environments for computer supported cooperative work, as expert systems, and as decision support systems (Gazendam, Jorna, and Blochowiak, 1991).

Multi-agent systems in the role of simulation models of human organizations offer an intriguing perspective for studying those organizations. Instead of hierarchically structured computer programs, multi-agent simulation models of organization offer a bottom-up perspective: based on the capabilities of agents and the resulting communication and cooperation, emerging characteristics of the organization as a whole can be defined, discovered, or explained. Because most work that is done in human organizations is based on intelligence, communication, cooperation, and massive parallel processing, multi-agent simulation models seem to offer a promising, plausible model of organization. Moreover, the local intelligence and parallel processing features of multi-agent computer systems offer a promise of overcoming present bottlenecks in computerized problem solving.

Multi-agent theory combines the instruments of system theory with the attention for the agent level of the classical theories. Multi-agent organization theory tries to explain the behavior at the systems level based on insight into the agent level of organizations. This leads to a picture of organizations that has advantages over the systems theoretical picture at the following points:

1. a psychologically plausible account of human behavior can be given;
2. the operationalization of theoretical concepts into measurable indicators is less problematic;
3. the black box of the organization as a system is opened, and organizational behavior can be explained as based on the (aggregated) behavior of agents;
4. the agent concept and the idea of agent interaction leading to emergent organizational behavior offer a good starting point for computer simulation models.

These advantages of multi-agent theories over systems theory have as a consequence that multi-agent theory has more intriguing theoretical perspectives and empirical research questions to offer than systems theory. In other words, multi agent theory has the potential to be a more fruitful Lakatosian research program. The development of multi-agent organization theories following the decision-making oriented tradition of the Simon and March school can be seen as a reaction on the stagnation of system-theoretical organization theory.

Multi-agent theory tries to combine the *ecology metaphor* with the *mind metaphor*. In the ecology metaphor, agents are autonomous beings with metabolism, able to renew and reproduce themselves (Dyson, 1988). The ecological agents wander around in a natural and agent-made environment, seeking for fulfillment of their needs. The ecological agent is a well-developed response function agent (see paragraph 4). In the mind metaphor, organizations are seen as consisting of intelligent agents that communicate. The architecture of an organization is seen as composed of knowledge distributed over agents. The typical mind metaphor agent is a representational agent (see paragraph 4). The combination of the capabilities of a response function agent with those of a representational agent, and the combination of an ecology metaphor architecture with a mind metaphor architecture are main research questions in the construction of multi-agent systems.

Defining, measuring, and predicting performance based on knowledge-based multi-agent theories is a complex task. Firstly, there is the complexity and normative value hidden in the performance concept, as will be shown below. Secondly, there is the recursive nature of the knowledge concept (an agent knows of an agent that knows of an agent, and so on) leading to so-called endomorphic systems (Zeigler, 1990) that complicates modeling. Dennett (1978) has called this higher order intentionalities. Thirdly, only rather abstract, economical theories like agency theory and Williamson's theory are available to interpret the results of running multi-agent models like multi-agent SOAR (Carley, Hansen, Newell, and Prietula, 1992).

According to the pragmatic philosophers, performance should be measured in terms of maximum usefulness in serving our needs, reasonableness, and aesthetic value rather than using procedures, theories, abstract principles, or norms that are fixed and stem from dogmas or authority. The criterion of *usefulness* in serving our needs is always relative to some given point of view and purposes. The perceived usefulness is dependent on the biologically and socially evolved modes of adaptation and control of humanity. The criterion of *reasonableness* is based on the ideas of pluralism and of truth as a quality that is established intersubjectively, and that is time-dependent because of the ongoing investigation of reality. Peirce regards reasonableness as the ultimate end of all existence. According to Peirce, *aesthetic value* can be seen as being related to playfulness: the play element can be seen as the major factor in all art and even religious contemplation (Wiener, 1973: 569).

3. Philosophical views on organization

3.1. What are organizations?

Organizations do exist. This assertion is easily made, but presupposes a bundle of distinct beliefs, such as that the concept of organization has an extension (in the logical sense) or that entities that are not visible can be perceived or that reality can be described in layers or levels. We are talking here about organizations as entities, not as principles of arrangement. Because of the weak, that is to say theory-dependent, definition of organization, it is not uncommon in organizational theory to discuss the concept in terms of metaphors and paradigms (Morgan, 1986).

In preceding decades a (profit or nonprofit) organization consisted of buildings, factories, offices or comparable things in which people and machines in one way or another were tuned into one another. In the recent past this situation has already changed a bit and it will change more dramatically in the near future. Intelligent computers, data processing structures, information exchange procedures and communication architectures will lead to questions such as “where is the organization?", “where is the information?”, “where is the communication?” and even “where is the human being or agent?”. Before we continue this discussion, we firstly look at philosophical views on organizations. Doing things with and in organizations strongly depends on the view on organizations. At least two (extreme) positions are possible: the first position, realism, sees an organization as an object, the second position, constructivism, views an organization as a construction.

3.2. Realism

Hard core classical *realism* represents a perspective in which it can be defended that organizations are objects. Realism is the view that physical (and many nonphysical) objects exist independently of being perceived. Thus understood, “realism obviously reaffirms the standpoint of common sense” (Flew, 1979). “To exist, to be an entity, to have ontological status are the same.” (Bergmann, 1967: 3) and Hacking (1983) says about scientific realism that "entities, states and processes described by correct theories really do exist". (p. 21) Physical things are not the only real objects, as a matter of fact many perceivable things are part of the real world. In this sense organizations are part of reality independent of (specific) perceivers. Stated differently, although organizations may consist of many components and aspects, they exist on their own and can be studied on their own. As indicated, this position more or less corresponds to the common sense view on reality. Organizations as well as cars, electrons and stars are all part of reality and can be studied independent from perceivers. The differences become visible as soon as researchers try to describe the different objects in reality. Then, metaphors and analogies lead to different descriptions and sometimes even to different description languages. In the case of organizational theories (OMT) this leads to different predictions about the nature and the structure of organizations. In our opinion many researchers of organizations implicitly hold this realist point of view. Organizations are existing entities in a real world and through the different language descriptions an organizational "hard core" can be detected and revealed.

3.3. Constructivism

The other extreme position is nominal or social *constructivism*. Constructivism is the view that entities exist only if they can be constructed (or, intuitively shown to exist), and that statements are true only if a constructive proof can be given (Flew, 1979). Although constructivism originally was developed as a reaction against the platonist view on (mathematical) entities, its meaning has been expanded to other non-directly observable entities. This meant that organizations, institutions or companies can be studied as constructed entities. The prefix "social" refers to the sort of entities we are talking about. The other prefix "nominal" philosophically needs much more elaboration. Nominalism on the one hand is the view that universals such as redness are names and do not have an independent meaning, which makes nominalism to be contrary to platonic realism. On the other hand it is the modern doctrine that says that "the use of general terms that accounts for their meaning, and denotation, is in the mutual resemblance of the particular things to which they can be applied, or the recurrence in them of the general property indicated." (Flew, 1979: 232) This nominalist position is most strongly defended by Goodman (1972; Goodman and Elgin, 1988). "In describing an object, we apply a label to it. Typically that label belongs to a family of alternatives that collectively sort the objects in a domain. Such a family of alternatives may be called a scheme, and the objects it sorts its realm." (Goodman and Elgin, 1988: 7) This leads to the conclusion that in fact there only are labels that are expressed in symbols. This not only includes symbols used in languages, but also symbols in pictorial, notational or other representational systems. The only thing we can be sure about in the end is that in describing reality we work with all kinds of symbol systems. Symbol systems are artifacts (Goodman and Elgin, 1988; see also: Simon, 1969) which, certainly according to Goodman and Elgin, means that organizations also are artifacts. Applied to organizational theory a constructivist position implies that an organization is constructed on the one hand by the symbol systems we use in describing complicated action patterns between natural and artificial agents, and on the other hand results as a symbol system out of these interactions. An organization, therefore, is a constructed reality and exists as long as the interacting agents and their teamwork exist.

3.4. Ontological and semiotic engineering

The realist as well as the constructivist view are two rather well defined positions regarding organizations (Morgan, 1986; Gazendam, 1993). Intermediate positions are also possible, but we leave that for what it is worth. Interesting in relation to the topic of the article is the discussion about what the theoretical consequences are for multi-agent systems, for organizational architectures and for information and communication. The realist and constructivist view also have their repercussions for how to analyze and advice organizational structures and changes. A realist position implies that an organization exists in reality, independent from its perception by human beings and the conceptual apparatus they use. Therefore, an integrated framework for the description and design of organizations can be developed. This integrated descriptive framework can be seen as an *ontology*, a system of hypotheses about the objects and structures that exist. The ontology develops based on practical experience and discussion in the scientific community. Analysis, design and change use this relatively stable framework of description. The philosophically applicable term for this activity is *ontological engineering*.

The situation is largely different for a (nominalist) constructivist position. In this case there is no well defined object called "organization". An organization is a descriptive construction and therefore it is very difficult to develop an integrated descriptive framework for organizations. The description influences the constructed reality and at the same time the dynamics of the construction continually change the description. Analysis, design and change therefore very much depend upon the symbol systems used in the description. The attempt to develop an integrated framework, one language of description, is seen as uninteresting and even impossible. In contrast with ontological engineering we call this activity *semiotic engineering*.

Because of the artifact nature of the organization concept, we tend to a constructivist, semiotic engineering point of view. Organizations are constructs of the human mind, and can be studied by several conceptual systems or language systems that are not necessarily compatible. The consequences for the construction of multi-agent models is that one should be very careful in handling objects that represent the concept of ‘organization’. A direct representation of this concept should be avoided. If representation is necessary, it should be done as a kind of instrument in the agent’s mind, not as a separate entity. In the eye of the investigator using multi-agent models, emergent properties of the organization as a whole are interesting. In determining these emergent properties, the investigator should delineate the organization as a collection of agents that share a certain property such as a contract, or that have established a certain type of cooperation.

4. Psychologically plausible agents

4.1. Aspects of agents

Normally, an organization consists of an architecture being the cement, or the glue between many agents. The levels of complexity of architectures and agents define the complexity level of the organization. We discussed architecture in paragraph 3, now we will analyze agents. Agent sorts can be discerned regarding the presence or absence of the following components:

a) perception,
b) interaction (including learning in the sense of habit formation),
c) representation and interpretation (including learning in the sense of chunking) and
d) autonomy and self-repair.

With *perception* we mean that a system must be able to accept input in a general sense. This input may include visible, audible and tangible stimuli and the accepting system may vary from a lobster to a human being or even a computer system. *Interaction* is the process by which a system has contact with its environment. Stimuli as input in the system lead to output in the sense of responses. The reaction patterns of the system may result in learned behavior, that is to say that the habits are formed.

A system that internally symbolizes the environment is said to have *representations* at its disposal. Representations consist of sets of symbol structures on which operations are defined. Examples of representations are words, pictures, semantic nets, propositions or temporal strings. A representational system learns by means of chunking mechanisms and symbol transformations (Jorna, 1990; Jorna and Simons, 1992). A system is *self-reflective* if it is has a representation of itself, including its own position in the environment. This means that the system has self-representation, and can act based on this self-representation.

A system is said to be *autonomous* or *self-organized* if it can maintain itself in its environment based on its own action and its own learning. The knowledge of an autonomous system is only partially 'innate', and is largely acquired based on learning by experience or imitation. An autonomous system has the abilities to repair itself and return to a healthy state. It can handle goals in order to be able to attain a healthy or a desired state.

4.2. Sorts of agents

The four aspects contain a sort of agent hierarchy. An agent that only has perception at his disposal is at the lowest level and can hardly be called an (intelligent) agent, whereas an agent with self-organization is the highest level and normally has perception, interaction and representations. This last form is what we regularly call an agent that is reflective, intelligent and thoughtful. We consider humans beings to be good instantiations of these intelligent agents, whereas other mammals can ultimately go as far as to have representations. Computers do not fit in this picture at all or are said to have representations, but not self-organization. This depends on what one's definition of intelligence contains. One can understand that in the discussion of an organization as a multi-agent system it really depends on the characteristics of the agents, humans as well as computers, whether an organization is an intelligent system.

Not every agent is intelligent and not every intelligent system is an agent. The above described classification in perception, interaction, representation and autonomy can be related to qualification of agents. First we make a distinction in *single agents* and *multiple agents*. Second we subdivide agents in *response function agents, representational agents* and *representational response function agents*. We consider an agent to be a system consisting of several components, for example motor parts, sensory parts, including perception, and cognitive parts. The parts will not be discussed in detail in this article (see Posner, 1989; Newell, 1990). Concerning multiple agents the surplus component is a communication and coordination mechanism, that realizes the working of the interaction of agents. This mechanism has to be intelligent, but in contradistinction to the coordination within a single agent, does not have a physical or physiological basis as the carrier of the mechanism (Simon, 1945; Newell and Simon, 1976).

4.3. The response function agent

We start with an environment in which a system is present. The system is a cohesive, structured and organized entity. This entity operates in an environment, but no specifications of its operations are given. In a sense this entity is an agent, because it is self-contained, strives towards continuation and, if we look at the agent characteristics, it has perception and interaction including the possibility of learning in the sense of habit formation. It is emphasized that this agent does not have internal representations. Its cognitive domain is absent or empty. We call this agent a *response function agent* and it can be compared with the ant in the sand (Simon, 1969). Simon defends the position that the behavior of this ant can be called complex, although not intelligent, because it is a function of the complexity of the sandy environment that it has to cross. The ant perceives, interacts, uses his motor parts and even learns in the sense of habit formation.

4.4. The representational agent

Keeping the environment the same we can conceive another agent that we call a *representational agent*. This agent has representations at its disposal and is able to project external events internally into its cognitive domain. As already stated representations consist of symbol structures on which operations or manipulations are defined. Normally this agent is called an intelligent or cognitive agent. A representational agent has perception, representation and autonomy. The interaction is problematic, that is to say that there is no device that semantically interprets causal input and output. The most extreme view concerning interaction has been defended by Chomsky and Fodor, who in relation to cognitive processes argue in favor of philosophical rationalism. This position is closely connected to nativism. The nativist position sees agents as self-contained and self-oriented. This position also can be called methodological solipsism. As far as we talk about interaction it is rather low level reaction to stimuli. If we look at present day cognitive science most linguistically and logically oriented researchers have such a kind of agent in mind. Another problematic aspect of representational systems concerns the meaning of representation. Representation has many interpretations and it is not always clear which reading is the correct one: representation as process, as description structure or as one, two or three place predicate (Goodman, 1968/1981; Jorna, 1990).

4.5. The representational response function agent

If we keep the environment still the same, the third possible interpretation of an agent is the *representational response function agent*. This agent incorporates a really intelligent, interactive and cognitive system. It is able to perceive, to interact, to represent and to be autonomous. Cognitive processes include symbols, operations and semantic interpretable response functions. The goal of cognitive science is to model this sort of agents and to mimic them to a certain extent in computer systems. Representational response function systems behave on the knowledge level, as Newell called it. "There exists a distinct computer systems level, lying immediately above the symbol level, which is characterized by knowledge as the medium and the principle of rationality as the law of behavior." (Newell, 1982: 99) The agent equipped with the integration of representations and response has knowledge. "Knowledge", says Newell, "is whatever can be ascribed to an agent, such that its behavior can be computed according to the principle of rationality." (Newell, 1982: 105) The principle of rationality is expressed in the belief that an acting person will undertake those actions by which his goals are reached. In fact, we are dealing here with a variant of what Newell and Simon (1972) have called: "bounded rationality" (Simon, 1945). Until now cognitive scientists have not succeeded in reaching the goal of understanding in detail this sort of agents, although the latest developments in SOAR (standing for state, operator, and result) come close to what is ultimately possible with the symbolic approach in Artificial Intelligence.

4.6. Multi-agent systems

The hierarchy of single agents returns in the composition of *multi-agent systems*. In the first place it is possible to have multi-agents consisting of several response function systems. The situation is comparable to the single agent system in that the agents do not have internal representations. In this circumscription of multi-agents the other agents are considered as parts of the environment. The only difference is that there is a proximal and a distal environment. The system borders define an inner and an outer area. The agents all have perception and interaction at their disposal. To take up the example of Simon's ant we are talking here about a group of ants perceiving and interacting with each other. Coordination is absent or only defined in terms of reactions to behavior of other agents.

In the second place we have multi-agents consisting of representational systems. Each of these agents has internal representations in the sense of symbol structures and operations upon them, but there is no guarantee that symbol structures are similar and thereby communicative, although it might be expected that they all use one or another form of Mentalese, as Fodor (1975) suggested. In the same way as it holds for the single agent representational system it is of course doubtful whether interaction between the agents is semantically meaningful. For this is a weak point in the extreme version of present day cognitive science. In discussions about social cognition the issue of semantic interaction is ticked off, but not resolved. Intelligent coordination is hardly handled.

In the third place there are the representational response function systems in a multi-agent situation. The agents perceive each other and react to each other in a semantically rich and intelligent way. There might be a difference in the interaction patterns related to the environment compared to interactions with other agents within the area boundaries. Each agent has perception, interaction, representation and autonomy and manages to integrate this into the organization as a multi-agent system.

If we look at the six appearances of single and multiple agents and try to find examples of each of them in organizational settings (humans and computers) we see the following. Concerning single agents a connectionist machine is an example of a response function system, an expert system an example of a representational system without, however, the autonomy characteristic and a human information processing system an example of a representational response function system. Concerning multiple agents, examples of *multi-agent response function systems* are computer networks. Combinations of expert systems and humans are examples of *multi-agent representational systems*, whereas examples of *multi-agent representational response function systems* do not exist, besides perhaps an idealized group of human information processing systems. With respect to the hierarchies of single and multiple agents it should be noted that a higher level system exhibits the functions of a lower level system whereas the other way around is not applicable. This holds for single agents as well as for multiple agents.

4.7. Intelligent multi-agent systems?

One conclusion from the foregoing discussion is that modeling an intelligent agent is not yet completely realized within cognitive science and artificial intelligence. This sort of agent, however, is necessary for the construction of *intelligent multi-agent systems*. So, there is a gap between what is required and what can be realized by present day science. On the other hand organizational theory speaks about the coordination of multiple intelligent agents, but defines agents as sort of response function systems without any internal representation. These are not the sort of agents we strive after. So, cognitive science cannot model the multiple agent situation OMT should ask for, whereas OMT takes for granted a sort of multiple agent system that is not sufficiently equipped in order to behave intelligently.

The interesting point of looking at an organization as a semiotic entity cannot only be found in semiotic engineering, but also in the different sorts of signs that turn up in the communication and information structures of the various single and multiple agents. In semiotics it is normal to distinguish signals from signs and to subdivide signs in icons, indexes and symbols. Icons emphasize the similarity aspect, indexes the contiguity aspect and symbols the conventional aspect of signs. Signaling is equivalent with reporting and registration, while working with signs involves representing and interpreting. Signaling is a causal relation, whereas representing is mainly semantic.

Considering the sorts of single and multiple agents, only the response function systems work with causal relations, that is to say that the information exchange and the communication structures are in terms of signals. Representational systems work with icons, indexes and symbols, that is to say with semantic entities and relations. Response function systems and representational systems are mainly segregated. This means that the causal realm is largely isolated from the semantic realm. Developments within cognitive science and connectionism show that the latter is oriented at the symbolic or semantic domain, while the former is directed at the causal domain. The integration really takes place in the representational response function systems and these are the sorts of systems cognitive science as well as connectionism are striving to.

The analysis of the sorts of single and multiple agents shows that different sorts of agents emerge in organizations as well as in organizational theory. Combined with the semiotic analysis of signal and sign patterns between agents it is now possible to locate the assumptions within theories of organizations concerning the view on the concept of organization, the orientation on the agent and the sort of architecture in the organization including the level of communication.

Multi-agent representational response function systems are the most plausible type of multi-agent systems from a psychological point of view. The problem, however, is that they do not yet exist.

5. The organization as a multi-agent system

5.1. The architecture of a multi-agent system

Architecture is the way in which components make up a whole. The architecture of a multi-agent system is the way in which agents, processes performed by agents, and symbol structures used and produced by agents make up an organization. In order to be able to say something about the relation between architecture and organizational performance, we must be able to distinguish different sorts of architectures. This requires that we have an idea about which are the relevant aspects or dimensions of a multi-agent architecture, and which alternatives exist for each of those aspects or dimensions. Moreover, we can define plausibility requirements based on the architecture aspects or dimensions distinguished.

Our idea about the relevant aspects of the architecture concepts is based on an analysis of the architecture concepts from the perspectives of the machine metaphor, the organism metaphor, the ecology metaphor, and the mind metaphor of organization. An *organization* as an artifact, a human construct that is so predominant in our social life that we believe in the reality of the existence of it although it is no tangible object. This dual nature of organizations (see paragraph 2), artifact and reality, accounts for the importance of prescriptive theories in the field of organization and management theory. Modern organization theories do not state that there is one best way to organize, but make explicit the space of possibilities that exists for interpreting and designing organizations (Morgan, 1986; Gazendam, 1993).

Important characteristics of organizations, as expressed in modern organization theories, are (1) that the characteristics of the whole (the organization) are defined in terms of the characteristics of the parts (e.g., persons), while the characteristics of the whole in turn influence the characteristics of the parts; (2) that (symbol) structures created by agents like task descriptions and assignments are a source of actions by agents, while actions by agents use resources and create, in turn, symbol structures. This means that each analysis of organization theories must consider the following five categories of objects and their interactions:

- the agents in the organization:
- the organization as a whole;
- the symbol structures in the organization;
- actions by agents, together forming processes;
- resources.

Multi-agent models of organization represent a more sophisticated and refined way of looking at organizations (Bond and Gasser, 1988; Gasser and Hill, 1990). Besides persons or agents (which can also include intelligent computer agents), communication channels or blackboards are distinguished as basic elements of the organization. Within each agent, agent knowledge is considered to be important. Furthermore, there are dynamic processes of task allocation, cooperation, and communication. Therefore, one can no longer speak of part qualities, actions, and intermediate structure qualities only. A much more complicated model emerges that can have recursive properties. Based on these considerations, the following four aspects of the architecture of multi-agent systems are distinguished: *the topology of the components, the system function decomposition, the way of coordination and synchronization of processes, and the distribution of knowledge and language*.

5.2. The topology of the components

The *topology* of a multi-agent system is the way in which the basic components of an organization are ordered in space and time. The basic components of an organization are agents and material objects. Their ordering in space and time takes the form of work constellations and fixed communication paths. The most relevant alternatives that have been defined for the topology aspects are computational markets and computational hierarchies (Malone, 1987: Malone, Yates, and Benjamin, 1987; Miller and Drexler, 1988; Malone, 1992). Furthermore, computational ecologies can be distinguished. In the *computational market topology*, all agents have access to a common marketplace where information can be exchanged and negotiations take place. This topology is also known as the blackboard architecture (Engelmore, Morgan, and Nii, 1988). A variant on the blackboard architecture is the situation in which subdivisions of the blackboard exist that are specialized for a certain type of communication (like the fish market or the vegetable market). The *computational hierarchy topology* is characterized by the restriction of communication and negotiation of agents to a hierarchically structured organization: each agent can only communicate with its boss and its direct subordinates. According to Fayol (1916), this topology can lead to long communication lines and tedious communication processes. Therefore, Fayol invented the lateral communication (passerelle). This leads to the passerelle (gangplank) variant of the hierarchical topology. Hierarchical topologies are often connected to the notion of central power. This, however, is a subject to be subsumed under the process architecture discussed below. The fundamental concept in the *computational ecology topology* (Huberman and Hogg, 1988) is the environment, which is partially natural and partially agent made. Agents wander in the ecological environment seeking for fulfillment of their needs. (The environment concept of ecology has little to do with the passive and abstract environment of contingency theory.) The topological aspect of the ecological environment is that at certain places, resources can be found; and that at other places, agents, by convention, gather to do things together like buying and selling, negotiating, cooperate in work or in pleasure, and so on. Ecological architectures can technically be seen as a specific refinement of the multi-blackboard architecture, but based on a totally different basic concept.

In the multi-agent literature, there is a certain preference for decentralized architectures and connected decentralized power structures. An example can be found in the open system characteristics formulated by Hewitt (1985, 1986):

- massive parallelism;
- asynchrony;
- decentralized control;
- local knowledge and information that may be inconsistent;
- restriction of information exchange to explicit communication between agents;
- accommodation of the failures of individual agents by the community of other agents.

In the economical organization literature, on the other hand, there is a preference for hierarchical structures and centralized authority. An example is economic agency theory (Jensen and Meckling, 1976), which supposes centralized authority in combination with cheating behavior of the subordinates (see also Perrow, 1986: 224). We support the view of Hewitt that the real power of a multi-agent system lies in its decentralized and localized structure, and therefore reject the assumptions made beforehand by economical agency theory that centralized authority has to exist and cheating is a fundamental characteristic of subordinate agents.

We see the computational ecological topology as the most promising type of multi-agent topology. It extends the decentralized computational market topology by introducing active, mobile agents and notions of resources, cooperation, and communication.

5.3. The system function decomposition

The *system function decomposition* aspect of the architecture of an organization is the way the organization is composed of suborganizations fulfilling a specific function. These suborganizations can be decomposed into smaller suborganizations, and so on. The major alternatives for system function decomposition found in the organization and management literature are the aspect system decomposition, the check and balance decomposition, and the subsystem decomposition. The *aspect system decomposition*, also known as the 'functional' decomposition of organizations in smaller units, is based on specializations like production, marketing, finance, personnel, research, material management, and so on (Stoner, 1982: 267). An alternative to the functional decomposition based on specializations is the decomposition of functions in a web of actors that have to maintain a system of *checks and balances*, like in a social and political arena. Since Locke and Montesquieu we know that a division of powers is necessary to avoid arbitrariness of rule. A similar reasoning can be found in the field of the organization of financial administration aimed at avoiding corruption and embezzlement. The *subsystem decomposition* decomposes the system into smaller subsystems or organization modules (Kastelein), for instance based on product/market combinations. Each subsystem has all aspects of the system as a whole and is relatively autonomous. The aspect system decomposition generally has a relatively centralized power structure, while the subsystem decomposition generally has a more decentralized power structure.

The following *plausibility requirements* apply to the aspect of function decomposition based on a state space model of organization:

1. learning and recursiveness in learning, which means that the system must be able to rewrite its own program;
2. recursiveness of context in the form of several levels of response function complexity (e.g., response functions to be used depend on current tasks and goals, current tasks and goals depend on communication processes between agents);
3. recursiveness of data structures of input signals, state, and output signal;
4. recursiveness of functions;
5. knowledge representation or representation of the environment as part of the state;
6. representation of architectural components as part of the state;
7. massive parallelism of flows of input signals, flows of output signals, and internal processing;
8. agents and communication between agents as the basic phenomena to which data structures as well as functions are reducible;
9. clearly definable system boundaries and relations with the environment.

These plausibility requirements lead to a large field of conceptual, theoretical, and practical problems in factoring transparent, mathematically well-defined, models of organization. An important attempt to overcome the problems associated with the requirements 1 to 6 described is embodied in the SOAR system (Laird, Rosenbloom, and Newell, 1986; Laird, Newell, and Rosenbloom, 1987; Newell, 1990), a plausible simulation model of the human mind. There is no general guideline for building symbol systems fulfilling all requirements posed above, especially the requirements 7, 8 and 9 that are specific for multi-agent systems. Anyhow, the system function decomposition aspect of architecture as such is too abstract to be implemented directly in multi-agent systems. It has to be translated into the topological, process, and knowledge aspects of multi-agent architecture first. These aspects, again, have to be translated into agent capabilities and agent behavior as the basic dispositions and mechanisms driving a multi-agent system.

The system function decomposition aspect of a multi-agent system is a kind of description that stems from general systems theory. It is especially applicable in multi-agent models where agents have predetermined tasks, competences and power (and the emergence of such tasks , competences and power is not a subject of study). Especially the decomposition of competences and powers based on a system of checks and balances is interesting for multi-agent models of organization.

5.4. Coordination and synchronization of agent processes

The way in which the *coordination and synchronization of agent processes* take place is the next architecture aspect to discuss. Agent processes can be seen as being related to the wandering around of autonomous agents in a natural and agent-made environment, in which they cooperate in an occasional or regular way. The basic thrust of agents is to fulfill their needs like individual survival by maintaining their metabolism and self-renewal, and survival of the species by reproduction (Dyson, 1988). Agent processes also can relate to perception, communication and learning. The latter processes are explained more specifically under the knowledge and language aspect of architecture.

Process coordination and synchronization can be described from the agent viewpoint, adding coordination and synchronization later, or from the system viewpoint, which means distinguishing functions and decomposing them into processes. The agent viewpoint is explained here; the system viewpoint is identical with the system function decomposition aspect of architecture. The agent viewpoint presupposes capabilities of agents to perform certain processes. These capabilities are explained in paragraph 4. In addition, an approach to coordination and synchronization has to be formulated.

We have identified the following approaches to synchronization, which, by the way, do not seem to exhaust the possibilities in this field:

1. discrete event simulation;
2. Thompson's coupling mechanisms;
3. speech act theory;
4. the ecological environment in which symbol structures and signs reside;
5. protocols as describing communication standards;
6. grammars and lexicons as describing communication standards.

The *discrete event simulation* way to see the problem of coordination and synchronization has always been strongly dependent of the difficulties in realizing communication and synchronization between parallel processes realized in the computer. It elaborates variants of the semaphore mechanism (Birtwistle (1979); Goldberg and Robson (1983; 1989); Wieringa (1990)).

The *coupling mechanisms distinguished by Thompson* (1967) are pooled coupling, sequential coupling, and reciprocal coupling. In pooled coupling, agents share resources but are otherwise independent. This can be simulated effectively by discrete event simulation. In sequential coupling, agent processes are synchronized in a way that the results of precursor agents enable, as well as trigger, the processes of the subsequent agent. In reciprocal coupling, sequential coupling with feedback loops or feed-forward loops is present. From simulation studies it is well-known that feedback loops in simulation models can lead to cyclic or chaotic dynamic behavior of the simulation model (Gazendam, 1990).

Speech act theory as interpreted by Schäl and Zeller (1991) elaborates structures following from commitments already made. Coordination (denoted coordination(1) in order to distinguish it from our general coordination concept), collaboration, and co-decision are distinguished as forms of cooperation. Coordination(1) is based on a decomposition of a task in subtasks to be performed by each agent. In coordination(1), agents work independently, like in pooled coupling. In collaboration, agents have to synchronize their actions in order to perform a task correctly. An example is the movement of a table by human agents (Schäl and Zeller, 1991). Coordination based on sequential coupling would be a form of collaboration. Collaboration also can involve reciprocal coupling. In co-decision, agents construct a decision (which is a kind of symbol structure) or other artifact in a way where the contributions of each agent are not fixed beforehand, the feed-backs during the process of co-decision can change the task or role of each agent during the process. It is clear that co-decision encompasses elements of reciprocal coupling.

In the *ecology metaphor of organization*, where agents wander around in an environment, the existence of symbol structures and signs in the environment offers a basis for coordination and synchronization. Remember that the environment is partially natural and partially agent-created, the meeting places of agents, for example, are typically places distinguished based on conventions established in communication and expressed in symbol structures or signs. The coordination and synchronization that happens is rooted in the environment, especially the symbol structures and signs related to place and time.

Another way of looking at the synchronization in the ecological environment is to see the coordination and synchronization of agent processes as, albeit facilitated by symbol structures and signs, mainly dependent on the capabilities of agents, and the standards of communication they have realized. There are two ways for describing these standards of communication: by protocols and by grammars (and the related lexicons).

Standards of communication can be described in terms of protocols. This is the most simple solution to realize coordination and synchronization that is based on the exchange of symbol structures between agents rather than on posting signs at a semaphore. Most protocols implement a specific theory of negotiation, choice, planning and contracting. The protocols often suppose a market-like communication structure and sometimes a coordinating agent. Well-known protocols are contract net negotiation (Smith, 1980; Davis and Smith, 1983), partial global plans (Durfee, Lesser, and Corkill, 1987; Durfee and Lesser, 1988), and mediated negotiation (Sathi and Fox, 1989).

Another, in many ways complementary, method for describing communication standards is to describe the signs or symbols used, their composition into sign complexes or symbol structures in the form of a *lexicon*, and the *grammar rules* that govern the use and production of these sign complexes or symbol structures. In order to understand messages, and to produce meaningful messages, agents must have a world view. In order to understand each other, agents must not only share (share in the sense of 'having compatible') their world views, but also share their language systems. The use of existing organization theories as the basis for such a world view and language system, to be expressed in terms of grammar rules, has been explored in earlier publications (Gazendam, 1990; 1992; 1993).

The choice between these alternative mechanisms of coordination and synchronization is difficult, because the most sophisticated, language-oriented mechanisms that seem to be the most plausible ones from a philosophical and psychological point of view, are also the most difficult to implement. In practice, more simple mechanisms like speech act theory or the ecological environment mechanism seem to be more appropriate for the state of affairs in multi-agent modeling at this moment.

5.5. Distribution of knowledge and language characteristics

The *distribution of knowledge and language characteristics* over the agents of the multi-agent system, is an important aspect of multi-agent architecture. In the case of (1) the co-decision type of coordination and synchronization (Schäl and Zeller), (2) the coordination and synchronization based on conventions, symbol structures, signs, and (3) coordination based on the agent communication that depend on agent world views and language capabilities, the level of description that focuses on identifying processes and the coordination and synchronization often becomes too complex to be useful. In these cases, it is often more useful to concentrate on the contents of communication (knowledge expressed in signs and symbol structures) and the language system enabling communication. Culture, language and knowledge are connected to the use and production of symbol structures (or signs, in terms of Peirce).

Knowledge and language can be studied from a declarative, logical point of view or from a constructivist point of view. The *declarative, logical point of view* tries to capture the possible states of the knowledge in a multi-agent system by describing the logical conditions that have to be fulfilled by the knowledge state. This description can be a description in general, and/or in terms of the relationships between knowledge states in time (where time is often defined as relative to actions considered or taken by agents), and/or in terms of relationships between the knowledge that agents have. It turns out that the description of the possible states of knowledge of a single agent is already very difficult, and requires a specialized logical language that is expressive enough to describe agent knowledge, agent actions, decision-making by agents about actions, and the relationship between knowledge and action. Attempts to formulate such a language often take the form of a new logical system with its own symbol set, grammar rules, axioms, and derived statements (Masuch, 1990; Masuch and Warglien, 1992). Classical logical models of knowledge and action are those of Moore (1985) and Halpern and Fagin (1989). In Halpern and Fagin' s model, a fixed number of agents are present, agents have a common memory (a kind of blackboard), there is no private agent memory other than states, and time moves in discrete steps (which means that it is not a discrete event model). A promising development in which several of the limitations of Halpern and Fagin's model have been removed is the action logic defined by Huang, Masuch, and Pólós (1991). A declarative, logical approach can fruitfully be applied to the analysis of existing organization theories, pointing out the weaknesses or even contradictions in those theories. Because of the complexities in the field of logical languages, we expect that the role of the declarative logical approach in constructing new theories will be relatively small. Its power lies in analyzing theories, not in inventing new ones.

The *constructivist point of view* does not aim at clarifying the possible states of knowledge of a multi-agent system, but at clarifying the processes that lead to knowledge and language capabilities. Once these processes have been understood in enough detail to simulate them, states of knowledge, and of language capabilities can be produced by the simulation model rather than derived from logical circumscription. Two important types of processes in the constructivist point of view are interpretation (of the world as well as of signs, symbol structures, and messages in a language) and learning. Other processes, like decision-making, acting, and language production probably can be understood better if we have a better understanding of interpretation and learning.

The interesting point of *interpretation* is that it sees the formation of knowledge not as a process of perception in which the agent is rather passive, but as a process in which the agent is active and forms a representation of the world governed by an interpretation framework. Interpretation frameworks can differ from agent to agent, they have been described as microtheories by Hewitt and Lenat, and as agent grammars by Gazendam (1992).

The interesting point of *learning* is that knowledge is not seen as fixed, but as fundamentally changing. All elements of agent knowledge must have been created once by processes of interpretation and/or learning. Interesting models of learning are cumulative learning by neural networks and classifier systems (Wasserman, 1989; Goldberg, 1989), learning by chunking (Laird, Rosenbloom, and Newell, 1986; Laird, Newell, and Rosenbloom, 1987; Newell, 1990), and learning by transformation of representations using grammars (Wand, 1993; Gazendam, 1993). Most learning, however, is social learning by imitation and the passing of *memes*, packages of culture and knowledge, in social interaction (Dawkins, 1976).

Multi-agent systems are necessary for learning processes in which language capabilities are learned, conventions arise, and specialized knowledge based on task division emerges. Language acquisition of human agents, based on interaction between these human agents, turns out to be a learning process in which the forming of conventions about signs and the formation of concepts interact in a way where the signs are dominant (Müller, 1991). Several multi-agent computer models have been made showing aspects of multi-agent learning. Two separate learning systems (two 'agents'), each of which develops its own specialization in a task that has to be accomplished together, has been shown to perform better than a single learning system of comparable total capacity (a single 'agent') (Cohen, 1992).

A structured theory of multi-agent learning, will have to account for (1) the existence, development and use of common knowledge in organizations, and (2) the learning processes of agents in interaction with this common knowledge. The common knowledge in organizations (the 'culture layer') consists of a common world model, rules and examples of expected behavior, and a common language of communication. The learning of agents in organizations encompasses the adaptation of agents to the organizational culture layer, changes in the culture layer, and learning patterns leading to the emergence of specializations and new language elements. Note that the organization, as an entity, is unable to learn. *Only agents can learn*. Such a theory of organizational, multi-agent, learning, however, has not emerged yet.

The construction of process models of communication, interpretation and learning seems to be a promising one for developing new perspectives, models and theories for multi-agent systems. It weakness might be the tendency to make overcomplex models in which it is difficult to determine the relations between the variables used.

6. Research perspectives

6.1. A research program relating organization theories and multi-agent systems

Organization theories sometimes predict organizational performance based on the organization type and task environment type. We can use organization theories for the design of multi-agent systems, and we can use multi-agent systems to test organization theories.

Existing organization theories in the OMT literature contain knowledge that can be used for constructing plausible multi-agent models. This knowledge, part of which can be described in the form of an organization grammar, is used by the organization theorist in constructing models of organization (Gazendam, 1993). Furthermore, multi-agent models presuppose that agents have some notion of what an organization is and how it functions. This knowledge has to be incorporated in an agent grammar.

Agent grammars (including a symbol and symbol structure lexicon), combined with a control mechanism, are sufficient to specify intelligent behavior (Gazendam, 1992). Such a specification would be fitting in the semiotic approach to multi-agent systems.

For the conceptual specification of organization theories, an ontological as well as a semiotic method can be used (Jorna, van Heusden and Posner, 1993). An ontological method for analyzing organization theories is, for instance, based on ontological engineering (Lenat and Guha, 1989) or Bunge's ontology as a frame of reference (Wand, 1993). A semiotic approach to analyzing organization theories is offered by the CAST method (Gazendam, 1993). CAST (Conceptual Analysis and Specification of organization Theories) aims at the translation of verbal organization theories into a conceptual model using a BNF-like conceptual modeling language. This conceptual model shows the theory's contents in a clear way, enabling a comparison and analysis of theories as well as a further formalization leading to a logic model or computer model. The CAST method shows that there is a variety of metaphors for describing and designing organizations, and that each metaphor leads to its own language system, system of concepts, and reasoning form.

Multi-agent models of organization can be used for testing organization theories. In order to do this, the following steps have to be taken:

1. the realization of the process frame of the multi-agent system;
2. the realization of a specific type of agent and attaching a number of these agents to the multi-agent frame;
3. feeding the agents with organizational knowledge;
4. exposing the multi-agent system to one or more task environments in order to record performance and to test predictions of organization theories about relationships between organization type and characteristics of the task environment.

Step 1 consists of the *realization of the process frame* of the multi-agent system. This architecture consists of a number of places where agents can be fitted in, a global communication channel, and agent-to-agent channels. The process frame is a kind of 'bus' to which processors or parallel processes representing agents can be attached. Assumptions about the utilization of the global channel versus agent-to-agent channels can give the multi-agent system a more hierarchic versus a more market-like communication topology. Synchronization of agent processes has to be done based on the rhythms of the behavior cycles of each agent. An example is the fundamental 'interpret-choose-act-learn' cycle of agents. Synchronization assumptions as well as communication topology assumptions have to be made in the organization theory to be studied. The variation of the number of agents is a subject of the process architecture. The communication channel architecture may be dependent on the task environment; for instance, communication may be dominated by a number of work places that are designed as part of a production process. These work places are parts of the global communication channel, which can have a structured (blackboard) nature. Other mechanisms to structure the information in the global communication channel (at the blackboard) are agenda's, plans, requests, and agent rights.

Step 2 consists of realizing *a specific type of agent* and attaching a number of these agents to the multi-agent frame. We have distinguished three basic agent types: the response function agent, the representational agent, and representational response function agent. Related to these agent types, we have described four layers of agent behavior: the response function layer for low-level automated behavior, the representational layer for task-related routine behavior, the cognitive layer for reasoning about tasks and the world, and the communication layer for understanding and producing language. Furthermore, agents can differ according to the learning mechanisms they use at the behavior layers distinguished. As we have discussed above, it is necessary to choose the agent's characteristics according to the organization theory to be studied.

Step 3 consists of *feeding the multi-agent system* with organization design information. This can be done by offering a training to the response function layer, symbol structures about tasks to the task execution layer, world models to the cognitive layer, and grammars to the communication layer. Generally, organization theories distinguish design parameters, design goals, and organization types related to these design goals. Different organization types will require different organization design information to be fed into the system.

Step 4 consists of *exposing the multi-agent system* (or the collection of multi-agent systems that has been created as a result of differences in design information and other parameters) to one or more task environments in order to record performance and to test predictions of organization theories about relationships between organization type and characteristics of the task environment.

6.2. Building multi-agent systems

The main question in building multi-agent systems is the integration of the following components:

- a database program;
- a discrete event simulation shell;
- an expert system or problem solver;
- a learning component.

In hybrid, object-oriented expert system shells the expert system is dominant (for example: NEXPERT OBJECT). In Plural SOAR (Carley, Kjaer-Hansen, Newell, and Prietula, 1992) , the problem solver and integrated learning component are dominant. In most existing multi-agent systems, the discrete event simulation shell is dominant enabling the (semi) parallel existence of processes for each agent and each blackboard during runtime. In our research, we aim at comparing these multi-agent architectures based on their use in simulation experiments. One of these architectures is Multi-Agent SOAR, another is a multi-agent organization modeling shell explained below.

The Information Strategy Model (ISM) is a multi-agent organization model written in Smalltalk80 (Gazendam, 1990). It simulates the choice and implementation of information management strategies. The basic structure of the Smalltalk80 multiprocessing and simulation classes is used to model agents, objects, resources and blackboards. Some agents use a personal knowledge base to fulfill their tasks. These knowledge bases use HUMBLE, a Smalltalk expert system shell with MYCIN-like features. The main results of ISM are insight in the stability or instability of strategies and the effects of long implementation trajectories on strategy choice. Based on ISM, an multi-agent organization modeling shell has been developed, adding a better integration of simulation shell and expert system components, a revised expert system shell, goal-oriented behavior, learning, and representation of space as well as time. Experiments with this multi-agent shell are a subject of ongoing research.

7. Conclusions

An organization is an artifact that is so predominant in our social life that we believe in its existence although it is not a tangible object. Because of the artifact nature of the organization concept, we tend to a constructivist, semiotic engineering point of view. Organizations are constructs of the human mind, and can be studied by several conceptual systems or language systems that are not necessarily compatible. Semiotic engineering uses semiotic theory because of the absence of a physiological carrier for coordination and communication between agents.

Multi-agent systems presuppose a representational response function agent. However, clear theories of representational response function agents do not exist, nor do computer models of representational response function agents exist. Therefore, plausible multi-agent systems have not been realized yet. Connectionism and cognitive science is restricted to response function agents and representational agents. In theory as well as in simulation and modeling environments there are several problems in combining response function agents with representational agents.

With respect to the organization or architecture of a multi-agent system we have distinguished the aspects of topology, system function decomposition, coordination and synchronization mechanisms, and knowledge and language. We see the computational ecological topology as the most promising type of multi-agent topology. It extends the decentralized computational market topology by introducing active, mobile agents and notions of resources, cooperation, and communication. The system function decomposition aspect of a multi-agent system is a kind of description that stems from general systems theory. It is especially applicable in multi-agent models where agents have predetermined tasks, competences and power (and the emergence of such tasks , competences and power is not a subject of study). Especially the decomposition of competences and powers based on a system of checks and balances is interesting for multi-agent models of organization. The choice between alternative mechanisms of coordination and synchronization is difficult, because the most sophisticated, language-oriented mechanisms that seem to be the most plausible ones from a philosophical and psychological point of view, are also the most difficult to implement. In practice, more simple mechanisms like speech act theory or the ecological environment mechanism seem to be more appropriate for the state of affairs in multi-agent modeling at this moment. A declarative, logical approach can fruitfully be applied to the analysis of existing organization theories, pointing out the weaknesses or even contradictions in those theories. Because of the complexities in the field of logical languages, we expect that the role of the declarative logical approach in constructing new theories will be relatively small. Its power lies in analyzing theories, not in inventing new ones. The construction of process models of communication, interpretation and learning seems to be a promising one for developing new perspectives, models and theories for multi-agent systems. It weakness might be the tendency to make overcomplex models in which it is difficult to determine the relations between the variables used.

Our final conclusion has to be that plausible knowledge based multi-agent theories do not exist yet. Ingredients and components do exist, but cognitive science as well as OMT do not succeed in the complete understanding and explanation of these multi-agent systems.

Literature

Ashby, W.R.

1956 "Self-regulation and requisite variety." in: W.R.Ashby. *Introduction to Cybernetics*. New York: Wiley. Reprinted in: F.E.Emery (ed.). Systems thinking. Harmondsworth: Penguin Books, 1970: 105-124.

Bergmann, Gustav

1967 *The Metaphysics of Logical Positivism: 2nd Ed*. Madison: University of Wisconson Press.

Birtwistle, G.M.

1979 DEMOS: A system for discrete event modeling on Simula. Houndmills: MacMillan.

Bond, A.H., and L.Gasser

1988 "An analysis of problems and research in distributed artificial intelligence." In: A.H.Bond and L.Gasser (eds.). *Readings in Distributed Artificial Intelligence.* San Mateo, CA: Kaufmann.

Burns, T., and G.M.Stalker

1979 *The Management of Innovation.* London: Tavistock Publishing.

Carley, Kathleen, Johan Kjaer-Hansen, Allen Newell, and Michael Prietula

1992 "Plural-Soar: A Prolegomenon to Artificial Agents and Organizational Behavior." In: Michael Masuch and Massimo Warglien (eds.). *Artificial Intelligence in Organization and Management Theory: Models of Distributed Activity.* Amsterdam: North-Holland: 87-118.

Cohen, M.D.

1992 "When can two heads learn better than one?: Results from a computer model of organizational learning." In: Masuch, Michael, and Massimo Warglien (eds.). *Artificial Intelligence in Organization and Management Theory: Models of Distributed Activity.* Amsterdam: North-Holland: 175-188.

Davis, R. and R.G.Smith

1983 "Negotiation as a metaphor for distributed problem solving.", *Artificial Intelligence* 20: 63-109.

Dawkins, Richard

1976 *The Selfish Gene*. Oxford: Oxford University press.

Dennett, Daniel C.

1971 "Intentional Systems." *Journal of Philosophy*, 8: 87-106.

1978 *Brainstorms*. Montgomery, VT: Bradford Books.

1991 *Consciousness Explained.* London: Allen Lane / The Penguin Press.

Durfee, E.H., and V.R.Lesser.

1988 "Using Partial Global Plans to Coordinate Distributed Problem Solvers.", In: A.H.Bond and L.Gasser (eds.). *Readings in Distributed Artificial Intelligence*. San Mateo, CA: Morgan Kaufmann: 285-193

Durfee, E.H., V.R.Lesser, and D.D.Corkill

1987 "Coherent cooperation among communicating problem solvers.", *IEEE Transactions on Computers* C-36: 1275-1291.

Dyson, Freeman

1988 *Infinite in all Directions.* London: Penguin.

Engelmore, R.S., A.J.Morgan, and H.P.Nii

1988 "Introduction." In: R.S.Engelmore and A.J.Morgan. Blackboard Systems. Wokingham: Addison-Wesley: 1-22.

Fayol, Henri

1916/1956

"Administration, industrielle et générale". Extrait du *Bulletin de la Société de l'Industie Minérale*, 3e livraison de 1916. Quarantième Mille. Paris: Dunod, 1956.

1916/1984

"Administration, industrielle et générale". *Bulletin de la Société de l'Industie Minérale*. Translated and Revised by Irwin Gray as "General and Industrial Management", London: Pitman, 1984.

Flew, Antony

1979 *A Dictionary of Philosophy*. London: Pan Books.

Fodor, J.A.

1975 The Language of Thought. Hassocks, Sussex: Harvester Press.

Gasser, Les, and Randall W.Hill, Jr.

1990 "Coordinated Problem Solvers." *Annual Review of Computer Science*, 1990, 4: 203-253.

Gazendam, Henk W.M.

1990 "Expert Systems Supporting Organization and Information Management." In: Masuch, Michael. (ed.). *Organization, Management, and Expert Systems: Models of Automated Reasoning.* Berlin: De Gruyter: 123-153.

1992 "Organization Theories as Grammar." In: M.Masuch and M.Warglien (eds.). *Artificial Intelligence in Organization and Management Theory: Models of Distributed Activity.* Amsterdam: North-Holland: 41-63.

1993 *Variety Controls Variety: On the Use of Organization Theories in Information Management*. Groningen: Wolters-Noordhoff, 1993.

Gazendam, Henk W.M., René J.Jorna, and Kenneth R. Blochowiak

1991 "The Mind Metaphor for Decision Support Systems." In: A.A.Verrijn-Stuart, H.G.Sol, and P.Hammersley (eds.). *Support Functionality in the Office Environment.* Amsterdam: Elsevier (North -Holland) / IFIP: 203-223.

Goldberg, Adele, and David Robson

1983 *Smalltalk-80: The Language and its Implementation.* Reading, MA: Addison-Wesley.

1989 *Smalltalk-80: The Language.* Reading, MA: Addison-Wesley.

Goldberg, D.E.

1989 *Genetic Algorithms in Search, Optimization, and Machine Learning.* Reading, MA: Addison-Wesley.

Goodman, Nelson

1968/1971 *Languages of Art: Second Edition*. Brighton: Sussex: The Harvester Press.

1972 *Problems and Projects*. Indianapolis: Hackett Publishing Company.

Goodman, Nelson, and Catherine Z. Elgin

1988 *Reconceptions in Philosophy and Other Arts and Sciences*. London: Routledge.

Hacking, I.

1983 Representing and Intervening. Cambridge: Cambridge University Press.

Halpern, J.Y., and R.Fagin

1989 Modeling knowledge and action in distributed systems. *Distributed Computing* 3, no. 4: 159-179.

Hewitt, C.E.

1985 "The challenge of open systems." *Byte*, 10, no. 4: 223-242.

1986 "Offices are open systems." *ACM Transactions on Office Information Systems* 4, no. 3: 271-287.

Huang, Zhisheng, Michael Masuch, and László Pólós

1991 "A Preference Logic for Rational Action." *CCSOM Report 92-45*. Amsterdam: University of Amsterdam.

Huberman, B.A., and T.Hogg

1988 "The Behavior of Computational Ecologies." In: B.A.Huberman (ed.). *The Ecology of Computation.* Amsterdam: North-Holland: 77-115.

Jensen, Michael C., and William H.Meckling.

1976 "Theory of the Firm: Managerial behavior, Agency Costs and Ownership Structure". *Journal of Financial Economics*, 26: 305-360.

Jorna, R.J.

1990 *Knowledge Representation and Symbols in the Mind: An Analysis of the Notion of Representation and Symbol in Cognitive Psychology*. Tübingen: Stauffenburg Verlag.

Jorna, R.J., and J.L.Simons

1992 *Kennis in organisaties: Toepassingen en theorie van kennissystemen.* (Knowledge in Organizations: Applications and Theory of Knowledge Base Systems). Muiderberg: Coutinho.

Jorna, R.J., B. Van Heusden, and R.Posner (eds.)

1993 *Signs, Search, and Communication: Semiotic Aspects of Artificial Intelligence*. Berlin: Walter de Gruyter.

Kephart, J., T.Hogg, and B.Huberman

1989 "Dynamics of computational ecosystems: Implications for DAI." In: In: L.Gasser and M.N.Huhns (eds.). *Distributed Artificial Intelligence: Volume 2.* San Mateo: Morgan Kaufman: 79-86.

Kieser, A., und H.Kubicek

1983 *Organisation: 2.Auflage.* Berlin: De Gruyter.

Laird John E., Allen Newell, and Paul S. Rosenbloom

1987 "SOAR: An Architecture for General Intelligence." *Artificial Intelligence*, 33: 1-64.

Laird, John E., Paul S. Rosenbloom, and Allen Newell

1986 *Universal Subgoaling and Chunking: The Automatic Generation and Learning of Goal Hierarchies.* Boston: Kluwer.

Lawrence, P.R., and J.W.Lorsch

1967 "Differentiation and integration in complex organizations." *Administrative Science Quarterly* 12 (1967): 1-47

Lenat, Douglas B., and R.V.Guha

1989 *Building Large Knowledge-Based Systems: Representation and Inference in the CYC Project.* Reading, MA: Addison-Wesley.

Malone, T.W:

1987 "Modeling coordination in organization and markets." *Management Science* 33: 1317-1332.

1991 "Analogies between Human Organizations and Artifial Intelligence Systems." In: Masuch, Michael. (ed.). *Organization, Management, and Expert Systems: Models of Automated Reasoning.* Berlin: De Gruyter: 21-40.

Malone, T., J.Yates, and R.Benjamin

1987 “Electronic Markets and Electronic Hierarchies”. *Communications of the ACM*, 26: 430-444.

Masuch, Michael (ed.)

1990 *Organization, Management, and Expert Systems: Knowledge-based Models of Organization.* Berlin: De Gruyter.

Masuch, Michael, and Massimo Warglien (eds.)

1992 *Artificial Intelligence in Organization and Management Theory: Models of Distributed Activity.* Amsterdam: North-Holland.

Miller, Mark S., and K. Eric Drexler

1988 "Markets and Computation: Agoric Open Systems." In: B.A.Huberman (ed.). *The Ecology of Computation*. Amsterdam: North-Holland: 133-176.

Moore, R.C.

1985 "A formal theory of knowledge and action". In: J.Hobbs and R.C.Moore (eds.). *Formal Theories of the Commonsense World*. Norwood, NJ: Ablex Publishing Corp.: 319-358.

Morgan, Gareth

1986 *Images of organization.* Beverly Hills: SAGE.

Müller, A.

1993 "On Knowledge Representing Interacting Systems." In: R.J.Jorna, B. van Heusden, and R.Posner (eds.). *Signs, Search, and Communication: Semiotic Aspects of Artificial Intelligence*. Berlin: Walter de Gruyter: 271-305.

Newell, A.

1982 "The Knowledge Level." *Artificial Intelligence*, 18, 1: 87-127

1990 *Unified Theories of Cognition: The William James Lectures, 1987.* Cambridge, MA: Harvard University Press.

Newell, A., and H.A.Simon

1972 *Human Problem Solving.* Englewood Cliffs, NJ: Prentice-Hall.

1976 "Computer Science as Empirical Enquiry: Symbols and Search." *Communications of the ACM*, 19 (March 1976): 113-126.
Reprinted in: J.Haugeland (ed.). *Mind Design*. Cambridge, MA: The MIT Press, 1981: 35-66.

Peirce, Charles Sanders

1878 "How to Make Our Ideas Clear." *Popular Science Monthly* (Jan 1878): 286-302. Reprinted in Philip P. Wiener (ed.). *Charles S. Peirce: Selected Writings: Values in a Universe of Chance*. New York: Dover, 1958: 113-136.

1905 "What Pragmatism Is." *The Monist,* 15 (April 1905): 161-181. Reprinted in Philip P. Wiener (ed.). *Charles S. Peirce: Selected Writings: Values in a Universe of Chance*. New York: Dover, 1958: 180-202.

1931 *Collected Papers: Volume 1.* Cambridge, MA: Belknap Press of Cambridge University Press. Reprinted in 1965.

1933 *Collected Papers: Volume 4.* Cambridge, MA: Belknap Press of Cambridge University Press. Reprinted in 1960.

Perrow, Charles

1986 *Complex Organizations: A Critical Essay.: Third Edition.* New York: Random House.

Posner, M.I.

1989 *Foundations of cognitive science.* Cambridge, MA: The MIT Press.

Sathi, A., and M.S.Fox

1989 "Constraint-directed negotiation of resource reallocations." In: L.Gasser and M.N.Huhns (eds.). *Distributed Artificial Intelligence: Volume 2.* San Mateo: Morgan Kaufman: 163-194.

Schäl, T., and B.Zeller.
1991 "Design Principles for Cooperative Office Support Systems in Distributed Process management.", In: A.A.Verrijn-Stuart, H.G.Sol and P.Hammersley (eds.). *Support Functionality in the Office Environment*. Amsterdam: North-Holland.
Schmidt, Kjeld
1991 "Cooperative Work: A Conceptual Framework.", In: Jens Rasmussen, Berndt Brehmer and Jaques Leplat (eds.). *Distributed Decision Making: Cognitive Models of Cooperative Work.* Chichester: Wiley.
Simon, Herbert A.
1945/1976 *Administrative Behavior: A study of decision-making processes in administrative organization: Third edition.* New York: The Free Press.
1969/1981 *The sciences of the artificial.* Cambridge, MA: The MIT Press.
Smith, R.G.
1980 "The contract net protocol: High level communication and control in a distributed problem solver.", *IEEE Transactions on computers* C-29: 1104-1113.
Stoner, J.A.F.
1982 *Management: Second edition.* Englewood Cliffs, NJ: Prentice-Hall.
Thompson, J.D.
1967 *Organizations in action*. St.Louis: McGraw-Hill.
Wand, Yair, and Ron Weber
1993 “On the Ontological Expressiveness of Information Systems Analysis and Design Grammars.” *Journal of Information Systems,* 3: 217-237.
Wasserman, Philip D.
1989 *Neural Computing: Theory and Practice*. New York: Van Nostrand Reinhold.
Wiener, Philip P.
1973 "Pragmatism." In: Philip P. Wiener (ed.). *Dictionary of the History of Ideas: Studies of Selected Pivotal Ideas: Volume III.* New York: Charles Scribner's Sons: 551-570.
Wieringa, R.J.
1990 Algebraic foundations for dynamic conceptual models. *Ph.D. Thesis,* Amsterdam: Vrije Universiteit.
Williamson, O.E.
1975 *Markets and Hierarchies.* New York: The Free Press.
1985 *The Economic Institutions of Capitalism.* New York: The Free Press.
Woodward, Joan
1965 *Industrial Organization: Theory and Practice.* London: Oxford University Press.
Zeigler, Bernard P.
1990 *Object-Oriented Simulation with Hierarchical, Modular Models: Intelligent Agents and Endomorphic Systems*. Boston, MA: Academic Press.