Keynote Address: Formal and Informal Aspects of Intelligent Agent-based Systems

John-Jules Ch. Meyer

Utrecht University, Department of Computer Science,
Intelligent Systems Group,
P.O. Box 80099, 3508 TB Utrecht,
The Netherlands

email: jj@cs.uu.nl

Key words: Intelligent agents, Agent-oriented programming, Modal logic, Multi-agent systems

Abstract: Recently on the borderline between computer science, software engineering and artificial intelligence the subject of *intelligent agents* has become in the limelight very much, cf. [WJ94]. Interestingly much work has been done on both the foundations and practical applications of agent technology. However, as yet there remains a big gap between these two in the sense that it is not obvious at all how the foundational work on agent logics and other formalisms relate to the practice of designing and implementing agent-based systems for such applications as the web and other user assistants more in general, but also for the control of hardware entities such as intelligent robots. In this paper we will discuss some of the work that has been done on the foundations of agents as well as the problems that remain to 'bridge the gap' between theory and practice.

1. INTRODUCTION

The area of intelligent agents is an emerging technology which is aimed at constructing more 'intelligent' software in the sense that such software does not merely react to a user but is also supposed to anticipate on its user's needs and interests [Mae95]. A key concept here is that of *autonomy*. Agent software should possess a certain degree of autonomy in order to be able to

The original version of this chapter was revised: The copyright line was incorrect. This has been corrected. The Erratum to this chapter is available at DOI: 10.1007/978-0-387-35500-9_30

E. D. Falkenberg et al. (eds.), *Information System Concepts: An Integrated Discipline Emerging*

fulfil the requirement of being *proactive* instead of merely *reactive*. In the literature many words are spent (or spilt as some people would say) on the definition of an agent, not unlike the endless discussions one or two decades ago on the definition of an object in the object-oriented programming paradigm.

In this paper I will not attempt to try and give yet another definition of what an agent is, nor do I intend to repeat the many definitions from the extensive literature on this topic. Rather I like to focus on a few aspects concerning the concept of an intelligent agent. Viewed from a distance it is of course quite interesting to see a community searching for the 'right' definitions and concepts concerning a new technology, and I myself think it is certainly not a waste of time to think about this carefully and make an inventory of the possibilities. On the other hand the field should not be blocked by this activity, since there is already a lot of interesting things to be done employing partial definitions or by maintaining various distinct variants of definitions / notions. As an aside we remark that Lakatos [LM70] argued for the rationale of the practice of having multiple distinct theories available in the methodology of the exact sciences. In the case of agent technology I believe this is even more rational since it is quite possible that in the end there will be several kinds of intelligent agents. For example, it is not *a priori* evident at all that agents that will be designed for obtaining information from the Web will have the same characteristics as those that will be employed for controlling pieces of hardware like e.g. intelligent robots.

What in my opinion makes an agent interesting from a scientific point of view is that in order to display some kind of autonomy it will have to be endowed with several *attitudes* for which some kind of 'mental state' of the agent is a convenient metaphor. Of course, one can now engage in very troublesome discussions as to the nature of such a mental state, and if / how it relates to human mental states, but for the purpose of this paper I will not discuss this in much depth since in my view the use of a mental state as a metaphor is already a very potent means for describing and specifying (the behaviour of) intelligent agents. This is especially so since in philosophy there has been done much work on several of these mental or cognitive aspects from a perspective of *logic*. The notions that are arising from this work are particularly fit to be used for a foundation of intelligent agents since these abstract away from the human / psychological aspect, so that they are in principle amenable to artifacts as well. In robotics this has been realised recently in a new subarea called *cognitive robotics* ([D98]).

The logic referred to in the last paragraph is known under the name *modal logic*, the logic of *modality*. In this context by modalities are meant things like knowledge, belief, time, desires, goals, intentions, obligations, etc., in fact quite a range of distinct notions that have to do with the mental or cognitive state of an agent.

Interestingly, although these are all formal logics in the strict sense of the word, the notions they describe are still rather vague or abstract in the sense that still many more concrete instances of these notions are captured by them, and no particular choice is made. I regard this 'vagueness' as a feature rather than as a bug: as argued before this enables us to use these logics for describing the mental states of agents in an abstract way which allows for a non-human implementation of these mental states. This is what I mean by the metaphoric use of these mentalistic concepts: they allow us to structure the architecture of the systems we like to build without having to look too much at implementations as they occur in nature (such as the human brain).

On the other hand, unfortunately the abstract notions of the modalities as defined by modal logic also cause in some sense the problem of 'bridging the gap' between (logical) theory and practice, since by their very abstract nature they do not give much indication as to how these notions should be implemented into a real system. However, as I shall argue in the sequel this problem might be overcome by means of the use of a modal logic that is close to the notion of programming (viz. dynamic logic [Har84]).

2. MENTAL STATES AS SYSTEM STATES

A very important notion when considering systems — whether artificial or not — is that of a (system) state. Since the beginning of computer science this has been realised, and one encounters states in e.g. automata theory and the semantics of programming languages. Usually there is also some notion of state transition capturing how computation proceeds stepwise by means of executing a program, for example. In the semantics of programming this idea has given rise to the formal notion of a transition system which is a formal system by means of which one can derive transitions of the system / program at hand.

In programming agents or agent-oriented programming, as it is called since [Sho91, Sho93], the notion of a state also occurs naturally, but whereas in traditional programming the state typically involves such matters like the values of variables, here things are a bit more involved. Using the mentalistic metaphor one arrives at the concept of a *mental state* of an agent [HS93], which evolves over time. So what kind of things should we think of when talking about an agent's mental state? Well, it depends on the application and the particular kind of agent one needs for this. Generally one can say that in order to display some kind of autonomy the agent needs to have the disposal of *information* and, perhaps more controversially, *motivation*. So the agent has to maintain certain *knowledge* about the environment it is to work in (or perhaps rather *beliefs*, since the information maintained may not be actually true of the world / environment; for instance it may be outdated or just obtained erroneously by means of a faulty communication or sensor device). But in order to be proactive and anticipate on things that happen in its environment it is also necessary to equip an agent with some motivational attitudes, such as the maintenance of goals and commitments. These things should all be recorded in the agent's mental state. So we have to find some means to express this in a precise and formal way.

3. MODAL LOGIC

As said before, for this we may resort to work in philosophical logic, and *modal logic* in particular. Modal logic is a very elegant and general framework to express the modalities as mentioned above [Che80]. It is an extension of classical logic by a special operator $\Box$ (or multiple of these) denoting some kind of 'modality'. Typical examples of modalities concern the notions of knowledge and belief, but also time, action, obligation, desires and intentions. Therefore modal logic seems an excellent tool to be used for describing and specifying intelligent agents. A 'neutral' reading of the $\Box$ operator is 'necessarily', but depending on the application it can also be read as 'it is known / believed that', 'always in the future it holds that', 'after performance of the action it holds that', 'it is obligatory / desired / intended that'. Of course, the operator will enjoy different properties for these different readings, but there are some properties that are common to all. The most prominent among these is a property called (K):

$$\Box(\varphi\rightarrow\psi) \rightarrow (\Box\varphi\rightarrow\Box\psi)$$

or equivalently

$$(\Box(\varphi\rightarrow\psi) \wedge \Box\varphi) \rightarrow \Box\psi)$$

which expresses that if is necessary that φ implies ψ, then if it is necessary that φ then it also necessary that ψ. For knowledge we additionally have the following properties:

$\Box\varphi\rightarrow\varphi$ *'knowledge is true'*
$\Box\varphi\rightarrow\Box\Box\varphi$ *'knowledge is known'*
$\neg\Box\varphi\rightarrow\Box\neg\Box\varphi$ *'ignorance is known'*

The last property is rather controversial. Philosophers generally deny this property while computer scientists normally adopt it. Although the reasons for this adoption by computer scientists of this third property, called *negative introspection* varies, one might for simplicity's sake think of situations in computer-based systems like databases where the knowledge ascribed to the system is finitary, so that in principle it becomes possible for such a system to infer that something is *not* known (see further [MH95], [FHMV95]).

For belief we have similar properties (the negative introspection is deemed even more acceptable in this case, also for philosophers!). However, it is not reasonable that beliefs are true, so the first property is replaced by a weaker one (called (D)):

$\neg\Box\bot$ *'belief is not inconsistent'*

For deontic and intentional notions like desires and goals mostly not many properties beyond the basic ones like (K) is assumed. Obligations and goals generally are taken to satisfy (D).

4. MENTAL STATES IN MODAL LOGIC: EPISTEMIC STATES ETC.

While states in traditional (imperative) programming contain information about such things as the values of programming variables, in agent-oriented programming we need to store / represent more sophisticated information. Models for modal logic are suited very well for this. These Kripke models as they are called [Kri63] consist of a set of so-called *possible worlds*, which can be taken to represent situations in which the 'world' (the environment of the system, or the system as a whole) can be. Formally this is reflected by a truth assignment function π which assigns truth-values to atomic formulas per possible world. But, most importantly, these models contain a relation R on the set of possible worlds, called the *accessibility relation* R(w, w') indicates that world w' is accessible from world w. Thus Kripke models for this basic kind of modal logic are of the form $\langle W, \pi, R\rangle$, where W is a set of possible worlds, π is a truth assignment function, and R is an accessibility relation.

It depends on the reading of the modal operator how the accessibility relation should be interpreted. For instance, for knowledge the relation yields so-called epistemic alternatives of a world: if R(w, w') holds it means that for an agent in world w the world w' might be considered an alternative. This must be viewed in the following sense: the agent considers on the basis of his/her knowledge (or rather the lack of it) the world w' as a possible state of the world. The interpretation of the modal operator is given by means of this accessibility relation: in a model M and a world w in that model it holds that the formula $\Box\varphi$ is true, if φ is true for every R-accessible world from w. More formally,

$$M, w \models \Box\varphi \text{ if and only if } M, w' \models \varphi \text{ for every } w' \text{ such that } R(w, w')$$

For knowledge the relation R is generally taken to be an equivalence relation. One can show that this choice has as a consequence that the three properties of knowledge that we gave above hold with respect to the models thus obtained. This gives rise to a natural notion of an epistemic state of an agent, viz. an equivalence class of worlds, all considered equally possible from the perspective of the agent. We can generalise this idea to mental states more in general by considering a world in a Kripke model together with the R-related worlds as the representation of the mental aspect associated with the modality under consideration. In the case of an obligation, for instance, the accessible worlds represent so-called deontic perfect alternatives in which all obligations are fulfilled (cf. [MW93]).

5. SPECIFYING AGENTS BY MEANS OF MODAL LOGIC

Now we know how we can model mental states of agents by means of the technique of modal logic we can specify the (intended / desired) behaviour of agents by means of this logic. To distinguish between the various modal operators we write the operators as K, B, D, ... for knowledge, belief, desires, ... So, for example $K\varphi$ means that φ is known. In general, we thus obtain a multi-modal logic, a modal logic with several distinct modal operators. Models for multi-modal logics have accessibility relations for each of the modal operators. Now a 'multi-modal mental state' of an agent is represented by a world and all the worlds that are accessible from this world by some of the accessibility relations.

Examples of such multi-modal logics are the well-known BDI logic proposed by Rao & Georgeff [RG91], in which the modalities belief, desire (in the form of goal) and intention are combined with a (branching-time) temporal logic. As said before, also time itself can be viewed as a modality and its associated accessibility relation is time-accessibility (the flow of time) (cf. [Eme90]). Although this idea sounds simple, the logic becomes rather technically involved, so that we will not go into any details here. Although thinking of agents in terms of BDI has also given rise to more or less concrete (BDI) architectures of agent systems, it remains a problem to relate agent programs with BDI logic. The logic is in some sense too abstract or 'ungrounded'. Another problem is that the properties of agents, although abstract in the sense that they are implementation-independent, are sometimes nevertheless regarded as 'too ideal'. The very basic property (K) of modal logic is an example of this: it states that modalities are closed under logical implication. For modalities such as knowledge, belief and certainly intentions / goals this is not very realistic. For instance, if it is intended to go to the dentist, but it is necessarily so that this will result in having toothache, closure under logical implication yields that it is then also intended to have toothache! For knowledge and belief this problem is called the *logical omniscience problem* [MH95]. By its very nature of being at the core of modal logic, it is hard to circumvent if one would like to (mostly by looking at a non-standard Kripke-style semantics, which complicates matters further, or introducing additional non-modal operators).

An attempt to remedy these problems is our own KARO logic [Lin96, LHM96, DMWK96, HLM98] where we base the attitudes of agents on a well-known logic of *action* (rather than time), viz. *dynamic logic*, which is then enriched with such notions as knowledge, belief, desires, goals and commitments. In this way we get a BDI-like logic which is founded on the agent's performance of actions rather than just the flow of time. In our view this renders a formalism that is better equipped to specify agent programs, since there is a direct and obvious relation between such programs and the actions of agents! In other words, unlike the original BDI logic, the KARO logic is grounded in the actions of agents.

6. AGENT-ORIENTED PROGRAMS AS MENTAL STATE TRANSFORMERS

In computer science and the semantics of programming languages in particular programs are often viewed as state transformers. Shoham [Sho93] extends this idea to viewing agent programs as *mental state transformers*. So, one can regard as the meaning / semantics of an agent program the way how mental states are transformed. One might think of an agent's performing of an observation resulting in a different epistemic state reflecting the update of the agent's beliefs about his/her environment. Another example is the agent's committing to do some action resulting in a change of the 'motivational / intentional' state of the agent. An elegant technical tool of treating these state transitions in a formal way is that of a *transition system* developed in the area of semantics of programming languages ([Plo81]). This idea is directly connected to the idea of agent programs as (mental) state transformers: a transition system can be used to derive transitions

$$<P, s> \rightarrow <P', s'>$$

representing a transition from state s to s' under the (partial) execution from (agent) program P to (remaining program) P'. We have used transition systems for giving formal semantics to agent-oriented languages [HBHM].

7. FROM AGENTS TO MULTI-AGENT SYSTEMS

Most researchers in the field agree that the biggest promise of agent technology lies in building and applying multi-agent systems (MAS), collections of more or less autonomous intelligent agents that are co-ordinated in some way as to be able to solve problems and perform involved tasks in a distributed manner. The applications seem to be endless: from e-commerce where agents representing companies negotiate contracts with each other to concrete multi-robotic systems used for transport and military purposes.

In order to describe such systems we need (at least) to extend our agent logics with the possibility to distinguish distinct agents. So we may use modal operators K, B, D,..., now marked with an index to indicate the agent concerned. For example $K_i\varphi$ means that agent i knows that φ, etc. Models for these logics now contain accessibility relations R_i for every agent-indexed modality $\Box_I$. But this is not all. In a multi-agent context it is also natural to consider notions of shared / distributed knowledge, common knowledge, joint intentions etc., group variants of the notions that are associated with individual agents. There has already been done quite some work on this, especially on the group notions of knowledge ([FHMV95], [MH95], [HLM99]). For example, there is the notion of *distributed knowledge* within a group. Intuitively, something is distributed knowledge within a group of agents, if by 'pooling' knowledge the group could conclude that it is the case. Formally one can define a modal operator D, where $D\varphi$ is read as φ is distributed knowledge, and gets a formal semantics by means of an accessibility relation R_D which is obtained by taking the intersection of all accessibility (knowledge) relations for the individual agents in the group: $R_D = \cap_i R_i$. An obvious property of distributed knowledge is

$$K_i\varphi \rightarrow D\varphi$$

expressing that if some agent knows φ then φ is also distributed knowledge in the group. A subject that has obtained much attention is that of *common knowledge*. Something is common knowledge if it not only the case that everyone knows it, but also everyone knows that everyone knows it, and everyone knows that again, *ad infinitum*. This notion plays an important role in game-theoretical situations but also in the analysis of communication protocols among agents where the communication channels are faulty and unreliable. A typical axiom for common knowledge, denoted by the operator C, is the following:

$$C\varphi \rightarrow EC\varphi$$

expressing that when something is common knowledge then everyone knows that it is. Here Eφ stands for 'every agent knows that φ', formally defined as (n is the number of agents):

$$E\varphi \leftrightarrow K_1\varphi \wedge \ldots \wedge K_n\varphi$$

Recently also some work has been done already on motivational attitudes in MAS, such as *obligations* in MAS (e.g. [Kro95, DL96]) and *joint intentions* of a group (e.g. [DDV99]), but much more need to be done here.

Clearly *communication* plays an important role in MAS, and indeed this topic has also attracted much attention. There are even (semi-) standard agent communication languages (ACLs) proposed like KQML [FMFM94] and FIPA-ACL [FIPA97], together with a more or less formal semantics using some kind of modal logic and speech act theory. Although in our opinion it seems a bit early for standards, this indicates the need for communication primitives for agents together with a well-defined meaning. However, also the 'formal' semantics that is provided [LF94] is not yet quite satisfactory: it still contains a lot of imprecision and moreover it seems to be too abstract from a computational point of view, cf. [EBHM99, HBHM99]), so that we have tried to give a computation-based semantics using the concept of a *transition system* as we have mentioned above. In [EBHM99] we have been able to adapt ideas of modelling communication in 'classical' concurrent programming languages as Hoare's CSP [Hoa78] to the context of agent communication where instead of sending mere values more involved information is communicated. For details we refer to the aforementioned paper.

Another problem with multi-agent systems is that concerning the *ontology* used, or rather the ontologies used, since in general for autonomous agents there is no reason why they should talk the same language. Obviously some assumptions should be made here for a fruitful communication to work. Some approaches take certain (fixed) translators into account, but in our opinion at least some part of the translations between agent languages should be made adaptable (cf. [EBHM99]).

8. CONCLUSION AND OUTLOOK

Agent technology, and especially the use of multi-agent systems, is an enormously promising area. It has many potential applications, several of which are already in advanced state of realisation (see e.g. [PAAM'97, PAAM'99]). Examples are (see [JW98]):

- industrial applications such as process control, manufacturing, air traffic control
- commercial applications such as information management, electronic commerce, business process management
- medical applications such as patient monitoring, health care
- applications for entertainment such as games, interactive theatre and cinema

For these non-trivial applications, particularly those involving human lives and huge financial consequences, it should be very desirable to have formal methods around to verify them in the sense that one can prove analytically that the agents devised behave in a desired and specified manner. It would also be very helpful if one could systematically (and perhaps even to some extent automatically) construct agent systems from formal specifications.

Although agent logics like BDI logic have helped to think about agent concepts, there are as yet no methods available to 'derive' agent programs from specifications in these logics, nor are there as yet formal ways to verify agent programs. Within the area of 'traditional' software engineering there has been done a lot of work on the development of formal methods in order to obtain correct software from formal specifications. This work has to be extended towards agent-oriented software as well, where the main challenge is to cater for the much more intricate notion of state as we have seen above. I believe that in principle it is possible to extend traditional logic-based methods to the agent-oriented paradigm employing the agent logics discussed as well as a formal semantics of agent programming languages in a style as above, but it is obvious that this will take quite some years of intensive research to come. In view of the many future applications of agent technology this will be worth the effort.

References

[Che80] B.F. Chellas, *Modal Logic. An Introduction*, Cambridge University Press, Cambridge, 1980.

[D98] G. De Giacomo, Cognitive Robotics, Papers from the 1998 AAAI fall Symposium, Orlando, Techn. Report FS-98-02, AAAI Press, Menlo Park CA, 1998.

[DDV99] F. Dignum, B. Dunin-Keplicz and R. Verbrugge, Dialogue in Team Formation, *in Proc. IJCAI'99 Workshop ABS-2 on Agent Communication Languages* (F. Dignum & B. Chaib-draa, eds.), Stockholm, 1999, pp. 39-50.

[DL96] F. Dignum and B. van Linder, Modelling Social Agents: Communication as Action, *Intelligent Agents III — Proceedings of the 3rd International Workshop on Agent Theories, Architectures, and Languages (ATAL-96)* (J. Müller, M. Wooldridge and N. Jennings, eds.), 1996, pp. 83-93.

[DMWK96] F. Dignum, J.-J.Ch. Meyer, R.J. Wieringa and R. Kuiper, A Modal Approach to Intentions, Commitments and Obligations: Intention plus Commitment yields Obligation, in: *Deontic Logic, Agency and Normative Systems* (M.A. Brown and J. Carmo, eds.), Springer-Verlag, 1996, pp. 80--97.

[EBHM99] R.M. van Eijk, F.S. de Boer, W. van der Hoek & J.-J. Ch. Meyer, Operational Semantics for Agent Communication Languages, *in Proc. IJCAI'99 Workshop ABS-2 on Agent Communication Languages* (F. Dignum & B. Chaib-draa, eds.), Stockholm, 1999, pp. 89-103.

[Eme90] E.A. Emerson, Temporal and Modal Logic, in: *Handbook of Theoretical Computer Science, Vol. B* (J. van Leeuwen, ed.), Elsevier, 1990, pp. 995-1072.

[FHMV95], R. Fagin, J.Y. Halpern, Y. Moses and M.Y. Vardi, *Reasoning about Knowledge*, MIT Press, Cambridge MA, 1995.

[FMFM94] T. Finin, D. McKay, R. Fritzson and R. McEntire, KQML: An Information and Knowledge Exchange Protocol, in: *Knowledge Building and Knowledge Sharing* (K. Fuchi and T. Yokoi, eds.), Ohmsha and IOS Press, 1994.

[FIPA97] Foundation for Intelligent Physical Agents, FIPA'97 Specification Part 2 - Agent Communication Language, October 1997.

[Har84] D. Harel, Dynamic Logic, in: *Handbook of Philosophical Logic, Vol. 2* (D.M. Gabbay and F. Guenthner, eds.), D. Reidel, Dordrecht, 1984, pp. 497-604.

[HBHM99] K.V. Hindriks, F.S. de Boer, W. van der Hoek & J.-J. Ch. Meyer, Semantics of Communicating Agents Based on Deduction and Abduction, *in Proc. IJCAI'99 Workshop ABS-2 on Agent Communication Languages* (F. Dignum & B. Chaib-draa, eds.), Stockholm, 1999, pp. 105-118.

[HBHM] K.V. Hindriks, F.S. de Boer, W. van der Hoek & J.-J. Ch. Meyer, Agent Programming in 3APL, *Int. J. of Autonomous Agents and Multi-Agent Systems*, to appear.

[Hoa78] C.A.R. Hoare, Communicating Sequential Processes, *Comm. ACM* 21(8), 1978, pp. 666-677.

[HLM98] W. van der Hoek, B. van Linder and J.-J. Ch. Meyer, An Integrated Modal Approach to Rational Agents, in: M. Wooldridge and A. Rao (eds.), *Foundations of Rational Agency*, Applied Logic Series 14, Kluwer, Dordrecht, 1998, pp. 133-168.

[HLM99] W. van der Hoek, B. van Linder and J.-J. Ch. Meyer, Group Knowledge Isn't Always Distributed (Neither Is It Always Implicit), *Mathematical Social Sciences* 38, 1999, pp. 215-240.

[HS93] J. Horty and Y. Shoham (eds.), Reasoning about Mental States: Formal Theories & Applications, Papers from the 1993 AAAI Spring Symposium, Stanford CA, Techn. Report SS-93-05, AAAI Press, 1993.

[JW98] N.R. Jennings & M.J. Wooldridge (eds.), Agent Technology, Foundations, Applications, and Markets, Springer, Berlin, 1998.

[Kri63] S. Kripke, Semantic Analysis of Modal Logic, *Zeitschrift für Mathematische Logik und Grundslagen der Mathematik* 9, 1963, pp. 67-96.

[Kro95] C. Krogh, Obligations in multiagent systems, in: *SCAI'95 —Fifth Scandinavian Conference on Artificial Intelligence* (A. Aamodt and J. Komorowski ,eds.), IOS Press, 1995, pp. 19-30.

[LF94] Y. Labrou and T. Finin, A Semantics Approach for KQML - A General Purpose Communication Language for Software Agents, in: *Proc. 3rd Int. Conf. on Information and Knowledge Management (CIKM'94),* 1994.

[LM70] I. Lakatos and A. Musgrave, *Criticism and the Growth of Knowledge*, Cambridge Univ. Press, Cambridge, 1970.

[Lin96] B. van Linder, Modal Logics for Rational Agents, Ph.D. Thesis, Utrecht University, 1996.

[LHM96] B. van Linder, W. van der Hoek and J.-J. Ch. Meyer, Formalising Motivational Attitudes of Agents: On Preferences, Goals and Commitments, *Intelligent Agents Volume II — Agent Theories, Architectures, and Languages* (M. Wooldridge, J.P. Müller and M. Tambe, eds.), LNAI 1037, Springer-Verlag, 1996, pp. 17-32.

[Mae95] P. Maes, Intelligent Software, *Scientific American* 273(3) (Special Issue on Key Technologies for the 21st Century), 1995, pp. 66-68.

[MH95] J.-J. Ch. Meyer and W. van der Hoek, *Epistemic Logic for AI and Computer Science*, Cambridge University Press, 1995.

[MW93] J.-J. Ch. Meyer and R.J. Wieringa, Deontic Logic: A Concise Overview, in: *Deontic Logic in Computer Science* (J.-J. Ch. Meyer and R.J. Wieringa, eds.), John Wiley & Sons, 1993, pp. 3-16.

[PAAM'97] *Proceedings of the First Int. Conf. on the Practical Applications of Intelligent Agents and Multi-Agent Systems*, London, 1997.

[PAAM'99] *Proceedings of the Fourth Int. Conf. on the Practical Applications of Intelligent Agents and Multi-Agent Technology,* London, 1999.

[Plo81] G. Plotkin, A Structural Approach to Operational Semantics, Techn. Report DAIME FN-19, Aarhus University, 1981.

[RG91] A.S. Rao and M.P. Georgeff, Modeling rational agents within a BDI-architecture, in: *Proc. 2nd Int. Conf. on Principles of Knowledge Representation and Reasoning (KR'91)* (J. Allen, R. Fikes and E. Sandewall, eds.), Morgan Kaufmann, 1991, pp. 473-484.

[Shoh91] Y. Shoham, Implementing the Intentional Stance, in: *Philosophy and AI: Essays at the Interface* (R. Cummins and J. Pollock, eds.), MIT Press, Cambridge, MA, 1991.

[Sho93] Y. Shoham, Agent-Oriented Programming, *Artificial Intelligence* 60, 1993, pp. 51-92.

[WJ94] M. Wooldridge and N.R. Jennings, Intelligent Agents: Theory and Practice, *The Knowledge Engineering Review* 10(2), 1995, pp. 115-152.