Lecture Notes in Artificial Intelligence 1193

Subseries of Lecture Notes in Computer Science
Edited by J. G. Carbonell and J. Siekmann

Lecture Notes in Computer Science
Edited by G. Goos, J. Hartmanis and J. van Leeuwen

Jörg P. Müller Michael J. Wooldridge
Nicholas R. Jennings (Eds.)

Intelligent Agents III

Agent Theories, Architectures, and Languages

ECAI'96 Workshop (ATAL)
Budapest, Hungary, August 12-13, 1996
Proceedings

Springer

Series Editors
Jaime G. Carbonell, Carnegie Mellon University, Pittsburgh, PA, USA
Jörg Siekmann, University of Saarland, Saarbrücken, Germany

Volume Editors

Jörg P. Müller
Michael J. Wooldridge
Mitsubishi Electric Digital Library Group
103 New Oxford Street, WC1A 1EB London, United Kingdom
E-mail: (jpm/mjw)@dlib.com

Nicholas R. Jennings
Distributed AI Unit, Queen Mary and Westfield College, University of London
Mile End Road, E1 4NS London, United Kingdom
E-mail: n.r.jennings@qmw.ac.uk

Cataloging-in-Publication Data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Intelligent agents III : agent theories, architectures, and languages ; proceedings / ECAI '96 workshop (ATAL), Budapest, Hungary, August 12 - 13, 1996. Jörg P. Müller ... (ed.). - Berlin ; Heidelberg ; New York ; Barcelona ; Budapest ; Hong Kong ; London ; Milan ; Paris ; Santa Clara ; Singapore ; Tokyo : Springer, 1997
(Lecture notes in computer science ; Vol. 1193 : Lecture notes in artificial intelligence)
ISBN 3-540-62507-0
NE: Müller, Jörg P. [Hrsg.]; ECAI <12, 1996, Budapest>; ATAL <3, 1996, Budapest>; GT

CR Subject Classification (1991): I.2, D.2, C.2.4

ISBN 3-540-62507-0 Springer-Verlag Berlin Heidelberg New York

Typesetting: Camera ready by author
SPIN 10549836 06/3142 – 5 4 3 2 1 0 Printed on acid-free paper

Preface

Intelligent agents are computer systems that are capable of flexible autonomous action in dynamic, typically multi-agent domains. Over the past few years, the computer science community has begun to recognise that the technology of intelligent agents provides the key to solving a range of complex software application problems, for which traditional software engineering tools and techniques offer no solution. The aim of the *Agent Theories, Architectures, and Languages* (ATAL) workshop series is to bring together researchers interested in this new technology. The ATAL workshops address the issues of agent specification via agent theories, the ability of agents to make decisions in time-constrained environments, and software tools for building, experimenting with, and evaluating agents. In particular, the workshops focus on the link between theories of agents and the realization of such theories using software architectures or languages.

Some 56 papers were submitted to the ATAL-96 workshop, from twenty different countries: nine from the UK, six from the USA, five each from France, Germany, and Portugal, four from the Netherlands, three each from Australia, Italy, and Spain, two each from Finland and Poland, and one each from Brasil, Canada, Hungary, Japan, India, Korea, Romania, Switzerland, and Taiwan. After 32 submissions to ATAL-94 and 54 submissions to ATAL-95, the further increase for ATAL-96 — as a primarily European event — reflects the continued growth of interest in agents on the part of the international computer science community. After reviewing, 24 papers were accepted for presentation; revised versions of these papers are published herein.

November 1996

Jörg P. Müller,
Michael Wooldridge, and
Nicholas R. Jennings

Workshop Organization

Organizing Committee

Jörg P. Müller (CHAIR)	(Mitsubishi Electric Digital Library Group, UK)
Michael Wooldridge	(Mitsubishi Electric Digital Library Group, UK)
Nicholas R. Jennings	(Queen Mary & Westfield College, UK)

Program Committee

Hans-Dieter Burkhard	(Germany)	Cristiano Castelfranchi	(Italy)
Paul Cohen	(USA)	Phil Cohen	(USA)
Ed Durfee	(USA)	Tim Finin	(USA)
Klaus Fischer	(Germany)	Michael Fisher	(UK)
John Fox	(UK)	Fausto Giunchiglia	(Italy)
Piotr Gmytrasiewicz	(USA)	Hans Haugeneder	(USA)
Sarit Kraus	(Israel)	John-Jules Ch. Meyer	(The Netherlands)
Anand Rao	(Australia)	Jeff Rosenschein	(Israel)
Yoav Shoham	(USA)	Munindar Singh	(USA)
Aaron Sloman	(UK)	Donald Steiner	(Germany)
Kurt Sundermeyer	(Germany)	Milind Tambe	(USA)
Jan Treur	(The Netherlands)	Gerd Wagner	(Germany)

Additional Reviewers

Alastair Burt	Rino Falcone	Peyman Faratin	Catholijn Jonker
Christoph G. Jung	Lisa Hogg	Marco Mulder	Susanne Kalenka
Michael Kolb	Bernd van Linder	Stephen Moon	Tim Norman
Simon Parsons	Luciano Serafini	Phillip Turner	Gero Vierke

Acknowledgments

The organization of this workshop was made possible by financial support from the German Research Center for Artificial Intelligence (DFKI GmbH).

Contents

Part V: Architectures

Part VI: Languages

Introduction

Only three years ago, when the first workshop on Agent Theories, Architectures, and Languages (ATAL) was held at ECAI-94 in Amsterdam, most interest in intelligent agents came from academic researchers with a background in artificial intelligence. Today, agent technology is rapidly migrating from research laboratories to commerce and industry, and is the object of study in many non-AI computing disciplines. Perhaps the most significant indication that agents are here to stay is that the technology is increasingly being adopted by software engineering practitioners.

The objective of the *Intelligent Agents* series is to document this fast development, by providing a carefully selected set of papers that reflect the current trends and research topics in the area. The content of these papers ranges from the theoretical foundations of agency through control architectures and agent languages to tools and methodologies that allow software engineers and agent designers to cope with complex applications.

This book, the third in the *Intelligent Agents* series, is structured into six sections, which can be logically grouped in two parts: The first part comprises Sections I–III; it addresses a number of current topics that interconnect the agent-related research areas of theories, architectures, and languages: *agent methodologies* (Section I), *agent definitions and taxonomies* (Section II), and *the role of theory in agent-based systems*. The second part (Sections IV–VI) provides collections of high-quality technical papers on agent *theories* (Section IV), *architectures* (Section V), and *languages* (Section VI).

Section I: Design Methodologies

The first section deals with an area of increasing significance: tools and methodologies that support the design of agents and agent-based systems. It consists of a paper by David Kinny and Mike Georgeff in which they present the methodology that is employed at the Australian Artificial Intelligence Institute (AAII) to build industrial agent-based systems. This paper is based on an invited presentation of the authors at the ATAL-96 workshop. Given the considerable advances in agent theories, architectures, and languages, the development of useful methodologies is expected to become one of the hottest topics in agent research. In the ATAL-97 Workshop there will be a special track devoted to agent design methodologies.

Section II: What is an Agent? — Definitions and Taxonomies

There is an ongoing debate about the essence and the boundaries of agenthood in the agent community. Many researchers feel that agreement on the basic characteristics of agency is a crucial prerequisite to scientific advance. This section of the book summarizes a panel session under the title *What is an agent?* held at the ATAL-96 workshop. It contains a paper by Stan Franklin and Art Graesser in which they compare several known agent definitions and come up with what they call the beginnings of an agent taxonomy. This article is followed by a summary of panel statements on the subject.

Section III: The Role of Theory in Building Agent-Based Systems

There is a considerable gap between theoretical agent research on the one hand and practical, often *ad hoc* approaches that are used in most attempts to actually build agent-based systems. The question of the necessity of narrowing this gap and of how this can be accomplished is the subject of the third section of this book. Like Section II, this section is based on a panel dicussion held at the 1996 workshop. It starts with a contribution by Michael Luck, Nathan Griffiths, and Mark d'Inverno that presents a case study demonstrating the design of an agent simulation environment that is strongly based on a formal agent theory. Subsequently, the reader will find summaries of the panel responses.

Section IV: Theories

This section is made up of four blocks: the first four papers by Lomuscio and Colombetti, Moreno and Sales, Dragoni and Giorgini, and Benerecetti et al, deal with different aspects of the formal representation of beliefs. The next two papers (Sandu, and Brazier et al) address important motivational and intentional concepts for the design of multi-agent systems. The third block in this section holds four papers describing agents involved in dialogues, authored by Traum, Noriega and Sierra, Bretier and Sadek, and Dignum and van Linder.

Lomuscio and Colombetti — quantified logic for belief: Logics of knowledge and belief are a key component of agent theory: they are essential for representing the information that agents have available to them. The most important tradition in the logical analysis of knowledge and belief is *possible worlds* semantics, in which an agent's information state is characterized as a set of possible states of affairs. An agent is said to know/believe a statement if that statement is true in all the states of affairs. While there are many desirable aspects of the possible worlds tradition, there are also many unsolved problems. *Quantified* logics of knowledge and belief have long been recognised as a difficult area. Lomuscio and Colombetti, working within the possible worlds tradition, investigate one particular aspect of quantification: that of quantifying over the agents in a belief formula. Specifically, they develop a logic QLB, which permits formulae of the form $B_t\varphi$, where B is a modal belief connective, φ is a statement, and t is an arbitrary term generally denoting an agent. They present an axiomatization of QLB, as an extension to the normal modal logic KD45n, and investigate the soundness of variations of the well-known Barcan formulae.

Moreno and Sales — dynamic belief analysis: Moreno and Sales work within the same tradition of normal modal logics for belief as Lomuscio and Colombetti (above). However, they address a slightly different, but very well-known problem with normal modal logic: that of *logical omniscience*. Normal modal logics predict that agents believe all the logical consequences of their beliefs. While humans generally believe *some* of the logical consequences of their beliefs, they never believe *all* of them. This makes normal modal logics unsuitable for reasoning about any real agent. Moreno and Sales adapt the normal modal logic semantics to counter this problem. They do this by allowing worlds (which are normally considered to be complete, consistent, deductively closed descriptions of the world) to be incomplete and possibly inconsistent. They show how this approach avoids the key problems of logical omniscience, and present a tableaux-based proof method for reasoning within their framework.

Dragoni and Giorgini — multi-agent belief revision: Belief revision is about how best to modify beliefs in the face of new information. Dragoni and Giorgini address the issue of performing this difficult process in a multi-agent environment, where there are many sources of new information. They make use of the *belief-function* formalism. This formalism supports conveying the source of a piece of information together with the information itself, and thus makes it possible to establish a relationship between the reliability of a source and its credibility.

Benerecetti et al. — context based specification: This work shows how two seemingly different schemes for formalizing belief are in fact complementary. The first scheme investigates the properties of contexts for formally specifying mutual belief between the agents and the second scheme considers the use of contexts for representing propositional attitudes. In this paper, the authors bring together these two strands of work: they investigate the formal specification of multi-agent systems using the multi-context framework.

Sandu — collective goals: The use of temporal/modal logics for the formal analysis of collective goals and intentions is an area of much ongoing work in multi-agent systems. Gabriel Sandu contributes to this area by extending the "standard" possible worlds semantics for knowledge with operators for representing the goals of agents and the performance of actions by agents. He then extends the framework further, by introducing operators for characterizing the *collective* goals of agents. He then shows how this framework satisfies various desirable properties of mutual goals and intentions.

Brazier et al. — joint intentions: Like Gabriel Sandu (above), Brazier and colleagues are concerned with the formalization of collective attitudes. They formalize a variation of the cooperation model proposed by Jennings using DESIRE, a framework for the formal compositional specification of complex reasoning systems. To illustrate their approach, they use as a case study a simple cooperative problem: the concurrent design of an aircraft interior.

Traum — reactive-deliberative dialogue agency: David Traum describes an agent theory that supports the design of agents that are capable of engaging in natural language dialogue. The model is driven by the observation that dialogue is an inherently reactive process. In particular, this theory comprises the modeling of social attitudes, i.e., mutual beliefs, obligations, and multi-agent plan execution. Traum illustrates his model through an example of an implemented social agent, i.e., the dialogue managing agent in TRAINS-93, an integrated natural language conversation and plan reasoning system.

Noriega and Sierra — layered dialogical agents: This paper describes an approach to unify different mental attitudes involved in the design of rational agents by a layered architecture of agent models. Individual communicational, informational, and motivational attitudes are described by theories. They form the layers of the model. The mappings between the layers are achieved by means of so-called *bridge rules*. The language chosen by the authors to model multi-agent systems is Concurrent Descriptive Dynamic Logic. The model is illustrated by the example of the Spanish fish market, where different types of agents, i.e., auctioneers, admitters, and buyers, engage in a bidding process.

Bretier and Sadek — ARTIMIS: This paper describes a formal theory of rational interaction. The theory expresses principles of rational behavior, communication, and cooperation in a uniform logical framework. The theory has been implemented using an inference engine based on quantified KD45 modal logics, which provides modalities representing concepts such as uncertainty, intention, and action. The authors describe the use of their implemented theory of interaction in the ARTIMIS system, which models the rational unit in a prototype of a cooperative spoken dialogue system. This system, which is hosted by France Télécom, is a directory service that is able to recognize and to answer natural language queries in real time.

Dignum and van Linder — modeling social agents: A considerable amount of research work on autonomous agents has been invested in developing logical formal models of informational and motivational attitudes such as beliefs and goals, and their connection to agent abilities (e.g., actions). Dignum and van Linder extend this research by integrating social components into such a framework. In order to cater for communication between agents, the authors employ four types of *speech acts*, i.e., commitments, directions, declarations, and assertions. These speech acts are formalized as meta-actions using a Kripke-style possible worlds model. The framework supports the modeling of relationships among agents such as obligations. In contrast to the approaches of Traum and Sadek and Bretier, that focus on natural language human-machine dialogue, the model given in this paper is geared to inter-agent communication, where it is sufficient to consider a formal language.

Section V: Architectures

The agent architecture section in this book subsumes seven papers that can be classified in three categories. The first three papers describe individual capabilities of an agent, such as learning (the contributions by Kraines and Kraines, and by Zeng and Sycara), shift of attention (Botelho and Coelho) and coordination (Müller), as well as aspects of their integration and operationalization in an agent. In addition to covering issues related to the architectural design of individual agents, the next two papers, i.e., those by Norman et al. and Schroeder et al. describe architectures of agent-based application system. Finally, Davis's work focuses on the integration of different agent capabilities in a uniform control framework.

Kraines and Kraines — threshold of cooperation: The investigation of cooperation among agents has had a long tradition in Distributed Artificial Intelligence. Kraines and Kraines's paper starts from the question why selfish behavior is relatively rare in animal societies, whereas it seems to be almost impossible to eliminate in human societies. The hypothesis set up in this paper is that cooperation tends to be less stable in societies of rapidly learning agents than it is in the case of agents that adapt more slowly. This hypothesis is backed up by a Markov chain model and by a series of experiments with stochastically learning agents in a simple cooperative game, the *Stag Hunt*. The most striking result of the paper is that the authors identify what they call the threshold of cooperation, i.e., a learning rate above which cooperation in a society declines rapidly.

Zeng and Sycara — BAZAAR: Automated negotiation is gaining increased importance in the context of the growing interest in interacting software agents and electronic commerce. The authors address the topic of agent learning in the context of negotiation. They describe BAZAAR, a sequential decision-making model of negotiation. Agents in BAZAAR employ Bayesian learning techniques in order to modify their knowledge about the parameters of the environment (e.g., interest rate, overall product supply and demand) and about object- and meta-level beliefs of other agents. The model is illustrated by a buyer–supplier negotiation example.

Müller — cooperation model for autonomous agents: This paper describes a cooperation model and its integration into an operational agent architecture. It is based on the game-theoretic negotiation model by Rosenschein and Zlotkin. The basic constructs in Rosenschein and Zlotkin's model, i.e., negotiation protocols and strategies, form the application-independent meta-level in Müller's cooperation model. It is complemented by a theory of joint plans to describe one specific instance of the object level, i.e., the negotiation set. The implementation of the cooperation model in the cooperative planning layer of the INTERRAP agent architecture is described.

Botelho and Coelho — emotion based attention shift: Autonomous agents that are situated in a particular environment have to attain a balance between performing cognitive tasks and examining their environment for potentially relevant changes. As agents are resource bounded, they cannot continually interrupt their reasoning for each and every external event, otherwise they would never achieve any goals. Hence an attention shift mechanism needs to be designed which enables the agent to pursue tasks until relevant or important changes occur. This paper describes and examines the properties of three such mechanisms: activation-based attention shift, attention shift by event driven emotion, and attention shift by anticipation driven emotion.

Norman et al. — architecture for business process management: Norman *et al.* describe a service-oriented architecture for autonomous agents. Agents are conceptualized in terms of the services that they can offer. As agents are autonomous, if they require a service from another agent they must negotiate for it. The multi-agent community in which the individual agents are situated can be structured as a series of peer agents or as loose confederations of related agents. The architecture and system are demonstrated using the domain of business process management.

Schroeder et al. — diagnosis agent based on logic programming: The authors describe the formal specification and implementation of a diagnostic agent using principles from logic programming. An agent architecture is developed and applied to the domain of distributed diagnosis. The architecture is hybrid in that it consists of a deliberative layer with a knowledge base and an inference engine and a reactive layer for communication and control. The development of the architecture also aims to show how logic and logic programming is a suitable paradigm for specifying and implementing multi-agent systems.

Davis — reactive and motivational agents: This work is part of an on-going study whose ultimate aim is to understand and develop complete agent architectures. The study encompasses both natural and artificial agents. In particular, Davis examines how a number of different agent types (with different architectures and capabilities) can be combined into agent systems with various properties. The ensuing system is demonstrated in a creche scenario, where a minder agent is responsible for looking after a number of babies (minibots).

Section VI: Languages

This section describes four very different languages which developers could use to program their agents. The languages differ in their scope, their basic paradigm, and their domain of demonstration. As regards scope, some of them aim at describing the whole agent (e.g., Poggi and Adorni's paper), whereas others focus on specific aspects of the agent's functionality (e.g., the cooperation language developed by Barbuceanu and Fox). With respect to the underlying paradigm, a distinction can be made between logic-based (Wooldridge), object-oriented (Poggi and Adorni, Barbuceanu and Fox), and knowledge-based approaches (Li and Pereira). Finally, different application domains were used to demonstrate the usefulness of the languages, e.g., robotics (Poggi and Adorni) and supply chain management (Barbuceanu and Fox).

Poggi and Adorni — HOMAGE: HOMAGE is an environment that integrates agent and object programming paradigms so that multi-agent applications can be programmed more easily. HOMAGE has two different programming levels. Firstly, there is an object level which allows object-oriented languages (presently, C++, Common LISP, and JAVA) to be used to develop the agent's domain problem solving capabilites. Secondly, there is an agent level which allows the agent control center to be specified. This level also facilitates the development of multi-agent systems by distributing and inter-connecting agent instances. Concepts from the two levels are illustrated in the context of a multiple robot scenario.

Barbuceanu and Fox — coordination language for agent systems: The focus of this paper is in defining a language which can be used for specifying coordination activities in multi-agent systems. Coordination is conceptualized as a structured conversation among agents. Thus the language provides KQML-based communication, support for describing interactions as multiple structured conversations, rule-based approaches for selecting appropriate conversations, and conversation execution tools. In addition to the development facilities, support tools for coordination level knowledge acquisition and debugging are analyzed and detailed.

Wooldridge — semantics for Concurrent METATEM: Concurrent METATEM is a multi-agent programming language, in which agents are programmed by giving them a temporal logic specification of the behavior it is intended they should exhibit. An agent's specification is directly executed in order to generate its behavior. Wooldridge addresses the issue of giving a semantics to Concurrent METATEM. He first gives a set of conditions that define what constitutes an acceptable *run* of a Concurrent METATEM system. The semantics of a particular Concurrent METATEM system is then its set of runs. In order to conveniently express properties of a system, Wooldridge introduces a temporal logic of knowledge, and shows how certain axioms of the logic correspond to properties of Concurrent METATEM. The axioms can then be used to reason about Concurrent METATEM systems.

Li and Pereira — knowledge based situated agents: This paper addresses some of the issues involved in building agents who have to operate in dynamic and evolving environments. The agent architecture is a hybrid system — having a set of pre-compiled rules for known situations and a deliberative planner for unpredicted situations. An agent specification language (SICSLOG) defines the agent's initial knowledge base, its atomic actions, its initial intentions, and the predefined plans. The architecture and language are demonstrated by specifying the actions of a commuter traveling between two cities to get to work.

November 1996

Jörg P. Müller
Michael Wooldridge
Nicholas R. Jennings