

A Characterization of Sapient Agents

Martijn van Otterlo
TKI, Dept. of Computer Science,
University of Twente, The Netherlands
e-mail: otterlo@cs.utwente.nl

Marco Wiering, Mehdi Dastani and John-Jules Meyer
Intelligent Systems Group,
Utrecht University, The Netherlands
e-mail: {marco,mehdi,jj}@cs.uu.nl

Abstract—This paper presents a proposal to characterize *Sapient Agents* in terms of cognitive concepts and abilities. In particular, a sapient agent is considered as a cognitive agent that learns its cognitive state and capabilities through experience. This characterization is based on formal concepts such as beliefs, goals, plans and reasoning rules, and formal techniques such as relational RL. We identify several aspects of cognitive agents that can be evolved through learning and indicate how these aspects can be learned. Other important features such as the social environment, interaction with other agents or humans and the ability to deal with emotions, will also be discussed. The paper ends with directions for further research on sapient agents.

I. INTRODUCTION

Intelligent agents have already found their way to the public, and during the last decades many intelligent agents have been shown to be effective in solving particular tasks. However, an intelligent agent is often used for a single task only, e.g., think about a chess playing program which is only used to play chess. If an agent has to fulfil multiple tasks, more complicated issues arise, such as a decision method for choosing the current goals based on current information about the environment and refining the decision method based on learning capability. Making the transition from intelligent agent to agents that decide autonomously and learn to refine their decision making capability, requires some new type of agent. This type of agent will be referred to as a *sapient agent*.

We will look upon sapient agents, from the starting perspective of cognitive agents extended with (relational reinforcement) learning capabilities. A cognitive agent is assumed to have some internal state consisting of mental attitudes such as beliefs, goals, and plans, receives inputs through its sensors, and performs actions. The actions are decided on based on its mental state in such a way that its effort to attain a goal will be minimal. For this many reasoning mechanisms could be useful such as logical deduction, neural networks, fuzzy logic, Bayesian networks etc. We consider an agent in which all of these mechanisms may run in parallel. E.g. pattern recognition may be done using neural networks, whereas communication is best done using logical languages.

In this paper we will first characterize sapient agents, and then describe how we can use learning methods for sapient agents.

In section II we define sapient agents starting from the notion of a cognitive agent. We discuss a general cognitive architecture, the deliberation cycle and possible influences stemming from emotions. In section IV we will describe the reinforcement learning paradigm, as well as recent extensions important for sapient agents. Section V will discuss the learning opportunities in the cognitive architecture defined in section III-A and proposes some solutions in the framework of reinforcement learning. Section VI will discuss the broader context of sapient agents; the interaction with other agents and communication and social issues, with both humans and agents. In section VII we will reflect on the characterization given in this paper and give directions for further research.

II. DEFINING SAPIENT AGENTS

Sapient agents are assumed to have accumulated learning and knowledge, the ability to discern inner qualities and relationships, often called the agent's *insight*, and good sense or *judgments*. These concepts and properties are, however, not intuitive and informal without explicit formal semantics. In this paper, we consider sapient agents as a specific type of cognitive agents for which many formalizations are proposed. In particular, we believe that properties such as knowledge, insight, and judgments of sapient agents are related to, and should be defined in terms of, mentalistic concepts such as beliefs, goals, and plans as used for cognitive agents. Therefore, we propose an interpretation of properties of sapient agents based on mentalistic concepts of cognitive agents and identify certain problems such as the integration of learning and decision making processes that together influence the behavior of sapient agents.

We assume that insight and judgment properties of sapient agents determine their course of actions. For cognitive agents the course of actions can be specified in terms of their mental attitudes which contain at least beliefs, goals, and norms, capabilities such as actions and plans, reasoning rules that can be used to reason about the mental attitudes, communication, and sensing. Given the above mentioned entities, the decision making ability of agents can be considered as consisting of reasoning about mentalistic attitudes, selecting goals, planning goals, selecting and executing plans, etc[1].

In our view, the judgment of an agent can be considered at the lowest level as making choices about how to reason about its mental attitudes at each moment in time. For example, an agent's judgment can be established by reasoning about

its goals or by reasoning about its goals only when they are not reachable using any possible plan. Some more moderate alternatives are also possible. E.g. the agent can create a plan for a goal and execute the plan. If this leads to a stage where the plan cannot be executed any further, then the agent can start reasoning about the plan and revise the plan if necessary. If the goal can still not be reached, then the agent can revise the goal. So, this leads to a strategy where one plan is tried completely and if it fails the goal is revised or even abandoned. In general, an agent with the judgment ability should be able to control the relation between plans and goals. For example, an agent should control whether a goal still exists during the execution of the plan to reach that goal. If the corresponding goal of a plan is reached (or dropped), the agent can allow or avoid continuing with the plan.

We consider the insight of agents to be directly related to the ability of agents to evaluate their mental states and mental capabilities. We therefore assume that the insight ability is the ability to learn how to reason about mental attitudes and thus how to make decisions. The reasoning capability determines the agent's decision making behavior and the learning capability determines the evolution of the reasoning capability through experiences. The focus of this paper with respect to the agent's insight is on the aspects of the agent's mental state and mental capabilities that are influenced by the learning process. These could be the goals, beliefs, desires, reasoning rules of even basic capabilities (the agent can learn new actions).

III. COGNITIVE AGENTS

In this section, we consider various aspects of the mental states and mental abilities of cognitive agents that may evolve through learning and from experiences resulting in properties associated with the sapient agents. In general, cognitive agents are assumed to have mental states consisting of mental attitudes such as beliefs, goals, plans, and reasoning rules [2], [3], [4], [5], [6], [7]. For example, a cognitive agent may believe there is no coffee available, desire to drink coffee, and desire to have tea if there is no coffee. Moreover, the behavior of cognitive agents, i.e. the actions it chooses and performs, are assumed to be determined by deliberating on the mental attitudes [2], [1], [8]. The deliberation process is a continuous and iterative process that involves many choices and decisions through which actions are selected and performed. For example, a deliberation process may select one goal, plan the goal and execute the plan. If the goal cannot be planned, it may either drop the goal or revise it. The revised goal may be planned. It is also possible that a plan cannot be executed since some of its constituting actions are blocked. In such a case, the agent may either decide to drop the plan or revise it. The existing proposals of cognitive agents [4], [5], [1], [6], [7] assume that many of these choices and decisions are fixed. These choices and decisions are based on predefined criteria and remain unchanged during agent's lifetime. In this section, we introduce a general architecture

for cognitive agents and discuss possible choices and decisions that are involved in the deliberation process.

A. A Cognitive Agent Architecture

In this section, we consider a general architecture for cognitive agent consisting of the representation of mental attitudes and the deliberation process. This agent architecture is illustrated in figure 1. According to this architecture, an agent observes the environment and communicates with other agents. The observation of an agent provides the facts that the agent recognizes from its sensory information. These facts can be used to update the agent's mental state. The communication provides information that an agent receives from other agents. The received information are messages that are stored in the message box (Mesg.) of the agent. These messages can be represented in terms of the identifier of the sender and receiver, a logical sentence that determines the content of the message, and a performative which indicates the modality of the message, i.e. whether the content is meant to inform the receiver, contains requests for the receiver, etc.

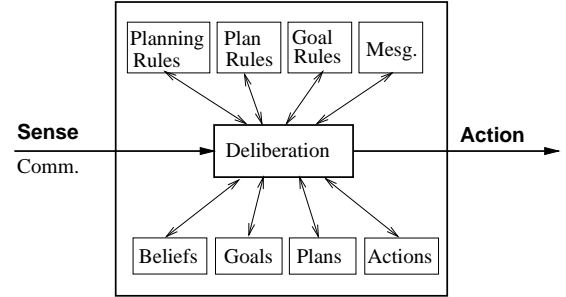


Fig. 1. A general architecture for cognitive agents.

The beliefs of an agent represent its general world knowledge as well as its knowledge about the surrounding environment. The beliefs are usually represented by sentences of a logical language, e.g. sentences of a first order predicate language. The goals represent the states that the agent desires to reach. Like beliefs, goals are represented by sentences of a logical language as well. Actions represent basic capabilities that agent can perform. These actions can be cognitive actions such as belief updates, or external (physical) actions such as communication or movement actions. The actions are usually specified by pre- and post-conditions which are belief formulae. The plans represent structured patterns of actions that agent can perform together.

B. Acting, Planning and Deliberating

A planning rule expresses that a goal can be achieved by performing a plan under a certain belief condition. A planning rule has the form $\phi \leftarrow \beta \mid \pi$, which indicates that goal ϕ can be achieved by plan π if belief condition β holds. A goal rule determines how to modify a goal under a certain belief condition. A goal rule has the form $\phi \leftarrow \beta \mid \psi$ which indicates that goal ϕ can be revised as goal ψ if belief condition β holds. Likewise, a plan rule determines how to modify a plan under a certain belief condition. A plan rule has the form $\pi \leftarrow \beta \mid \pi'$

which indicates that plan π can be revised as plan π' if belief condition β holds.

Cognitive agents deliberate on these concepts to decide which actions to perform at each moment of time [9]. The deliberation process involves many activities such as applying a reasoning rule for the above mentioned purposes, selecting a goal to achieve, selecting a plan to execute, generating a plan to achieve a goal, etc. In particular, a cognitive agent decides at each moment of time which activity to perform. It should be noted that different applications require different deliberation processes and that there is not one single universal deliberation process. An example of a deliberation process is the following iterative procedure:

```
repeat
  Find and apply goal rules that are applicable
  Find and apply plan rules that are applicable
  Find a goal and a planning rule which is applicable to it
  Apply the selected planning rule to the selected goal
  Find and execute a plan
end repeat
```

In order to specify, design, and implement a cognitive agent one needs to initialize its cognitive state and specify, design, and implement various decisions and choices involved in the deliberation process beforehand. For example, the agent designer should develop beforehand various selection functions to select goals, plans, and various types of rules at various stages of the deliberation process. Also, the agent designer should indicate beforehand how goals and plan are generated and dropped. For many types of agents, especially sapient ones, it is not possible, or even desirable, to specify all these concepts at design-time. Therefore, sapient agents should be capable of learning.

C. Emotions

Emotions will also be important for truly sapient agents. Emotional attitudes towards agents, objects, events etc. can become important in the process of acting, planning and deliberation. Emotions motivate and bias behavior, but they do not completely determine it. They play a *reflective* role in decision making and learning [10], may monitor planning and may be *prospect-based* [11]. By focusing on *emotion-inducing* events, the agent can decide more effectively. Basic emotions such as fear can trigger behavior needed to act fast, or to quickly change plans. Emotions such as happiness can influence choices for certain goals or plans. In some sense, emotions complement *ratio* so that the agent becomes wiser, more sapient.

It is acknowledged that, at least in humans, emotions are not a separate process from cognition, but both are inextricably intertwined. It can even be stated that without emotions, decision-making and acting is hardly possible and that reason itself uses emotions to guide its decision making processes [12]. Even though some may argue that it is not important for machines (agents) to actually *have* emotions, it surely is important to be able to *reason about* emotions. Especially in

situations in which natural language understanding and co-operative problem solving are important. When interaction with humans is involved, a capability to deal with emotions, whether to express or to understand, becomes highly desirable or even needed. [10].

IV. LEARNING

There is general agreement nowadays that intelligent agents should be *adaptive*, i.e. capable of learning. For learning to work, agents should be able to make the proper generalizations to reuse learned knowledge to apply it to new situations similar to encountered ones. For sapient agents learning as a capability should be extended to learning how to organize the deliberation cycle. Sapient agents can learn how to solve multiple tasks in parallel, how to deal with multiple goals and also how to set the right priorities. They can use their own experiences but also the social context for doing this.

There are roughly three learning paradigms. *Unsupervised* and *supervised* learning are used mainly for isolated *clustering* and *classification* tasks, respectively. However, for agents, the *reinforcement learning* paradigm is dominant, given the fact that it deals with *behavior* learning. In the following sections we will describe its main features and discuss extensions useful for learning within (cognitive) sapient agents.

A. Reinforcement Learning

In reinforcement learning (RL) [13], an agent learns how to behave by interacting with its environment (including other agents) using a trial and error process. The agent has to learn a policy to decide on actions based on its mental state in such a way that its cumulative intake of rewards will be maximized. In general, RL methods learn estimates of utility values for certain belief states and certain actions. These values can be used to determine optimal actions for the agent's current state. We can see an agent with RL capabilities as an agent that tries to find out which goal to select and how to achieve the selected goal with minimal efforts. RL has already been successfully applied to learn to play the game of backgammon at human world class level [14].

B. Abstraction in Reinforcement Learning

Although RL is a general method for behavior learning, standard RL is not powerful enough for the rich knowledge structures and capabilities of sapient agents. Recently a number of extensions to the RL framework have been developed that deal with various kinds of higher-order *abstractions*. Abstractions *over actions* (or time, i.e. *temporal* abstraction) can be used to abstract over different ways an abstract action can be instantiated. For example, an action *moveTo(room1)* can abstract over a number of motor-actions actually needed for a robot to move to *room1*. Whole action sequences can be abstracted into a *plan* or a *macro-action*. Methods that use abstraction over time or action sequences are termed *hierarchical* RL methods [15], [16], [17]. Hierarchies of actions and behaviors can be defined or even be learned.

Another recent direction in RL involves abstraction by using more powerful representation languages. Quite naturally, cognitive concepts like beliefs about the world and goals are expressed in terms of *objects* and *relations*. Traditionally, RL has been using feature-based and propositional representations for representing cognitive concepts and actions, although for logic-based agents, richer representational formalisms are needed. Recently progress has been made in closing the gap between logic-based agent formalisms and methods for learning behavior such as RL.

On the one hand, formalisms such as the Situation Calculus [18] have been extended with means to calculate values for actions and mental states [19]. These values enable the agent to choose rationally between goals and actions. On the other hand, progress has been made in upgrading RL methods towards richer representational formalisms. *Relational Reinforcement Learning* ([20], [21], [22]) methods learn values for relational expressions over mental concepts and actions. By means of logical induction, useful concepts that are important for optimal behavior of the agent, can be learned from experience. These concepts are expressed in terms of knowledge the agent has, and based on actual experience gained in performing different actions.

By integrating value-based behavior learning methods such as RL, the notion of a logic-based cognitive agent, and concept induction methods, behavior learning of an agent can be directly connected to cognitive notions, represented in terms of *objects* and *relations* [23].

V. LEARNING IN COGNITIVE AGENTS

In section III-B we mentioned the fact that in cognitive agent architectures, many aspects are specified beforehand. However, for a sapient agent we believe that various choices and decisions involved in the deliberation process should be learned through experience instead of being fixed and defined beforehand.

A. Adapting the Deliberation Cycle

There are many opportunities for learning in cognitive agents. In particular, the agent should learn at run time various concepts ($car(x)$), facts ($car(p)$), and rules (e.g. $bird(x) \rightarrow fly(x)$) that constitute its beliefs. Moreover, given goal formulae ϕ and ψ , plan expressions π and π' , and belief formula β , the following can be the subject of learning with regards to the agent's goals and plans:

- Which goal to select in order to plan, and which plan to select in order to execute? Two types of selection functions can be learned. The goal selection functions should be learned based on the agent's beliefs and the plan selection functions should be learned based on the agent's beliefs and goals.
- Which goal or plan to generate in certain situations? This can be achieved by learning goal or plan rules of the form $\top \leftarrow \beta \mid \phi$ and $\top \leftarrow \beta \mid \pi$, respectively.
- Which goal or plan to drop in certain situation? This can be achieved by learning goal or plan rules of the form

$\phi \leftarrow \beta \mid \top$ and $\pi \leftarrow \beta \mid \epsilon$, (where ϵ is the empty plan) respectively.

- Which goal or plan to modify in certain situation? This can be achieved by learning goal or plan rules of the form $\phi \leftarrow \beta \mid \psi$ and $\pi \leftarrow \beta \mid \pi'$, respectively.
- How to plan a goal in a certain situation? This can be achieved by learning planning rules of the form $\phi \leftarrow \beta \mid \pi$.

Finally, for each type of rules a selection function should be learned that selects a rule to apply at each moment of time. These selection functions should be learned based on agent's mental state and differ for each type of rules. In particular, the selection function for planning rules should be learned based on agent's beliefs, goals, and plans, the selection function for plan rules should be learned based on agent's beliefs and goals, and the selection function for goal rules should be learned based on agent's beliefs. In the following section we discuss how these aspects can be learned by various learning techniques.

B. Learning Goals, Plans and Concepts

In order to cope with the demands of a sapient agent described in the previous section, we can use hierarchical RL (using relational representations). We consider goal-selection and plan-selection first.

For goal-selection, the agent has to map its beliefs to a particular goal which it will adopt. Each time-step the agent can change its mind about the goal, but in order to allow the agent to continue with one particular goal, we can use a mapping from beliefs and the previous goal to a newly selected goal. Reinforcement learning algorithms learn value functions for this by trial and error using e.g. Q-learning [24]. The goal is to learn to select goals leading to the maximal average reward intake per time-step. By trying out goals, and using plans or actions to achieve these goals, the agent gets estimates about the quality-value (Q-value) of selecting each of its goals given some mental states. Since the agent can at any time change its goal, it can drop previous goals and continue with new ones. It can also learn that committing to some goal is good until some mental state tells the agent to adopt another goal. Thus, using the hierarchical RL framework, selecting and revising goals may be learned just as learning action sequences.

For learning to select plans, the agent has to map a goal and beliefs to a plan. There can be multiple plans, some plans may even consist of single actions. Although some plans take longer than single actions, this is not any problem if hierarchical RL is being used. The agent can even choose to invoke a planner which will then plan at a specific time-step. If this planner returns useful plans given some mental state, it will be invoked more often in that context. Plans can be dropped or revised at any time, since the agent selects a plan or action at each time step. Thus, again using the hierarchical RL framework, selecting and revising plans can be learned just as learning action sequences.

Finally, the agent has to learn to map sensory information obtained by for example cameras to concepts. This can be

done by using pattern recognition methods such as neural networks or support vector machines. Each time the agent receives sensory information and does not understand what it sees, it should get feedback about the concept it is looking at. This can only be done in a social setting in which humans communicate with the agent, and the agent is also able to communicate with other agents. We will examine this issue further in section VI.

C. Emotions in Learning

Emotions may influence behavior as was explained in section III-C. Emotions may also influence learning. For example, a negative emotion that produces a bad feeling may trigger reassessment of what causes the bad feeling, followed by learning how to avoid it in the future. A sapient agent can also predict that by not doing an action, it will feel even worse, and by feeling this, it can interrupt its current behavior to do that action. Emotions can also help focus on a goal, or trigger to reassess a situation (e.g. by *insight*, *reflection*) and look for a way to improve it, thus adapting the behavior.

VI. THE SOCIAL ENVIRONMENT

Agents, especially sapient ones, will usually be *situated* in complex, multi-agent, social environments in which they have to interact with other agents and humans. Such complex environments create difficulties, but also opportunities, especially in learning. We will discuss some of these in this section.

Reinforcement Learning has already been applied successfully for solving particular multi-agent problems such as network routing [25], elevator control [26], and traffic light control [27]. For all these problems, the agent still has to solve a particular task such as controlling a specific traffic light and therefore these agents are not sapient at all. We can use multi-agent systems to make it easier for agents to learn to become sapient agents, however.

If the agent has to learn to achieve a goal and it can choose which task to learn, there are several complicated issues. In some sense, the agent has to devote its time to learn something useful. But what if the agent is unable to learn to solve a particular task? When should it stop trying to learn the task? And also, how much reward can it expect when it would be able to learn to perform the task? A solution is to let the agent learn from other agents. For example, the agent can estimate its learning time by looking at other agents, or by communicating with them. The agent can also ask the reward functions of other agents, it can estimate the learning time by asking or looking at the other agent, and the agent can even ask the decision skill to solve a particular task to another agent. Thus, some issues seem complicated, but may become easier when the agent is not alone in the world. Although the whole system would become much more complex, particular subproblems are easier to solve. Some problems would even be impossible if the agent cannot learn by imitating other agents. For example, suppose one agent, a robot, approaches a deep canyon and just near the edge, it slips and falls into the canyon. Because of the fall, the agent is destroyed and it discontinues

to exist. The only way of learning that one should not come too close to the edge of the canyon is to look at the results of other agents approaching it. Since it is easy to see that coming too close to the edge of the canyon was bad for the other agent, a sapient agent can learn that this is a wrong action in this context.

In multi-agent settings we have to distinguish between competitive, co-operative, and semi-competitive settings. Although we would like all agents to be cooperative, this is not realistic since each agent tries to maximize its own average reward intake per time-step. However, even in (semi-)competitive settings it makes sense to let agents communicate (e.g. if two agents try to walk through the same corridor and bump against each other, they can signal to which side they will go). By communicating knowledge, agents can share experiences, concepts and procedural knowledge of how to solve tasks. Using communication, possibly with humans, is also a good way to get a lot of examples for learning to classify sensory information into concepts. These are examples of *learning by communicating*. Classical experiments show that in many cases, communication between agents can have a positive influence on learning behavior, provided that communicated information is useful and not superfluous [28]. On the other hand, agents can also *learn how to communicate* [29]. This involves learning *what*, *when*, *with whom* and *how* to communicate. Social laws, protocols and shared *ontologies* are important factors in communication.

Agents can learn to judge just like other agents, and agents can reward each other using ethical or social laws which have already existed for a long time and therefore may be evolved or preprogrammed. Thus, in multi-agent systems judgment and insight can also be learned, obtained, and refined using communication. For communication between agents some issues such as trust (insight in relationships) plays an important role and has to be learned based on the experiences of the agent. If another agent provides wrong estimates about the learning time or reward for solving a particular task, or it gives a wrong decision skill for solving the task, the agent can learn that this agent cannot be trusted. The agent can also ask other trusted agents, whether they trust another agent. In this way social relationships among agents can evolve.

The problem of using reward functions is that it is difficult to say how much reward one should get for task A relative to task B. The decision of the agent will be to do the task leading to maximal average reward per time-step. However, if these relative reward values are incorrect, the agent could always do one single task at which it is good. Therefore the reward function should also be dynamic, where a reward is given only under particular circumstances. The reward could be made dependent on the agent's emotions such as boredom, pride, pity, disappointment, satisfaction, anger, etc. In this way, an agent who is angry with another agent may learn not to communicate interesting information. Also if the agent is bored with its current task, it will get less reward for doing it, and therefore may switch to another goal.

VII. CONCLUSIONS

In this paper we have given a characterization of *sapient agents*. By starting from the notion of a *cognitive agent*, for which many formalizations exist, we place cognitive notions such as *beliefs*, *desires*, *goals*, and *plans* at the core of the deliberation cycle of a sapient agent. Furthermore, with this as a starting point, we have a firm basis for a model of true *sapience* as well as that we can take advantage from existing knowledge and formalizations concerning the modelling of cognitive notions, logic-based systems and agent programming languages such as 3APL.

Furthermore, we have emphasized the need for managing control over different tasks that can be performed in parallel, by choosing constantly between actions, goals and plans in the deliberation cycle. Various tasks can also be run in parallel on different cognitive levels. On the perceptual level, pattern recognition can transform visual images to (logical) concepts, while planning and acting can be performed on a higher cognitive level.

We have also stressed the importance of *emotions* as a possible factor in both behaving and learning. In a single agent, emotions may influence decision-making and planning. In a multi-agent, social context, emotions may play an important role in the interaction, especially when humans are involved.

A very important feature of sapient agents that we discussed is *learning*. We discussed reasons, opportunities and solutions for learning. For sapient agents, learning transcends the idea of single-task learning by focusing on the whole deliberation cycle, emotional attitudes and the social context. Of much importance will be the integration of RL methods and logic-based, cognitive agents. Relational languages and hierarchical learning methods in RL may function as a bridge between cognition and learning.

One line of further research should focus first on formal definitions of the various parts discussed in this paper. Formal notions present in formalizations of cognitive agents and agent programming languages should be extended with learning mechanisms and emotional attitudes.

A second line of research should aim at *experimenting* with and *developing* concrete applications of increasingly sapient agents. By integrating learning mechanisms such as relational RL into agent programming languages such as 3APL, ideas can be put to a test in order to develop truly sapient agents.

REFERENCES

- [1] M. Dastani, F. de Boer, F. Dignum, and J.-J. Meyer, "Programming agent deliberation: An approach illustrated using the 3apl language," in *Proc. of the Second Int. Conference on Autonomous Agents and Multiagent Systems (AAMAS'03)*. ACM Press, 2003.
- [2] A. Rao and M. Georgeff, "BDI agents: from theory to practice," in *Proc. of the Int. Conference on Multi-Agent Systems (ICMAS'95)*, 1995, pp. 312–319.
- [3] —, "Modeling rational agents within a BDI architecture," in *Proc. of Second Int. Conference on Knowledge Representation and Reasoning (KR'91)*. Morgan Kaufmann, 1991, pp. 473–484.
- [4] A. S. Rao, "AgentSpeak(L): BDI agents speak out in a logical computable language," in *Seventh European Workshop on Modelling Autonomous Agents in a Multi-Agent World*, R. van Hoe, Ed., Eindhoven, The Netherlands, 1996.
- [5] M. d'Inverno, D. Kinny, M. Luck, and M. Wooldridge, "A formal specification of dMARS," in *Agent Theories, Architectures, and Languages*, 1997, pp. 155–176.
- [6] K. V. Hindriks, F. S. D. Boer, W. V. der Hoek, and J.-J. C. Meyer, "Agent programming in 3apl," *Autonomous Agents and Multi-Agent Systems*, vol. 2, no. 4, pp. 357–401, 1999.
- [7] J. Broersen, M. Dastani, J. Hulstijn, and L. van der Torre, "Goal generation in the BOID architecture," *Cognitive Science Quarterly*, vol. 2(3-4), pp. 428–447, 2002.
- [8] M. Dastani, F. Dignum, and J.-J. Meyer, "Autonomy and agent deliberation," in *Proc. of The First Int. Workshop on Computational Autonomy - Potential, Risks, Solutions (Autonomous 2003)*, 2003.
- [9] M. Dastani, B. van Riemsdijk, F. Dignum, and J.-J. C. Meyer, "A programming language for cognitive agents: Goal-directed 3APL," in *Proc. the ProMAS 2003 Workshop at AAMAS'03*, 2003.
- [10] R. Picard, *Affective Computing*. Cambridge, MA: MIT Press, 1997.
- [11] A. Ortony, G. Clore, and A. Collins, *The cognitive structure of emotions*. Cambridge, MA: Cambridge University Press, 1988.
- [12] A. Damasio, *Descartes' Error: Emotion, Reason and the Human Brain*. New York, NY: Gosset/Putnam, 1994.
- [13] R. Sutton and A. Barto, *Reinforcement Learning: an Introduction*. Cambridge: The MIT Press, 1998.
- [14] G. Tesauro, "Practical issues in temporal difference learning," in *Advances in Neural Information Processing Systems 4*, D. S. Lippman, J. E. Moody, and D. S. Touretzky, Eds. San Mateo, CA: Morgan Kaufmann, 1992, pp. 259–266.
- [15] T. Dietterich, "Hierarchical reinforcement learning with the MAXQ value function decomposition," *Journal of Artificial Intelligence Research*, vol. 13, pp. 227–303, 2000.
- [16] R. Sutton, D. Precup, and S. Singh, "Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning," *Artificial Intelligence*, vol. 112, no. 1–2, pp. 181–211, 1999.
- [17] R. Parr and S. Russell, "Reinforcement learning with hierarchies of machines," in *Advances in Neural Information Processing Systems 11*, 1997.
- [18] R. Reiter, *Knowledge in Action: Logical Foundations for Specifying and Implementing Dynamical Systems*. Cambridge, Massachusetts: The MIT Press, 2001.
- [19] C. Boutilier, R. Reiter, and B. Price, "Symbolic dynamic programming for first-order MDP's," in *Proc. of the 17th Int. Joint Conference on Artificial Intelligence (IJCAI-01)*, B. Nebel, Ed. San Francisco, CA: Morgan Kaufmann Publishers, Inc., Aug. 4–10 2001, pp. 690–697.
- [20] S. Dzeroski, "Relational reinforcement learning for agents in worlds with objects," in *Proc. of the Symposium on Adaptive Agents and Multi-Agent Systems (AISB'02)*, 2002, pp. 1–8.
- [21] M. van Otterlo, "Relational expressions in reinforcement learning: Review and open problems," in *Proc. of the ICML'02 Workshop on Development of Representations*, E. de Jong and T. Oates, Eds., 2002.
- [22] —, "Efficient reinforcement learning using relational aggregation," in *Proc. of the Sixth European Workshop on Reinforcement Learning, Nancy, France (EWRL-6)*, 2003.
- [23] L. Kaelbling, T. Oates, N. Hernandez, and S. Finney, "Learning in worlds with objects," in *The AAAI Spring Symposium*, Mar. 2001.
- [24] C. J. C. H. Watkins, "Learning from delayed rewards," Ph.D. dissertation, King's College, Cambridge, England, 1989.
- [25] M. Littman and J. Boyan, "A distributed reinforcement learning scheme for network routing," in *Proc. of the First Int. Workshop on Applications of Neural Networks to Telecommunication*, J. Alspector, R. Goodman, and T. Brown, Eds., Hillsdale, New Jersey, 1993, pp. 45–51.
- [26] R. Crites and A. Barto, "Improving elevator performance using reinforcement learning," in *Advances in Neural Information Processing Systems 8*, D. Touretzky, M. Mozer, and M. Hasselmo, Eds. Cambridge MA: MIT Press, 1996, pp. 1017–1023.
- [27] M. A. Wiering, "Multi-agent reinforcement learning for traffic light control," in *Proc. of the Seventeenth Int. Conference on Machine Learning*, P. Langley, Ed., 2000, pp. 1151–1158.
- [28] M. Tan, "Multi-agent reinforcement learning: Independent vs. cooperative agents," in *Proc. of the Tenth Int. Conference on Machine Learning*, June 1993, pp. 330–337.
- [29] G. Weiss, *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. Cambridge, Massachusetts: The MIT Press, 1999.