# Agenten

Die Agentenmetapher basiert auf Entwicklungen in verschiedenen Bereichen der Informatik wie verteilte Systeme, Softwaretechnik und künstliche Intelligenz. Sie wurde auch stark von den Forschungsergebnissen anderer Disziplinen beeinflusst, insbesondere der Soziologie, der Biologie, der System- und Entscheidungswissenschaften und vieler anderer. Diese verschiedenen Forschungsbereiche spiegeln sich in den zahlreichen Facetten wider, die Agenten charakterisieren.

Die Frage "Is it an agent, or just a program?" wurde von Franklin und Graesser (Franklin und Graesser, 1996) gestellt und zeigt bereits einige der Probleme bei der Definition eines Agenten auf. Das Papier gehört zu einer Reihe von Bemühungen, die in den 1990er Jahren auf die Definition von Agenten abzielten. Damals gab es ein großes Interesse an solchen Definitionen, da Agenten gerade als neues Paradigma eingeführt worden waren. Im Laufe der Jahre nahm dieses Interesse an einem formalen Konsens ab und der Schwerpunkt wurde auf die Anwendungsbereiche des Agentenparadigmas gelegt. Glücklicherweise ist, wie auch andere wissenschaftliche Bereiche zeigen, eine gemeinsam vereinbarte Definition keine Voraussetzung für den Erfolg eines Konzepts in der Praxis.

Um die obige Frage zu beantworten, diskutieren Franklin und Graesser mehrere Definitionen prominenter Agentenforscher.

Der von Russell und Norvig (Russell und Norvig, 1995) vorgeschlagene Agentenbegriff, wonach "ein Agent alles ist, was als seine Umgebung durch Sensoren wahrnehmend und durch Effektoren auf diese Umgebung einwirkend betrachtet werden kann", gilt für alle Programme, deren Ausgaben auf Eingaben beruhen.

Andere Definitionen berücksichtigen die Art der Umgebung, d.h. ob sie dynamisch und komplex ist, und gehen auch von einer gewissen Autonomie und Zielgerichtetheit des Agenten aus, z.B.: "Autonome Agenten sind Computersysteme, die eine komplexe dynamische Umgebung bewohnen, diese wahrnehmen und in ihr autonom handeln und dabei eine Reihe von Zielen oder Aufgaben verwirklichen, für die sie entworfen wurden." (Maes, 1995). Es wird nichts darüber gesagt, wie diese Autonomie oder Zielgerichtetheit von einem konkreten Agenten erreicht werden soll. Im Gegensatz dazu unterstreicht (Hayes-Roth, 1995) die traditionelle Sichtweise der künstlichen Intelligenz auf Agenten, indem er davon ausgeht, dass alle Agenten explizit schlussfolgernde Mechanismen verwenden: "Intelligente Agenten führen kontinuierlich drei Funktionen aus: Wahrnehmung dynamischer Bedingungen in der Umwelt; Handeln, um die Bedingungen in der Umwelt zu beeinflussen; und schlussfolgerndes Denken, um Wahrnehmungen zu interpretieren, Probleme zu lösen, Schlüsse zu ziehen und Handlungen zu bestimmen.".

Andere Forscher folgen der Argumentationslinie von Brooks, der provokativ die Frage aufgeworfen hat, ob der Entwurf intelligent handelnder Systeme Repräsentation oder Denken erfordert (Brooks, 1991a,b). Brooks hebt die Reaktionsfähigkeit von Agenten als eine der intrinsischen Eigenschaften von Agenten hervor, die mit Mitteln zur Deliberation kombiniert werden müssen (Stone und Veloso, 2000).

Wooldridge und Jennings (Wooldridge und Jennings, 1995) definieren einen Agenten grundsätzlich als ein hardware- oder softwarebasiertes Computersystem, das die folgenden Eigenschaften aufweist:

- Autonomie: Agenten operieren ohne direkten Eingriff von Menschen oder anderen und haben eine gewisse Kontrolle über ihre Handlungen und ihren internen Zustand;

- Soziale Fähigkeit: Agenten interagieren mit anderen Agenten (und möglicherweise Menschen) über eine Art von Agenten-Kommunikationssprache;

- Reaktivität: Agenten nehmen ihre Umgebung wahr (das kann die physische Welt, ein Benutzer über eine grafische Benutzeroberfläche, eine Ansammlung anderer Agenten, das Internet oder vielleicht alles zusammen sein) und reagieren rechtzeitig auf Veränderungen, die in dieser Umgebung auftreten;

- Proaktivität: Agenten agieren nicht einfach als Reaktion auf ihre Umgebung; sie sind in der Lage, ein zielgerichtetes Verhalten an den Tag zu legen, indem sie die Initiative ergreifen.

Auch hier werden keine Annahmen darüber gemacht, wie bestimmte Eigenschaften erreicht werden können. Allerdings werden zentrale Probleme von Agenten berücksichtigt, nämlich wie zwischen Reaktivität und Proaktivität vermittelt werden kann und wie die Interaktion mit anderen Agenten und möglicherweise Menschen erfolgen kann. Die Umwelt wird nur implizit berücksichtigt: Um auf etwas reagieren zu können, muss der Agent es wahrnehmen. Dieses Manko wurde in späteren Veröffentlichungen behoben: "Ein Agent ist ein Computersystem, das sich in einer bestimmten Umgebung befindet und zu flexiblen, autonomen Handlungen fähig ist, um seine Entwurfsziele zu erreichen... Es gibt also drei Schlüsselkonzepte in unserer Definition: Situiertheit, Autonomie und Flexibilität" (Jennings et al., 1998).

Um die Aspekte hervorzuheben, die für die Systemtechnik von besonderem Interesse sind, schlagen wir im Rahmen dieser Arbeit die folgende Arbeitsdefinition vor:

- Der Agent ist situiert, er nimmt seine Umgebung wahr und er handelt in seiner Umgebung. Die Umwelt umfasst typischerweise andere Agenten, andere teilweise dynamische Objekte und passive Objekte, die z.B. durch den Agenten manipuliert werden können. Die Kommunikation mit anderen Agenten ist in Systemen mit mehreren Agenten von besonderem Interesse, da die Agenten zusammenarbeiten und um Aufgaben konkurrieren können. Diese letztere Eigenschaft wird auch als soziale Fähigkeit bezeichnet.

- Der Agent sollte in dem Sinne autonom sein, dass er ohne direktes Eingreifen von Menschen oder anderen agieren kann, und Autonomie erfordert die Kontrolle über seinen eigenen Zustand und sein Verhalten.

- Flexibel zu sein bedeutet für einen Agenten, dass er zwischen reaktivem Verhalten, d. h. der Fähigkeit, auf Veränderungen in seiner Umgebung zu reagieren, und der Absicht, seine Ziele zu verfolgen, vermitteln muss. Eine geeignete Vermittlung ist einer der entscheidenden Aspekte für einen Agenten, um seine Aufgaben in einer dynamischen Umgebung zu erfüllen. Ein Agent kann auf der Grundlage seines Wissens, seiner Regeln, Überzeugungen, Operatoren, Ziele und Erfahrungen handeln und sich bei Bedarf schnell an neue Einschränkungen und Anforderungen - oder sogar neue Umgebungen - anpassen. So können beispielsweise neue Situationen neue Ziele erfordern, und neue Erfahrungen können zu neuen Verhaltensregeln führen. Außerdem erhöht die Mobilität die Flexibilität eines Agenten.

## Situated Environment and Agent Society

In der Regel haben wir es nicht nur mit einem Agenten zu tun, sondern mit mehreren Agenten in ihrer Umgebung. Im Folgenden werden wir Agentensysteme als Systeme verstehen, die aus einer Reihe von Objekten bestehen. Es wird zwischen aktiven und passiven Objekten unterschieden. Passive Objekte ändern ihre Eigenschaften im Allgemeinen nicht durch ein eigenes Verhalten; sie sind oft ausschließlich der Manipulation unterworfen. Im Gegensatz dazu haben aktive Objekte ein eigenes Verhalten; Agenten sind eine Teilmenge dieser aktiven Objekte. In Multi-Agenten-Systemen verfügt jeder Agent in der Regel nur über unvollständige Informationen oder Fähigkeiten zur Lösung des jeweiligen Problems, Daten und Steuerung sind dezentralisiert, und die Berechnungen erfolgen asynchron.

In Abbildung 3.1 befindet sich ein Agent in seiner Umgebung, er nimmt seine Umgebung (einschließlich anderer Agenten) wahr, bildet die Wahrnehmung auf der Grundlage seines Wissens und seiner Ziele auf eine interne Repräsentation ab, und er kommuniziert und handelt mit anderen Agenten. Seine Umwelt enthält andere Objekte, passive wie Hindernisse, aber auch aktive wie den Ball, der (sobald er geschossen wurde) eine Eigendynamik hat, die berücksichtigt werden muss. Die Wahrnehmung der Umwelt durch einen Agenten kann unvollständig (er weiß nicht alles, was wahr ist, z. B. nimmt der Agent das Dreieck und den zweiten Block nicht wahr) und unsicher sein (es ist auch nicht alles wahr, was er glaubt, dass es wahr ist). Agenten sind in Bezug auf ihr Wissen und, aufgrund der dynamischen Umgebung, auf die Zeit, die sie für eine Entscheidung und zum Handeln benötigen, begrenzt.

In dieser Hinsicht sind sie dem Menschen nicht unähnlich: Menschen agieren beim Treffen von Entscheidungen in einer komplexen und sich häufig verändernden Umgebung mit begrenzten kognitiven Fähigkeiten, Zeit und anderen Ressourcen. Daher ist die Entscheidungsfindung nur innerhalb der den Entscheidungsträgern auferlegten Grenzen rational. Tversky und Kahneman (Tversky und Kahneman, 1974) haben eine Reihe von Heuristiken und Voreingenommenheiten identifiziert, die der Mensch bei der Entscheidungsfindung nutzt. Diese Studien zielen darauf ab, klassische und analytische Entscheidungstheoretiker mit den Erkenntnissen der kognitiven Psychologie in Einklang zu bringen. So führen Heuristiken im Vergleich zu klassischen Methoden häufig zu kostengünstigeren Lösungen in Bezug auf Zeit und geistigen Aufwand, eine Tatsache, die auch von traditionellen Methoden der künstlichen Intelligenz ausgenutzt wird.

*The agent metaphor is based on developments in diverse computer science areas such as distributed systems, software engineering, and artificial intelligence. It has been strongly influenced by the research results of other disciplines as well, in particular sociology, biology, systems and decision science, and many others. Those diverse areas of research are reflected in multiple facets that characterize agents.*

*The question “Is it an agent, or just a program?” was posed by Franklin and Graesser (Franklin and Graesser, 1996) and already reveals some of the problems in defining an agent. The paper belongs to a series of efforts aimed at defining agents during the 1990s. At that time, there was a lot of interest in such definitions, as agents had just been introduced as a new paradigm. Over the years, this interest in finding a formal consensus decreased and emphasis was placed on application domains of the agent paradigm. Fortunately, as other scientific areas show as well, a commonly agreed-upon definition is not a prerequisite for the success of a concept in practice.*

*To answer the above question, Franklin and Graesser discuss several definitions of prominent agent researchers.*

*The notion of agents suggested by Russell and Norvig (Russell and Norvig, 1995),that “an agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through effectors,” applies to all pro-grams whose outputs are based on inputs.*

*Other definitions take the type of environment into account, that is, whether itis dynamic and complex, and also assume some autonomy and goal-directedness of the agent, for example, “autonomous agents are computational systems that inhabit some complex dynamic environment, sense and act autonomously in this environment, and by doing so realize a set of goals or tasks for which they are designed.” (Maes, 1995). Nothing is said about how this autonomy or goal-directedness will be achieved by a concrete agent. In contrast, (Hayes-Roth, 1995) emphasizes the traditional artificial intelligence view on agents by assuming that all agents use explicitly reasoning mechanisms: “Intelligent agents continuously perform three functions: perception of dynamic conditions in the environment; action to affect conditions in the environment; and reasoning to interpret perceptions, solve problems, draw inferences, and determine actions.”.*

*Other researchers follow the argumentation line of Brooks, who provocatively raised the question of whether the design of intelligently behaving systems requires representation or reasoning (Brooks, 1991a,b). Brooks emphasizes the reactiveness of agents as one of the intrinsic features of agents that must be combined with means for deliberation (Stone and Veloso, 2000).*

*Wooldridge and Jennings (Wooldridge and Jennings, 1995) define an agent basically as a hardware- or software-based computer system that has the following properties:*

*• Autonomy: agents operate without the direct intervention of humans or others and have some kind of control over their actions and internal state;*

*• Social ability: agents interact with other agents (and possibly humans) via some kind of agent-communication language;*

*• Reactivity: agents perceive their environment (which may be the physical world, a user via a graphical user interface, a collection of other agents, the Internet, or perhaps all of these combined) and respond in a timely fashion to changes that occur in it;*

*• Proactiveness: agents do not simply act in response to their environment; they are able to exhibit goal-directed behavior by taking the initiative.*

*Again, no assumption is made as to how to achieve certain characteristics. How-ever, central problems of agents are taken into account, that is, how to mediate between reactiveness and proactiveness and how to interact with other agents and, potentially, humans. The environment is only implicitly considered: in order to react to something, the agent must perceive it. This deficiency has been addressed in later publications: “an agent is a computer system, situated in some environment, that is capable of flexible autonomous action in order to meet its design objectives... There are thus three key concepts in our definition: situatedness ,autonomy, and flexibility” (Jennings et al., 1998).*

*In the context of this paper, to highlight the aspects of particular interest to systems engineering, we propose the following working definition:*

*• The agent is situated, it perceives its environment, and it acts in its environment. The environment includes typically other agents, other partly dynamic objects and passive ones that are, for example, subject to manipulation by the agent. Communication with other agents is of particular interest in systems comprising multiple agents, as agents can collaborate and compete for tasks. This latter characteristic has also been referred to as social ability.*

*• The agent should be autonomous, in the sense that it can operate without the direct intervention of humans or others, and autonomy requires control over its own state and behavior.*

*•To be flexible for an agent means to mediate between reactive behavior, being able to react to changes in its environment, and deliberativeness to pursue its goals. A suitable mediation is one of the critical aspects for an agent to achieve its tasks in a dynamic environment. An agent can act upon its knowledge, its rules, beliefs, operators, goals, and experiences, and soon and adapt to new constraints and requirements – or even new environments – as required. For example, new situations might ask for new goals, and new experiences might lead to new behavior rules. Also, being mobile adds to the flexibility of an agent.*

## *Simulated Agent Society*

*Typically, we are not only faced with one agent, but with multiple agents in their environment. In what follows we will understand agent systems as systems that are comprised of a set of objects. We distinguish between active and passive objects. In general, passive objects do not change their attributes driven by a behavior of their own; they are often exclusively subject to manipulation. In contrast, active objects have a behavior of their own; agents are a subset of these active objects. In multi-agent systems, generally each agent has incomplete information or capabilities for solving the problem at hand, data and control are decentralized, and computation happens asynchronously.*

*In Figure 3.1 an agent is situated in its environment, it perceives its environment (including other agents), maps the perception to an internal representation, based on its knowledge and goals, and it communicates and acts with other agents. Its environment contains other objects, passive ones like obstacles, but also active ones, like the ball, which has a dynamic of its own (once kicked) and which must be taken into account. An agent’s perception of the environment can be incomplete (it does not know all of what is true, for example, the agent does not perceive the triangle and the second block) and uncertain (nor is all true what it believes to be true). Agents are bounded with respect to their knowledge and, due to the dynamic environment, to the time they must come up with a decision and to act.*

*In this regard, they are not unlike humans: in making decisions, humans operate within a complex and often changing environment with limited cognitive capabilities, time, and other resources. Hence, decision making is only rational within the bounds imposed on decision makers. Tversky and Kahneman (Tversky and Kahneman, 1974) identified a number of heuristics and biases that humans use to make decisions. These studies aim to bring classical and analytic decision theorists into conformity with findings in cognitive psychology. Thus, heuristics often yield cost-effective solutions compared to classical methods in terms of time and mental effort, a fact that is also exploited by traditional methods in artificial intelligence.*