# Agenten

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.