# Agent, autonomous

Annie Luciani

# Agent, autonomous

Annie Luciani [ACROE&INPG]

*Contributors: Ronan Boulic [EPFL]*

The expression autonomous agents, widely used in virtual reality [→ Virtual reality and virtual environment], computer graphics [→ Computer graphics], artificial intelligence and artificial life, corresponds to the simulation of autonomous creatures, virtual (i.e. totally computed by a program), or embodied in a physical envelope, as done in autonomous robots.

Intelligent characters [→ Intelligent characters] are a specific category of autonomous agents. The distinction between both comes from the difference between intelligence and autonomy, these two concepts being not synonymous.

Strictly speaking, autonomous and agent are redundant, in the sense that the basic definition of an agent is "an autonomous adaptive entity", having the following basic properties:

- Able to satisfy goals in a complex environment;
- Autonomous: able to do so by itself, i.e. without any external help, by controlling its actions and its internal states;
- Adaptive: able to improve its performance.

As a remark, autonomy is sometimes restrictively identified with the capability of perceiving the environment through sensors and act on it through effectors. By adding others properties to these three basic ones, especially the following, the autonomous agent goes toward the status of intelligent creature (or intelligent agent, or intelligent character):

- Social: interacting with other agents through a language ;
- Pro-active: able to define its own goals by itself;
- Able to learn: as an improvement of adaptive capabilities.

Two main types of autonomous agents are usually confronted, based on two different approaches on the nature of the autonomy: (1) Deliberative agents, cognitive agents, rule-based agents; and (2) Reactive agents or stimulus/reaction–based agents.

**Deliberative agents, cognitive agents, rule-based agents**

Such agents lead to the Symbolic Artificial Intelligence paradigm, as founded by the Newell and Simon's concept of symbol system [Newell & Simon, 1976]. They contain an explicit symbolic representation of the world, on the basis of which decisions are taken by logical symbolic reasoning. One of the most relevant implementation of this type of agent is the Believe-Desire-Intention of Rao [Rao & Georgeff, 1991].

The main difficulties on such approaches are:

- The need of a completeness of the system of rules
- The lack of reactivity, directly linked to the number of rules
- The maintaining of the coherence of the rules systems during their evolution

**Reactive agents, stimulus/reaction–based agents**

Such agents are able to exhibit autonomous behaviours without any explicit symbolic representation nor abstract reasoning. Well-known implementations of those are Brooks' robots [Brooks, 1991]. They are deterministic and passive systems, but they are not capable of defining their own goals and long lasting planning.

The software and hardware architecture implementing such concepts are called behaviour-based architectures. They are composed of modular behavioural blocks and perception-decision-action processes. Many hardware and software architectures have been proposed. The main three are based either on artificial neural networks, finite state automata, or logical rules systems.

### Basic underlying assumptions

Each of these approaches are grounded on specific hypothesis and philosophical enlightening on what are intelligence, autonomy, living organisms. Let us sum up some of the main assumptions.

- The principle of modularity.

Modularity is the principle according to which autonomous behaviours may emerge from a large collection of interacting simple and specialized behaviours. It corresponds to the ant paradigm, implementing the Taylorism organisation of labour, based on division of labour according to predefined task specializations.

- A modelling inspired by neural biological systems.

Especially, artificial neural networks, or finite state automata networks, are considered to model neural biological systems. These approaches are close to non-representationalist cognitive approaches. The behaviours emerge from the network structure. In neural networks, the data processed are "bits" (i.e. representing logical states) while in finite state automata networks, they can be real and integers (i.e. representing analogical states).

- A pure logical interpretation of decisional and reasoning capabilities.

Especially, logical rules systems are usually used, based on deduction and inference processes, which are at a high level of abstraction [Jennings, 2000]. This approach is directly inspired by the computational theory of mind [→ Computational paradigm]. The perception-decision-action process, supposed to be inspired by processes observed in human, is often implemented as a basic functionality. The main difficulty of this approach when implementing artificial creatures is that the specification of a complex performed behaviour is unwieldy because of a few numbers of rules compatible with a reasonable time of computation, and thus the method is strongly confronted to the automation of the creation of the rules, and thus it strongly depends on learning level.

### Role of autonomous agents in enaction and enactive interfaces

The mediations between humans and world through computerized representations can be split in five categories of interaction:

- Physical interaction between real human and the real physical world (such as is teleoperation).
- Physical interaction between real human and a virtual physical world.
- Symbolic or physical interaction between real humans and virtual agents.
- Interaction between real humans and the real physical world through virtual physical objects.
- Interaction between real humans and the real physical world through virtual agents.

Consequently, the assumptions on the nature of virtual autonomous agents as well as of the nature of real autonomous agent, intervene directly in the enactivity of the interaction:

- between real human and the real physical world through virtual agents;
- and between real humans and virtual agents.

This enables implementing situations from those that absolutely non-enactive (as those based on computational theory of mind for both interacting bodies) to those that are very close to the enaction paradigm (as those based on automaton networks).

## References

[Brooks 1991] Brooks R.A. (1991). "Intelligence without representation ». Artificial Intelligence 47 (1991), 139–159.

[Jennings 2000] Nicolas R. Jennings. On agent-based software engineering. Artificial Intelligence, 117:277–296, 2000.

[Newell & Simon, 1976] Newell A, Simon H.A. "Computer Science as Empirical Enquiry. Communications of ACM. Vol. 19. Pp. 113-126. 1976.

[Rao & Georgeff, 1991] Rao A.S., Georgeff M.P. "Modeling rational agents within BDI-Architecture. Proceedings of the 2nd International Conference on Principles of Knowledge representation ahd Reasonning". Pp 473-484. Morgan Kaufmann Publishers Inc. April 1991.

## Related items

Autopoïesis
Computational paradigm
Computer graphics
Dynamic systems
Intelligent characters
Virtual reality and virtual environment

## Related documents

EI_Enaction&ComputerTechnologies_Luciani.pdf