A view on the architecture and design of highly autonomous and situated agents

Masthoff, J., & Van Hoe, R.
Institute for Perception Research (IPO)
P.O. Box 513
5600 MB Eindhoven, The Netherlands
{masthoff, vanhoe}@prl.philips.nl

Abstract

The architecture and design process of a highly autonomous, situated agent is described. The architecture consists of a population of agents which are designed in an incremental, bottom-up fashion according to what we call a behavioural engineering approach. The agents operate in parallel and have a close coupling between perception and action. Emergent behaviour and memory-based learning determine the level of adaptability of our approach.

Background

As there is so much different agent-related work going on, a three-dimensional framework is used to situate our agent research. The first dimension of this framework is the situatedness of an agent, i.e., the extent to which the agent reacts to changes in its environment rather than executing a pre-defined plan. The second dimension is the level of autonomy of the agent. Agents vary from the butler-type of agents with limited autonomy (the agent only performs a particular function for another agent) to the salesman-type of agents with extensive autonomy (the agent not only performs a particular function, but also has a self-interest). The third dimension is the task domain of the agent.

As far as the task domain is concerned, we have been working for the last two years on an agent-based coaching application. Because of the unpredictable and dynamic environment in which a coaching system has to function, namely the student, an approach based on highly autonomous and situated agents was adopted.

Architecture

As a first architectural principle, the incremental nature of Brooks' subsumption architecture has been used (Brooks, 1986): it is desirable to have a working system at each step of the design, where each component has a closed loop between perception and action. However, in contradistinction with Brooks' subsumption architecture, the behaviours of an agent are not hierarchically organized. Furthermore, a complex agent is divided into simpler, interacting agents, which are rather independent of each other.

In our coaching application, we currently have five agents, operating in parallel. There is no central agent "in control": the behaviour of the coaching application is an emergent property of the interaction among agents, and among agents and the student. The agents mainly communicate through changes they cause in the environment, i.e., the student and screen contents. This independent way of communication makes it easy to add, maintain, extend, and remove agents.

Design

The design process is based on two rules. First, when there are different skills related to the task of an agent which require independent investigation, an agent should be associated with each of these skills. Second, when different strategies can be used by an agent (e.g., different instruction strategies), the agent should be divided into sub-agents, where each subagent uses another strategy.

Agents are designed in a bottom-up fashion, beginning with a minimal functionality. Simple actions are used which cause situations to change in the desired direction. When a minimal functionality has been implemented, this can be tested by running simulations or experiments, the outcome of which determines the update and inclusion of behaviours in the architecture. The experiments are also used to construct models about human behaviour, which in their turn can be used to find new behaviours.

In our application, each agent is equipped with a rudimentary memory-based learning process. The agent stores interaction sequences into its memory, and on the basis of the statistics of these episodic memory traces, the behaviour of the agent is tuned.

References

Brooks, R. 1986. A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation* 2: 14–23.