# Entertaining Agents

Peter Wavish

*The Multiple Agent Systems project at Philips Research Laboratories (PRL) in Redhill, UK has developed a distinctive approach to the design and implementation of multi-agent systems and has applied it, amongst other things, to entertainment. This article explains some of the philosophy that underlies the approach.*

## Introduction

When we first started talking about agents some years ago we were told to stop using the word ‘agent’ because nobody understood what it meant! Now the situation has turned around completely: everybody understands the word, but it means quite different things to different people. So this article begins by putting forward a particular view – from an artificial intelligence perspective – of what it is to be an agent.

## What is an agent?

In 1987 the philosopher Daniel Dennett published a book called ‘The Intentional Stance’ [**?**]. The key idea in this is that there are different ways of understanding things and hence of developing expectations about their behaviour which allow people to deal with them. An example is a thermostat.

- *The physical stance.* This is the stance of viewing the thermostat as a system obeying the laws of physics. In the case of a thermostat this would involve the temperature dependent expansion and contraction of the bimetallic strip, the physical linkages with the contacts, the making and breaking of the electric circuit, and so on. By studying carefully how the thermostat is built, you can predict what it does.

- *The design stance.* This is the stance of treating the thermostat according to its purpose. The designer’s aim with all (domestic) thermostats is that they should turn the boiler/furnace on when the temperature falls too low, and you trust that the designer has done his job correctly. It does not actually matter how the thermostat works internally.

- *The intentional stance.* This is the stance of treating the thermostat like a person having mental states such as goals and beliefs. The goal of a thermostat is to keep the room at a constant temperature, from time to time it believes that the temperature is too low, and as a result acts by turning on the boiler/furnace.

A thermostat is unusual in that all three stances can be adopted (just about!). Most consumer electronic products rely on the consumer to adopt the design stance. As systems become more complex we may need to resort to the intentional stance so that users can understand how products work and interact with them successfully. This is where agents come in; they are systems towards which users tend to adopt the intentional stance and which (in some cases) can adopt an intentional stance towards each other.

Under this definition, it is often not possible to say with certainty that something is an agent; it depends on how it is perceived. But some things (like bricks or television sets) certainly are *not* agents, and others (like animals and people) certainly are, because the intentional stance helps you to predict their behaviour and the physical and design stances do not.

The key question is: how do you design and build an agent that is intentional (i.e., towards which the intentional stance would naturally be adopted)? There are two clear alternative strategies (with a few other less clear ones in between).

- *Goals and beliefs are represented explicitly.* This is the artificial intelligence, formal modelling, symbolic reasoning approach to agent design. It has its roots in cognitive science and mathematical logic. The big problem is tractibility.
- *Goals and beliefs are epiphenomena; they emerge from the agent's behaviour.* This is the nouvelle AI, behaviour-based [**?**], situated [**?**] approach to agent design. It has its roots in ethology, sociology, and artificial life. The big problem is lack of problem solving power.

## Behaviours and skills

At PRL, we have consistently pursued the second approach because it avoids the technical problems of the first approach. The foundation of our work has been the notion of representing the behaviour of the agent explicitly. Behaviour is to do with interaction taking place in time, and provides a uniform way of modelling agents and their interactions with each other and with the physical world in which they are embedded. RTA [**?**, **?**] is a language that we have developed for building multi-agent systems modelled in this way.

RTA has been used in a variety of research projects within Philips. Currently its main area of application is seen as being the development of agent-based characters in computer games, an early example of which is a word game embedded in a published CD-i title [**?**]. But it is also being applied in other areas such as interactive learning [**?**]

The first demonstrator that we built in RTA was a simulation of a sheepdog rounding up some sheep [**?**]. The sheepdog rounds up the sheep into a flock and drives them towards the shepherd in (apparently) a strongly goal-oriented fashion. In fact, there are no explicit goals, just hundreds of tiny rules of behaviour which together produce the observed overall 'emergent' behaviour of the system.

The second RTA demonstrator we built was of a group of three agents playing the tetris video game [**?**]. This was built in just the same way as the sheepdog demonstrator, by incrementally extending the agents' sets of rules. The result though is qualitatively different: the overall behaviour of the system is obviously skillful rather than simply displaying a particular dynamics.

The notions of behaviour and skill are the bedrock of our approach to agents. But we are concerned with multiple agents (there are no applications for single agents!) and so we are concerned with how agents can interact fluently with each other. How do agents take part in societies? Or looking at it the other way, how can properly functioning societies of agents be built from individual agents?

## Roles and Organisations

Our answer is that agents need complementary skills. When one agent interacts with another, it needs to be able to rely on particular responses to its actions in order to achieve a satisfactory outcome. Some interactions are symmetric, like shaking hands, while others are asymmetric, like buying and selling. Such interactions are optionally entered into the agent, according to circumstance. Once engaged in an interaction, the agent deploys considerable skill in acting, signalling its actions, and interpreting the other agent's actions and signals. Its behaviour is stereotyped; we can think of it using its skill to perform its chosen role.

One of the objections to the behaviour-based approach to agent design, typified by the sheepdog trials demo, is that the behaviour of the system is both produced and designed bottom-up, and hence is difficult to engineer. Roles, however, provide a way of designing multi-agent systems top-down, in terms of organisations. An example of an organisation is a restaurant with roles such as waiter or dishwasher. Instances of organisations occur when suitably skilled agents adopt particular roles and so start interacting with each other.

Another problem with the situated or behaviour-based approach to agent design is that it apparently lacks problem solving power compared with the approach (standard in artificial intelligence) of planning sequences of actions in order to achieve goals. For instance, if an artificial intelligence agent is hungry, it can plan a sequence of actions including obtaining food, preparing it, and eating it in order to satisfy its hunger. The kind of social agents that we are talking about use a different strategy: they can get what they want by adopting the appropriate role, relying on other agents to carry out most of the hard work. For example, if you are hungry, go to a restaurant and act out the role of being a customer. (Or if you need money, sign on as a dishwasher!).

To summarise the first part of this article, we have developed a methodology and technology for designing and implementing systems of interacting agents in terms of organisations, roles, skills and behaviour [**?**].

- *Organisations* are designed top down as sets of roles.
- *Roles* are designed to be complementary to each other.
- *Skills* are required for performing roles and for performing physical tasks.
- *Behaviours* are the basic means by which agents interact with each other and with their physical environment.
- *RTA* is used to implement behaviours and skills as sets of situation-action rules.

## The Humanoid 2 ESPRIT project

The second part of this article describes the work we are doing in the Humanoid 2 ESPRIT project in which these ideas are being used and further developed. We are collaborating with the Institute for Perception Research Eindhoven (IPO) and with EPFL Lausanne and the University of Geneva, who are amongst the world leaders in highly realistic 3-D graphics human figure modelling and animation [**?**].

Our role in the project is to equip their simulated human figures with autonomous agents that direct their actions to produce convincing human behaviour. The central problem we face is that, to do this properly, we would need agents with the intellectual capabilities of real people, and there is no technology on earth that can remotely approach this capability.

So we have to cheat. We have to deceive people into thinking that the simulated human is a real person. In fact deception of this kind is commonplace; it is what authors, directors and actors aim to do when they create or portray characters. A character performed by an actor has no real existence; it is the skill of the actor which is primarily responsible for sustaining the illusion that the character is real.

The approach we are taking in the Humanoid 2 project is therefore to automate acting skills [**?**]. We use our agent technology to build, not characters as such, but artificial actors equipped with the skills required to portray characters. As with a real theatrical production, the agent-actors perform a script which tells them what to do and how to behave. The agent-actors interpret the script according to the circumstances of the performance.

Only foreground activity is normally scripted, so actors must also be capable of improvising unscripted background activity when they are not being explicitly directed. You can probably see already how the ideas on organisations and roles can be applied directly to creating unscripted background activity; routine social interactions, such as buying a drink in a bar, can be designed in terms of complementary roles (drinker/bartender, buying/selling, giving/taking) and implemented by programming the skills that the agents need in order to act out these roles.

We are using the Humanoid 2 project as an opportunity to provide higher level support for building agent-based systems, so that agent developers do not necessarily need to be able to program in RTA. The generic agents that we are developing are modular and extensible. They will consist of the following.

- A generic agent *framework* into which skills can be plugged.
- A library of *skill modules* implementing skills and roles.
- A simple English-like *scripting language*.

We believe that agents of this kind will be useful, not just for games or VR, but for other applications in which fluent interaction between intelligent software systems (or between such systems and people) is desired.

### Concluding remarks

I have described just one of many different technical approaches to the creation of agent-based characters that are being pursued in research laboratories around the world. Our own thinking owes a lot to pioneers in this field such as Brenda Laurel and Joseph Bates. The broader programme of creating agents that are genuinely intentional has been a topic of research in artifical intelligence almost since its inception. The whole field of agents is broader still; the diversity of current research on agents can be seen clearly from the programme of the forthcoming Agents But how do we then design for entertainment?'97 conference which covers the application areas of autonomous robots, softbots, expert assistants and synthetic agents (the kind we are working on).

Ultimately what matters is not the technology in itself, but how convincing the characters are that are designed with its help. Our own work feeds into other research projects that bring together artistic and technical expertise to search for new ways to inform and to entertain.

□

## References

[1] Agre, P., and Chapman, D., 1987, *'PENGI: An implementation of a theory of activity'*, Proceedings of the Sixth National Conference on Artificial Intelligence AAAI-87, pages 268-272, Seattle, WA, 1987.

[2] Brooks, R.A., *'A Robust Layered Control System for a Mobile Robot'*, IEEE Journal of Robotics and Automation, RA-2, April 1986.

[3] Dennett, D.C., *'The Intentional Stance'*, Bradford Books, MIT Press, Cambridge MA, 1987.

[4] Graham, M. and Wavish, P.R., *'Simulating and Implementing Agents and Multi-Agent Systems'*, Proc. 1991 European Simulation Multi-Conference, Copenhagen, June 1991.

[5] Magnenat Thalmann, N. and Thalmann, D., *'Digital Actors for Interactive Television'*, Proceedings of the IEEE, Vol. 83, no.7, July 1995.

[6] Masthoff, J. and van Hoe, R., APPEAL: A multi-agent approach to interactive learning environments. Proceedings of the 1994 MAAMAW workshop, Denmark.

[7] Wavish, P.R., *'Real Time ABLE'*, In: 'Review of 1990', Philips Research Laboratories, Redhill.

[8] Wavish, P.R., *'Exploiting Emergent Behaviour in Multi-Agent Systems'*, Decentralized A.I. 3, eds. Werner and Demazeau, North Holland, 1992.

[9] Wavish, P.R. and Graham, M., *'Roles, Skills and Behaviour: A Situated Action Approach to Organising Systems of Interacting Agents'*, In: 'Intelligent Agents' ed. M.J. Wooldridge and N.R. Jennings, Springer-Verlag, 1995.

[10] Wavish, P.R. and Graham, M, *'Situated action approach to implementing characters in computer games'*, Applied AI Journal, vol 10, no 1, Jan-Feb 1996.

[11] Wavish, P.R. and Connah, D.M., *'Virtual actors that can perform scripts and improvise roles'*, accepted for the First International Conference on Autonomous Agents, Marina del Ray, 1997.

Pasfoto Wavish

*Peter Wavish is project leader of the Multiple Agent Systems Project at Philips Research Laboratories (PRL) in Redhill, UK. He is currently responsible for PRL's involvement in the Humanoid 2 ESPRIT project.*

# Short News

**OOTI Input/Output**

In September 1996, 13 students started the OOTI programme, followed by 7 students in December. For March 1997, already one student is planned. The total number of OOTI students starting in the academic year 1996/97 is 21. This is nearly equal to last year. This year's intake was positively affected by the new possibility of part-time appointments.

The 21 new OOTI students have the following educational background: computing science: 12; physics: 4; information technology: 2; mechanical engineering: 1; electrical engineering: 1; mathematics: 1. With the group of new OOTIs, the OOTI population is becoming more and more heterogeneous in this respect.

Nine OOTI students plan to graduate in December 1996, followed by another OOTI graduate in March 1997. This will bring the total number of certified OOTI graduates at 115. □