Consciousness of Animat
NTT9904-01

Proposal for 1999-2000 Funding

Rodney A. Brooks

Research Target

How can animats have a "consciousness"? To investigate this ultimate question, we would like to employ this working hypothesis: consciousness is formed in the subjective mind of a human observer who is actually communicating with the animats. So the question can be transformed into: How can animats pretend to have a "consciousness"? To explore this question, we propose to study the communication of humans and robots in terms of social communication. Some tentative steps are:

1) build a human-robot communication environment. We need a physical robot that can communicate with humans in a natural way.
2) study how the robot should behave to pretend to have a "consciousness".
3) measure how the human feels about the robot and its behavior.

(1) We expect that MIT will build a small robot as a device for the human-robot communication experiment. It should be small enough to make the people feel comfortable with the robot. The shape of the robot might be similar to a dog or a human infant. One idea is to build human infant robots and developmentally implement behavior, but the range is from birth to about one year. Thus, the project would not include bipedal walking or language communication.

(2) This is main theme of this project. We would like to frequently exchange information with AI lab people about this project. NTT will build a behavior engine based on reinforcement learning, which will be optimized to keep the interaction going as long as possible. Thus, the criteria of the learning is not to bore the counterpart. We expect that MIT will study and test various ideas.

(3) We propose using human biotic information to measure human feelings, and the subjective evaluation. It could be an objective method of finding a correlation between the human subjective evaluation and biotic information.

NTT has some equipment for measuring biotic information and they will study how to utilize the measure.

Tentative agenda:

1999

MIT will build two robots. NTT will use one of them to share the same hardware environment.
At the same time, it is preferable to build an onscreen agent as a simulator for the robot. The virtual agent would be helpful in finding out how the response to humans differs between real and virtual agents.

NTT will study methods for measuring the human emotional state via biotic information, e.g., EEG, SPO2, heart rate, blood pressure, and so on.

2000

MIT and NTT will organize joint experiments to measure human responses to the robots. We should evaluate real/virtual agents and some behavioral learning schemes such as reinforcement learning, which is compared to static or random behavior.

2001-2002

Based on experimental results, we would build a second version of the robots, focusing on critical features. The robot should have some emotional expressions using its face, arms, and legs. Then, we will conduct more joint experiments to measure human responses.