

Software Agents and Situatedness: Being Where?

Sanjay Chandrasekharan,
Cognitive Science Program,
Carleton University, Ottawa, Canada
schandr2@chat.carleton.ca

Babak Esfandiari,
Department of Systems and Computer Engineering,
Carleton University, Ottawa, Canada
Babak.Esfandiari@sce.carleton.ca

Abstract

Software agents are very different from mobile robots. For instance, they exist in a world quite unlike the world robots exist in. Also, software agents are not embodied. We investigate how these two differences reflect on agent design principles like situatedness and embodiment. We find that the insights gained from designing mobile robots cannot be applied directly to software agent design. Similarly, in the other direction, insights gained from designing software agents cannot be readily applied to mobile robots. This leads us to a notion of situatedness that is different from the notion of being in the world.

The Feynman problem solving Algorithm:

- *Write down the problem*
- *Think real hard*
- *Write down the answer*

But that's Feynman. For an ordinary Agent, the algorithm has a lot more lines, and a lot of the extra lines read the same: *ask others*. This social nature of human problem-solving has been rediscovered by Artificial Intelligence, and there's a new AI developing, Distributed Artificial Intelligence, or DAI. Distributed AI is an "inherently social study of artificial agents existing in some society or world" (Marsh, 1994).

Most of the work in DAI tends to be engineering-oriented. Very little has been done in linking work in DAI to higher-level models of cognition. In this paper, we try to relate results from DAI to one high-level model of cognition—the "intelligence-without-representation" model proposed by Rodney Brooks (1991). We explore the relationship between real-world agents and software agents, and try to find out how the fields relate to one other.

First, we offer a brief overview of the Rodney Brooks model. The Brooks model is largely based on the robot Herbert, developed at the MIT Artificial Intelligence lab.

The robot does exactly one thing: roam around the lab and collect empty soft drink cans. The robot uses an architecture Brooks calls "subsumption architecture." It consists of a number of distinct activity-producing subsystems or layers. These layers do not pass explicit, symbolic encodings or recordings of outputs between them. Herbert queries the environment often, using sensors, and "reacts" according to this "raw" input. The reactions are hard-wired. Unlike traditional systems like SHRDLU (Winograd, 1973), Herbert has no central system where a model of the world is represented. Another significant difference is the fact that Herbert operates in the real world, and not in a microworld, like SHRDLU, which operated in world of blocks.

Herbert's success is the cornerstone of Brooks' argument for "intelligence-without representation". He argues against the representation-driven systems of traditional artificial intelligence. (Here representation is taken as centrally-stored models of the world, like Minskian frames). Brooks (1991) makes four core points, quite convincingly, using Herbert as illustration.

Situatedness: Traditional AI uses an objective world model, provided by the programmer. This model is not robust, and fails when ported to the real world. Herbert uses the real world as its model. The world, therefore, is its own best model.

Embodiment: The physical world grounds the regress of meaning-giving. Traditional AI's symbols have no meaning, because they don't refer to the real world. They are empty symbols, referring to other symbols. Unlike these, the processing going on within Herbert has meaning, because Herbert is embodied, and it has a connection to the real world.

Intelligence: This is determined by the dynamics of the robot's interaction with the world. It's hard to draw a line between what is intelligence and what is environmental

interaction. There is no stand-alone intelligence. All intelligent systems are/should-be situated in some world.

Emergence: Intelligence is in the eye of the observer. You cannot identify the seat of intelligence within a system. Intelligence is produced by the interaction of many components, and can be determined only by relating the behaviours of the components with the environment.

All the points are probably right, when looked at from the point of view of creating real-world robots. But what about software agents?

The Web Wide World

Consider the mobile software agent system Gossip, distributed by a company named Tryllian (tryllian.com). The system has Agents that go out into the Internet and search out pieces of information you need, like web pages, pictures, text etc. In the process, it will travel far, and query other Agents, web pages, databases, and search engines. It will come back with the information in a time-frame set by the user.

The system is almost the same as Herbert in functionality. If what Herbert is doing is intelligent, what the Gossip Agent is doing is intelligent as well¹. It is mobile, and it collects random pieces of information from a world. Both Agents search, and bring back something. The only difference is that the Gossip Agent is searching an artificial world, while Herbert is searching the real world.

Given this Agent, it would be interesting to revisit the points Brooks makes using Herbert.

Situatedness: The Gossip Agent is situated. However, it uses a world model, a world model not unlike the one used in traditional AI — an objective world model, provided by the programmer. The difference is that the world is the agent's own, and not *ours*. The Agent has representations about servers, agents, databases etc. This world model is robust, and works well for the Gossip Agent.

Moral: for situatedness, the Agent's world doesn't have to be our world. *Some* world will do. And the Web, a world made up of symbolic programming structures, will do just

¹Andy Clark, in his book *Being There*, considers software agents as simulations of real world agents. We don't agree fully. Here's one reason why: for a process to be a simulation, there should exist an equivalent, "more-real", version of the process in the real world. However, what the Gossip agent does *is* the real web search process, and there doesn't exist another "more-real" version of it in the real world. More on this presently.

fine. That is, there exists an environment of symbolic structures, an *Umwelt* as Andy Clark (1997) calls it, where a situated Agent needs² a centrally stored world model. Therefore, situatedness does not rule out central representations.

Embodiment: the Gossip agent is not embodied (At least not in the sense of having a physical body. It can be considered as having a virtual body.). The world that grounds the regress of meaning-giving for the Gossip Agent is the Web³.

Moral: It is not necessary to have a physical body, or a link to the real world, for an agent to solve the symbol-grounding problem. If the Agent exists in a rich and complex world of its own, and it can act on this world as we can act on ours, the symbols mean something to the Agent. However, this does not undermine Brooks' argument against microworlds. The microworld model tried to make the Agent understand *our* world by feeding the agent a symbolic version of it. This does not work, because what the symbols mean to us is different from what it means (if anything) to the Agent. The story changes significantly when the Agent has a world of its own. In this world, the Agent can act on the symbols, and it is not clear what those symbols mean to us. The embodiment argument does not apply for a world like the Web.

Intelligence: is determined by the dynamics of the interaction with the world. This is true of Web agents as well. Without a complex and unpredictable world like the Internet, it would be difficult to classify the actions of the Gossip agent as intelligent. The Gossip agent underlines the point Brooks' makes: there is no stand-alone intelligence. All intelligent systems are/should-be situated in some world.

Emergence: intelligence is in the eye of the observer. This is true of Web agents as well. However, the argument about the identification of the seat of intelligence appears weak. In its current form, we can identify the seat of intelligence within a mobile agent system. It is within the Agent. The

²Things could change later, when servers and databases and other Agents could come equipped with Agent-detection mechanisms, and can volunteer information to the mobile Agent. This would allow the Agent to go around without a centrally stored representation of anything.

³In fact, a case can be made that we, as humans, have a symbol-grounding problem when we talk about URLs. They mean *something* to a server or an Agent, but what do they mean to us? What is the referent of a URL for us? What is its sense? Is the typing in of a URL a speech-act? We consider this a philosophical problem by itself.

action that counts as Intelligence does require the interaction of the Agent with an environment, but the environment (Web) in its current state does not play a very active role in that interaction. So the intelligence that “emerges” is not very significant in the case of the Web Agent. The role of the Agent in the interaction is far greater than the role of the Agent in Brooks' robots.

Moral: Emergence involves a trade-off between what's stored in the world and what's stored within the Agent. If there is very little structure in the world that the Agent can exploit, the agent will need to carry a lot more within.

The above list shows that Agents in the Web significantly weaken the arguments that Brooks' makes with Herbert's example. They also seem to show that AI is heading for a split — into artificial-world agents and real-world agents. There are overlaps in some areas — like the requirement for situatedness — but insights gained from one cannot be uniformly applied to the other.

Situatedness, A & B

From the above list, we would like to look a bit more closely at the notion of situatedness. Consider the honeybee, which lives in colonies similar in many ways to situated robotics' favourite insects, ants and termites. Typically, the behaviour of a foraging bee consists of two kinds of actions. The first one is foraging, finding food. The second one involves communicating the location of food to other bees using a complicated dance sequence.

The foraging activity can be considered as not requiring representation, because the paramagnetic particles in a bee's stomach responds to magnetic fields. This may allow the bees to locate a site and return to it⁴. A foraging bee can be considered as being situated in a sea of magnetic fields, and the bees navigate this sea much like fish navigating the ocean.

However, the second activity, namely the communication of a food site's location to other bees, *require* representation. In the absence of a marked route to the site (like in the case of ants) the bee needs to use some kind of symbolic structure (here the dance) to convey the location of the site to other bees. The sea of magnetic field does not provide the bee any kind of handle while communicating a site location to other bees. The bee is forced to resort to a symbolic structure (the dance).

⁴See Gould (1980) for a different view. It has been shown that bees in a hive evaluate a forager's dance for the “plausibility” of a site. This suggests the possibility of bees using a map-like information structure.

The moral that's usually drawn here is: communication requires symbolic structures. Therefore, activities that require communication cannot be performed by agents modeled using reactive architectures like Brooks'.

However, we wish to take one step back, and draw the following general moral: there are two kinds of situatedness.

Situatedness A: This is situatedness in the world, the situatedness that allows the bee to forage and Herbert to collect cola cans.

Situatedness B: This is situatedness among other agents, or social situatedness. This kind of situatedness is different from situatedness in the real world, and involves trade in symbolic structures⁵. It is this situatedness, *and not communication*, that makes the bee *create* symbolic structures. Communication and the creation of symbolic structure is the same process. They take place at the same time, and the former does not explain the latter.

Herbert, as we see it, is not situated in this second sense, because Herbert does not interact with other agents. Interestingly enough, most software agents *are* situated in this second sense. This is because software agents are situated in a world consisting of other agents, and, *by design*, need to communicate with them. The other agents could be humans (like in the case of the Microsoft paper clip agent) or artificial agents.

High level intelligence (anything from the bee up) involves an interplay between these two kinds of situatedness. If Herbert has to scale up to higher levels of intelligence, he will need to be situated socially.

The Web-World relation

Most software agents are situated socially, but most real-world robots are not. We have also seen that software agents and robots are situated in different worlds. So what relation, if any, exists between these two classes of agents?

As mentioned before, Andy Clark makes the claim that software agents could be treated as simulations of real-world agents. We don't agree fully, but we do think that software agents could be treated as partial simulations — in the limited context of understanding social situatedness. For instance, social protocols developed for software agents could play a constructive role in developing social protocols for real-world agents.

⁵ We view pheromones as symbolic structures. Ants and termites “communicate” using pheromones.

However, we think the lack of a shared world limits the amount of simulation possible. For instance, consider a human (or a robot) asking another human (or a robot) the directions to a hotel. The first agent can point her finger in a direction and say, "that way." This is possible because the two agents share the same world. This communication would be hard to simulate using software agents.

The software agents' lack of a physical body also limits the simulation. Consider a real-world agent that receives an antagonistic glare from another real-world agent. This communicative action is meaningful because both agents possess a body and are familiar with the correlation between body states and internal states. This communicative action cannot be simulated using a software agent. The nearest equivalent would be a KQML message with content "glare".

Conclusion

Contrary to popular belief, software agents and real-world agents (robots) do not form a common class. They are different in distinct ways, and theories developed for one has limited applicability in developing the other.

What is common to both software agents and robots, however, are the notions of situatedness and action-oriented representations, representations that simultaneously describe aspects of the world and prescribe possible actions (Clark, 1998). In fact, the Web Agent is just such a representation. For situatedness, the Web agent comes with a tailor-made world of its own. This means it can get away with its brittle, symbolic structures, while the real-world mobile robots would need to use adaptive programming techniques.

There also seem to be two notions of situatedness emerging—situatedness in the world, and situatedness among other agents⁶. The second situatedness (social-adeptness) has not been explored much by work in robotics. This situatedness requires postulation of internal states, representations, communication, etc. Without the second kind of situatedness, the Agents, even if they are situated in the world, will not be able to scale up to human intelligence, because we are situated in a world consisting of not only objects, but other Agents as well.

In conclusion, it appears that every world creates its own agent, and every agent its world. The real-world is only one such world. Contrary to what Clark and Brooks argue, there

⁶ It follows that there are two kinds of cognition—cognition of the world and cognition of other agents. They parallel the two major problems in philosophy—the problem of the external world and the problem of other minds.

is nothing special to be said about insights gained from the real world.

References

- Brooks, R. (1991). Intelligence without representation, reprinted in *Mind Design II*, Ed. Haugeland, J. (1997), Cambridge, Mass.: MIT Press.
- Clark, A. (1997). *Being There: putting brain, body, and world together again*, Cambridge, Mass.: MIT Press.
- Finin, T., Labrou, Y., & Mayfield, J., (1997). KQML as an Agent Communication Language, in *Software Agents*, J. Bradshaw (ed.), Menlo Park Calif.: AAAI Press.
- Gould, J. (1980). Honeybee Cognition. *Cognition* 37, 83-103
- Marsh, S. (1994). *Formalising Trust as a computational concept*, Ph.D. Thesis, Department of Computing science and Mathematics, University of Stirling.
- Winograd T. (1973): A Procedural Model of Language Understanding, in *Computer Models of Thought and Language*, R. Schank, K.M. Colby (eds.), New York, Freeman
- Wooldridge, M.J., and Jennings, N.R.,(1995). Agent Theories, Architectures and Languages: A survey. In *Intelligent Agents: ECAI-94 Workshop on Agent theories, Architectures and Languages*, M.J. Wooldridge and N.R. Jennings, (eds.)1-39, Berlin: Springer-Verlag