A Long Term Proposal to Simulate Consciousness in Artificial Life

Joseph D. Horton[1], Michael Francis[1] and Eckart Sußenburger[2]
[1]*Faculty of Computer Science, University of New Brunswick, 550 Windsor Street, Fredericton NB, Canada*
[2]*Bonn-Rhine-Sieg University of Applied Sciences, 53757 Sankt Augutin, Germany*

Keywords: Artificial Life, Simulation of Consciousness.

Abstract: Computers will soon be powerful enough to simulate consciousness. The artificial life community should start to try to understand how consciousness could be simulated. The proposal is to build an artificial life system in which consciousness might be able to evolve. The idea is to develop internet-wide artificial universe in which the agents can evolve. Users play games by defining agents that form communities. The communities have to perform tasks, or compete, or whatever the specific game demands. The demands should be such that agents that are more aware of their universe are more likely to succeed. The agents reproduce and evolve within their user's machine, but can also sometimes transfer to other machine across the internet. Users will be able to choose the capabilities of their agents from a fixed list, but may also write their own powers for their agents.

SCIENCE AND TECHNOLOGY PUBLICATIONS

1 INTRODUCTION

A great deal has been written about consciousness in the last few years. A few of the books: (Chalmers, 2010),(Deacon, 2012), (Dennett, 1991), (Grim, 2009), (Pinker, 2009). Yet not much is agreed upon. A quote we like by Sutherland found in (Gazzaniga, 2011), from the 1989 International Dictionary of Psychology (Sutherland, 1989).

> Consciousness: The having of perceptions, thoughts, and feelings: awareness. The term is impossible to define except in terms that are unintelligible without a grasp of what consciousness means. Consciousness is a fascinating but elusive phenomenon; it is impossible to specify what it is, what it does, or why it evolved. Nothing worth reading has been written about it.

We are not so pessimistic as the last sentence, although we do not think a great more is either known or agreed upon now. All the books referenced above have interesting things to say about consciousness. But they do not agree upon a definition, in fact most do not even try to define it. The problem as we see it is that there is very little with which we can actually do experiments. It is difficult to study consciousness in living animals. We can only watch and see what actions agents take in certain circumstances. Not everyone agrees that animals are conscious.

As for humans, they can tell us what they are conscious of, at least some of the time. But from a logical viewpoint, we cannot even prove to one another that we are conscious, that we are not "zombies" that are only feigning consciousness without actually being conscious. People usually assume that others are conscious, if only from politeness. Now if you believe that strong AI hypothesis (Searle, 1997) page 9, that the mind is just a computer program, there is no contradiction. Acting as if conscious and claiming to be conscious implies consciousness. There is no possibility of such a thing as a Zombie that acts as if it were a conscious being and says that it is conscious, and is not conscious. Searle has the opposite viewpoint and thinks that simulating consciousness in a machine is irrelevant. "... the behavior *by itself* is irrelevant" (Searle, 1997), page 204. We think that if some computer program tries to take over your bank account for its own purposes, whether it is conscious or not is irrelevant. Of course this is a different probem from what Searle was discussing.

We ignore the "hard problem" of explaining the conscious experience. We leave that to the philosophers. Our goal is to study consciousness, or at least the simulation of consciousness, from an experimental computational viewpoint. We intend to build artificial universes in which the agents demonstrate many of the properties of conscious beings, whatever people think of as properties of conscious beings. Of course the animats will not be able to discuss the real world, but we would like them to become able to discuss their own artificial universe, among themselves.

D. Horton J., Francis M. and Sußenburger E..
A Long Term Proposal to Simulate Consciousness in Artificial Life.
DOI: 10.5220/0004332903890394
In *Proceedings of the 5th International Conference on Agents and Artificial Intelligence* (ICAART-2013), pages 389-394
ISBN: 978-989-8565-39-6
Copyright © 2013 SCITEPRESS (Science and Technology Publications, Lda.)

This paper is essentially a project proposal, rather than a scientific research paper, although we are building such a system. The project is too large for a small research team. We need to have help to make it successful. The project was originally proposed in (Horton, 2010).

2 AN OVERVIEW OF THE PROPOSAL

2.1 Why now?

It may be that computers will be powerful enough to simulate the human brain in the next few years. We do not think that we are close to being able to simulate the human brain, although there are people like the Blue Brain project, that are trying. It is known that the human brain contains close to 10^{11} neurons and that neurons on average have about 10^4 connections between them (Azevedo and et al, 2009). Such considerations has led to speculation that a human brain could be simulated in 10^{16} floating point operations per second, which is about the same as today's fastest supercomputers. We think that this is likely to be a low estimate. There has even been some evidence that not all thinking occurs in the brain itself.

There are several reasons not to be unduly influenced if you think that this is a severe underestimate of the computational power of the human brain. The first is Moore's Law type arguments. The computational power of computers is doubling every year or so. Just wait a few more years, and the computational power will be there. Another is that most of the human brain is not being used for consciousness. Only 19% of the neurons of the human brain are in the cortex (Azevedo and et al, 2009), which is the most likely part of the human brain in which consciousness is to be found. A third reason, if you believe that animals are conscious, is that many relatively intelligent and social animals, such as the crow, have much smaller brains than humans, yet have to deal with a very complicated physical world.

Is it possible for a computer system of some kind to become conscious on their own? Not likely. But as algorithms become more nondeterministic, and more and more autonomous agents are being written and sent to do their jobs online, it would be nice to have some evidence, however weak, that the computer systems are not likely to appear to be conscious on their own. If it turns out that it is relatively easy to simulate consciousness, that would be a very important discovery.

2.2 How to Start towards Consciousness?

The only beings that we currently recognize as conscious were produced thru evolution. We propose to use an evolutionary approach. We insert autonomous agents into an artificial life environment, and try to stimulate them to simulate consciousness. We do not think that this will be an easy task. It is not at all clear that there is a good reason for consciousness to have developed in the real world. What advantage does it give to an animal, or to a species?

If an animal lives in a community of other animals that it interacts with in multiple ways, it is certainly to its advantage to be able to predict what one of its companions is going to do in certain circumstances. To do that it is useful for the animal to feel the same as its companion, and perhaps be able to predict what it itself would do in the same circumstance.

The first steps to move in this direction is to develop the four C's: conflict, co-operation, communication, community. They may not be essential for consciousness, but they seem to be a start. These should be possible to make in our proposed universes.

We have not yet thought much beyond the stage of building communities. Building communities is a goal for the intermediate term.

2.3 The Gaming Interface

We do not expect that we will be able to simulate consciousness by ourselves. We hope that we can enlist others to aid us. Nor do we expect that the goal will be obtained in a short period of time. We need a great deal of computational power for many years. It might take several decades. The internet becomes the only real source of such resources. Ideally we need millions of computers working for many hours a day. So like seti@home or the protein-folding online game, we want to attract non-scientists to work with us, and lend us

their computers for us to work on. We need something to give them in return.

Many computer/internet games have an artifical life component. Will Wright's SPORE is the most obvious example, and it even has an evolutioinary component. Several game types could be developed. Can you evolve a bug, or community, with some propetiy or properties? One could have contest between bugs/communities.

A very important point here is that the agents evolved must be available to the project, so that any useful property that is evolved or developed can be made available to everyone. Not only should the players allow bugs to evolve on their computers, but also they should be able to write code, what are called powers below. Probably a restricted scripting language is needed. These user-developed powers should slowly become available to other players. In fact the possibility of serendipitous interactions is an essential idea of this paper.

3 THE ARTIFICIAL LIFE MODEL

We are writing a system called **SOCIAL**, simulation of consciousness in artificial life. It has many of the features that we have thought about, but is far from complete.

3.1 The Universe

The universe should be as simple as possible, so that the brains will not require much resources to handle the physics of the universe. What does a universe require? A being must be located somewhere in a universe, so locations are required. From a location a being must be able to move to some other locations. This suggests that the universe should be a graph: locations are nodes; connections between nodes are edges. Adding geometry just makes the universe more complicated. Conceivably even simpler universes without locations are possible, with the bugs only knowing facts about the universe, but this would be harder to describe as a universe.

A graph structure is very simple, yet it allows for many different possible universes with many different properties. A graph can mimic a two, three or higher dimensional space, using any finite tesselation of space R^n. A grid graph could be used, to make for easy display. Or one could take any random set of points in the space, and construct the Voronoi diagram, with the cells being the nodes. We have experimented with very simple universes, such as a cycle. Or the universe can be the union of many different types of graphs, possibly with gateway nodes connecting the different components.

Each node can contain beings and resources. There are caps upon the contents of a node, decided by the experimenter. The first resource that we added, we called energy. We have toyed with the idea of adding conservation laws, but have not done so yet. It could be done by specifying that whenever a resource is used, an equal amount of some other resource is created. We have not considered any entropy law, that would require the universe to run down.

We have considered making an expanding universe, in that new nodes could be added at random times and places, but have not done so yet.

The beings must be given the power to move between neighboring nodes. Resources as we have imagined them cannot move, although this is a feature that could be added.

Time we have decided should be discrete, like space. In our experiments so far we have allowed the beings one action per time step.

3.2 Bugs

The beings in the universe are called **bug**s, beings whose universe is a graph. One can also think of them as being like viruses, although viruses are much more complicated than our bugs so far.

The bugs have multiple powers. Powers are abilities that a bug can use. Each bug has three different kinds of powers: actions, like move, eat, turn, pickup, drop, which they can do; sensors, like bugsensor (how many bugs are there in neighboring nodes?) that return information about their environment; and a brain, which decides what action they will try to do next. The powers that specific types of bugs have can be decided by the user. All powers can also change thru learning or evolution, decided by the user.

Bugs can use the resources at the node where they are. When a bug performs an action, it consumes some resource. In most of the simulations that we have tried, we have only had one resource which we called energy, but the system has no such restriction, and we have run simulations with multiple resources.

They also have internal resources, which they can obtain by eating. They can pick up and carry resources as well. Presently they generally only eat energy, although in some of the simulations they eat other bugs. If their internal energy resources falls to 0, then they starve to death.

3.3 Evolution

Evolution can occur in many different ways in the system. The bugs do not have genes per se. The set of powers that the bug has, together especially with the structure of the brain which decides what the bug does next, is the “genetic code”. The parameters included in the powers must also be included. Mutations can occur in any power, most importantly in the brain. It is possible that a new power can be added, if this particular type of bug is allowed to add powers. How the brain changes during mutation is determined by the type of brain that the bug has. More importantly, a bug can inherit different powers from different parents, and thereby get an evolutionary advantage over both parents.

The experimenter determines when mutations occur. The most obvious time of mutation is when the bugs reproduce. Both sexual and asexual are possible. When bugs reproduce asexually, only mutation occurs. When two different bugs reproduce sexually, two different brains must be combined. How this is done is determined by the type of brain that they have. Sexual reproduction has not yet been tested.

The bugs also can mutate as they live. One of the factors that can mutate is their mutation rate. As bugs become adapted, their internal mutation rate become much slower. This may appear to be an unusual way to evolve, but consider jumping genes (McClintock, 1950). The human brain appears to have jumping genes that change what genes are applied in new neurons, even in adults.

We have considered Lamarckian evolution as well as Darwinian evolution. If some useful sequence of actions gets used a lot, then the bug should be able to increase the chance of that sequence occurring or possibly make some of the actions less costly. We have not yet implemented this idea directly, but it is feasible within the system. This is just a form of learning.

4 AN IMPLEMENTATION

A prototype of SOCIAL has been implemented using Java. It includes a graphical user interface for building an experiment. The user can define the graph, choose the type, number, and locations of the bugs, and also of the type and amount of resources. The user can then start the simulation and watch it play out if the graph is displayable. They can also watch a graphical display of data, such as number of bugs, amount of resouces, average age of each type of bugs among others.

The implementation presently allow only grid graphs or graphs imported from a file to chosen as a universe. There are twenty-six different actions that a bug can perform; eight sensors that the bug can have, most with multiple parameters; thirteen brains to choose from, also with multiple paramters. Some of these powers are general purpose, but others were developed simply to see if the SOCIAL implementation could copy results in the literature.

4.1 Some Simple Brains

Many different brains have been programmed, but we only mention those more general ones that might be useful as parts of future brains. The first brain that we tried was the random chance brain; each action has a probability of been chosen. The chances of each action can change thru mutation. The second brain was the sequential brain that has a sequence of actions that get repeated in a loop. The sequence can change by adding/ removing actions, interchanging actions, or splitting the sequence. This is a relatively effective brain compared with many others. These two brains do not use information from the sensors.

Brains can also use facts about the world, to help decide what to do. Facts are discovered using sensors. The simplest one is a hierarchical ruleset (Brooks, 1986), where a rule is an if-then construct consisting of a possible fact, which can be true or false, and an action. More complex approaches include Affect Logic (Ciompi and Baatz, 2008) where creatures follow plans and the Fungus Eater (Toda, 1982), which combines routines and urges to create a behaviour. These three brains have been tested in (Sußenberger, 2013).

Any deterministic brain can be implemented as a decision tree. Evolving good decision trees does not seem to be easy, because decision trees with similar or even the same outcomes, can be quite distant from each other in a tree edit model. We do not know a good way to implement a deterministic brain that is comprehensive (can model any deterministic brain), stable (does not change too much with an evolutionary change) and yet is compact (size is polynomial in the number of possible facts).

4.2 Some Simulations

We have not yet done any experiments that are of scientific interest. For the most part we have looked at other simpulations in the literature, and tested that SOCIAL could also perform them. We emulated a simple simulation writen for Repast Simphony (North et al., 2006), in which “zombies” chase “humans”. There was no evolution.

We have modeled an ecosystem with three different bugs, which we called grass, sheep and wolf. A grass bug ate energy and could not move, but could reproduce into a neighboring cell; a sheep could move and eat only grass; a wolf could eat only sheep. The bugs evolve somewhat, as bugs with poor brains died quickly. If the parameters are chosen well, eventually the system can stablize and keep all three types of bugs usually. Sheep worked well with sequential brains; wolves with decision tree brains.

We have tested SOCIAL with some other test setups, for example the StupidModel (Isaac, 2011), which is a reference implementation for agent-based modeling platforms. It supports all the necessary setup, output and display features to complete that model. The behaviour of the agents in the model is

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

implemented using Actions, and the action sequence is implemented using a simple Brain. This demonstrates one of the core features of the SOCIAL model: these actions can be reused by any other experiment, or the same experiment can be run with a different brain. Other models, such as one based on the Prisoner's Dilemma (Kim, 2010) game, have also been implemented. This showed that cooperation can evolve.

4.3 Motivators

Bugs must be motivated to do things. The only such system that we have yet implemented is something akin to hunger. The bug knows, or has a sensor that can detect how much energy that it has, and if the energy level falls to 0 then the bug dies. So one thing that the bug has to do to survive is to eat if the energy level gets low. Successful bugs become *motivated* to find and eat food, whatever their food is, or at least they act like they are motivated. Similarly bugs that reproduce successfully only reproduce if their energy level is high. Thus a single stored variable gives successful bugs something like hunger and a desire to reproduce under the appropriate conditions.

The action of the energy level is somewhat analogous to the level blood sugar in animals. The simulation is much simplified, but other motivators for animals can often be driven by the levels of chemicals in the blood or in the brain. We have not tested this yet, but we can make general motivators just by giving the bugs more variables that rise and fall on the basis of their actions and/or their sensors. Unlike (Grand, 2000), who had a simulated blood stream in his creatures, and simulated chemical levels, we do not intend to specify what the bugs do with their motivators. We intend to let evolution decide.

Whether motivators are part of the brain or an entirely different power of the bug is not clear. Every motivator needs sensors. They also need something to change the level of the sensors, which could be a brain of its own.

5 THE FUTURE

We have a great deal of work to do. Here are a few of our ideas, some much more difficult than others.

1. We need to get SOCIAL to work across multiple machines. There will be security problems to solve here.
2. In the short term, although the Prisoner's dilemma model has shown that the bugs can cooperate in at least one case, we need more examples in which the bugs to cooperate and communicate. (Wagner et al., 2003) provides a comprehensive overview of recent research in the emergence of communication in artificial simulations.
3. Add motivators to the system.
4. Give the bugs memory. (How?)
5. We would like to see a brain based on the Cerebral Code defined in (Calvin, 1996a; Calvin, 1996b), or some other darwinian machine.
6. Design games to attract users.
7. Make SOCIAL an open source project, to which others can contribute.
8. Show the universe from a bug's viewpoint, as well as examine every aspect of a bug. This could be rather interesting in some cases. For example, a bug wandering in a four-dimensional universe could be rather interesting.
9. Investigate multicellular creatures, in which multiple bugs stick together.
10. Have an interface to create new powers for bugs, maybe a scripting language. Currently powers have parameters can that can be set, but new powers need to be programmed.

6 CONCLUSIONS

We are proposing a system to evolve consciousness. The graph universe is very simple, yet at the same time will require the bugs to be very flexible if they move to different graphs.

The bugs themselves are very flexible. A power written for any bug up to now can be given to any other bug. Eventually there may be powers that conflict. Certainly it is not easy to combine two different brain types, but otherwise powers do not seem to conflict.

The internet is an almost ideal medium for evolution. Evolution occurs best on small island environments, but also requires there to be mixing between them. Individual computers act like small islands. The occasional transfer of individual bugs from one machine to another can allow the powers and the brains to combine in novel ways, with possible serendipity.

REFERENCES

Azevedo, F. A. C. and et al (2009). Equal numbers of neuronal and non-neuronal cells make the human brain an isometrically scaled up primate brain. *Journal of Comparative Neurology*, 513:532–541.

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

Calvin, W. H. (1996a). *The Cerebral Code*. MIT Press, Cambridge,Mass.

Calvin, W. H. (1996b). *How Brains Think*. BasicBooks, New York.

Chalmers, D. J. (2010). *The character of consciousness*. Oxford University Press, Oxford.

Ciompi, L. and Baatz, M. (2008). The energetic dimension of emotions: An evolution-based computer simulation with general implications. *Biological Theory*, 3:42–50.

Deacon, T. W. (2012). *Incomplete Nature: How Mind Emerged from Matter*. Norton, New York.

Dennett, D. (1991). *Consciousness Explained*. Backbay Books, New York.

Gazzaniga, M. S. (2011). *Who's in Charge? Free Will and the Science of the Brain*. HarperCollins, New York.

Grand, S. (2000). *Creation: Life and How to Make It*. Harvard University Press, Cambridge Mass.

Grim, P., editor (2009). *Mind and Consciousness: 5 Questions*. Automatic Press.

Horton, J. D. (2010). An idea for a project: A universe for the evolution of consciouness. Technical Report TR10-203, Computer Science University of New Brunswick.

Isaac, A. G. (2011). The abm template models: A reformulation with reference implementations. *Journal of Artificial Societies and Social Simulation*, 14(2):5.

Kim, J.-W. (2010). A tag-based evolutionary prisoner's dilemma game on networks with different topologies. *Journal of Artificial Societies and Social Simulation*, 13(3):2.

McClintock, B. (1950). The Origin and Behavior of Mutable Loci in Maize. *Proceedings of the National Academy of Science*, 36:344–355.

North, M. J., Howe, T., Collier, N., and Vos, J. (2006). A declarative model assembly infrastructure for verification and validation. In S. Takahashi, D. S. and Rouchier, J., editors, *Advancing Social Simulation: The First World Congress*. Springer.

Pinker, S. (1997,2009). *How the Mind Works*. Norton, New York.

Searle, J. R. (1997). *The Mystery of Conscousness*. The New York Review of Books, New York.

Sußenberger, E. (2013). Tracing motivation in virtual agents. Master's thesis, Bonn-Rhine-Sieg University of Applied Sciences and University of New Brunswick.

Sutherland, N. S. (1989). *The international dictionary of psychology / Stuart Sutherland*. Continuum, New York.

Toda, M. (1982). *Man, Robot and Society*. Nijhoff, The Hague.

Wagner, K., Reggia, J. A., Uriagereka, J., and Wilkinson, G. S. (2003). Progress in the simulation of emergent communication and language. *Adaptive Behavior*, 11:37–69.

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS