# Franklin's New Infant Theory of Mind:
## Review of *Artificial Minds: An Exploration of the Mechanisms of Mind* by Stan Franklin

**Luciano da Fontoura Costa**
Coordinator, Cybernetic Vision Research Group
IFSC, University of Sao Paulo
Caixa Postal 369,
Sao Carlos, SP, 13560-970,
Brazil

luciano@ultra3000.ifsc.sc.usp.br



PSYCHE, 5(29), October 1999
http://psyche.cs.monash.edu.au/v5/psyche-5-29-costa.html

KEYWORDS: Mind, brain, Artificial Intelligence, cognitive science.

REVIEW OF: Stan Franklin (1995) *Artificial Minds: An Exploration of the Mechanisms of Mind.* MIT Press. xi + 449pp. ISBN 0-262-06178-3. Price: $US32.00; hbk.

"Her feet are tender, for she sets her steps,
Not on the ground but on the heads of men:"
(Homer)

## 1. Preamble

"This tour can be injurious to your present concept of mind", warns Franklin on page 17 of *Artificial Minds.* Indeed, the 16 chapters in his book are aimed at changing the reader's mind about mind. The situation is all the worse (or should I say better?), since Franklin's tour is not an unbiased one, as he admits on page 17--the tour stops are orchestrated so as to help in persuading the reader of some of Franklin's personal beliefs. The warning on page 17 is followed by a brief description of seven of these beliefs, which are developed in the book and recapitulated in chapter 14. Though overlooked by most previous reviewers of *Artifical Minds* (see http://www.msci.memphis.edu/~csrg/) who preferred to organize Franklin's book in terms of the three central AI debates which it addresses, these points are especially important, since they provide the bedrock for the book. As a matter of fact, after much thought and many consultations, Franklin even called his set of seven principles a "new infant theory of mind" (p. 417). Were he a little less shy, I guess his book would have had that title, i.e., "A New Infant Theory of Mind". For this reason, the author of the present lines found it more suitable to focus on these seven important points rather than to provide yet another review organized around the three central AI debates. Accordingly, the present review is organized in seven principal sections, corresponding to Franklin's framework.

One of the hallmarks of Franklin's approach is his commitment to physicalism. Yet, it should be stressed at the outset that Franklin does recognize the problem of qualia (Chalmers, 1996) as a very difficult issue, and even wonders about qualia inversion. However, his is the pragmatic alternative to postpone this discussion for later, in the hope that advances in physical theories of mind may lead to explanations of qualia. Chalmers, and others who have addressed this controversial issue before, would say that the key issue in mind and consciousness has thus been overlooked. Though Franklin's position is quite understandable and coherent with his physicalist approach, Chalmers's separation of the mechanical and qualia aspects will be considered in this review in order to better assess Franklin's contributions to the mind-body problem.

# 2. Mind is Continuous and Not Boolean

The first of Franklin's seven points is that mind is graded. Indeed, there is no reason why minds, at least in the mechanical sense, should be Boolean (i.e., exist or not)--must one say that somebody who underwent a lobotomy either no longer has a mind or has it to the same degree as before? Though evident, the fuzzification of mind is a subtle attitude that provides the key for endowing animals (chapter 3), and even machines (chapter 5), with some degree of mind, even if it is restricted to the mechanical aspect. It is a characteristic of our times that such fuzzifications of concepts have become more and more common. Intelligence has become graded, and so has mind. The Cretan says all Cretans lie. Assume what he says is a half-truth (or half-lie), and this classical paradox is gone for good (Kosko, 1997). Although even logic can be fuzzified, it is worth observing, for the sake of completeness, that there are some important exceptions, such as the fact of you being or not being a Martian. How about qualia?

Franklin's first point does not address the distinction between the mechanical and qualia aspects of mind. I would have appreciated it if he had spent a little more time on this stop of his tour, elaborating a little bit on why some people seem to have a point (acknowledged by Franklin) when they make the distinction between objective and subjective aspects of consciousness. Though the concept of qualia as the hard problem behind consciousness has become popularized only recently through Chalmers' work (Chalmers, 1996), that idea can actually be traced back at least to Johannes Muller

(Muller, 1843; see also Julesz, 1995). Perhaps the reluctance of many people to come to terms with such a slippery topic is accounted by the fact that qualia are virtually untreatable and smell of mysterianism. Yet, as the Empiricists believed, our personal and subjective experience of qualia may well be the only thing about which we can be completely certain about. At the same time, its definition is most elusive.

Actually, the problem with qualia is not only what they are or where they are felt, but *why* they are at all. In other words, all the interactions between humans and their environment could be explained in terms of information only, as they are in a program. Has your hand touched fire? Just change the value of a temperature variable in the system (human or machine), and specific processes will be triggered accordingly. Why on earth should such a phenomenon be accompanied by feeling? As a matter of fact, the above example is particularly fortunate, since it illustrates these two possibilities. Firstly, the sensory impulse is processed locally at the spinal cord, and a muscular command is almost immediately sent to the arm. Only later will the event be communicated to the brain, there producing pain. But there is a series of other body processes, including homeostatic regulation and heartbeat control, among many others, which almost never bother sending qualia-inducing messages to the brain.

If you happen to believe in strong AI, the explanation for qualia is easily forthcoming: information IS feeling. However, as qualia are inherently subjective and personal, the strong AI hypothesis (and also the opposite proposition) cannot be falsified. In other words, there is no way in which we can be sure about the subjective feelings of a robot or a bat. Even if we were capable of exactly mapping the neural state of the bat, there would be no way to prove beyond doubt that its subjective feeling of, let us say, red, would be equal to that of a human or even another bat. True, there is good chance that similar information processing is very likely followed by similar feelings, but this fact is impossible to verify *beyond any doubt.* However, it must be observed that mysterianism does not automatically follow from the fact that current science can not handle the qualia problem properly. On the other hand, it is important to acknowledge our limitations, in order to seek alternative theories and approaches. Should we forget that the true essence of knwoledge is to know also about what we do not know?

## 3. Mind is Aggregate Rather Than Monolithic

In how many parts or agents can a mind be divided? Rather than being a single process, Franklin believes mind is organized in terms of several agents specific to determined competencies (chapters 2, 10 and 16). In addition, the interaction between such processes would be band-limited, restricted by the possible correlations between agents. For instance, itch detectors in your feet do not have to exchange messages with your tongue (unless you are a dog).

Personally, I have always found it rather difficult to separate mental processes. In what sense could that be done? In space or time? Furthermore, if the structure of the mind is indeed hierarchical, and there is much neuroscientific evidence that at least our neocortex is (see, for instance, Zeki, 1993), should not we take this into account? To judge from our personal experience, high-level processes of mind ought to be predominantly sequential and unified. That is why I say I am Luciano, and not Luciano, Steve and Bill. At the same time, modern neuroscience has provided strong indications that lower mental processes, such as those in the primary sensory areas, are indeed modular and undergoing continuous parallel execution (e.g., processing color and shape at the same time) (Zeki, 1993). To be fair, it should be also conceded that even the highest processes of mind could be thought of as many, but succeeding each other in time. For instance, I am now using my writing processes but later will switching to my "how to power off the micro" resources. Roughly speaking, neocortical processes can be thought of in terms of a pyramidal hierarchy, the highest levels being more abstract and sequential, and the lowest closer to the sensory world and more parallel.

Back to Franklin's second point, there seems to be nothing wrong in hypothesizing that mind is aggregate, as long as such models help us to better understand some of its aspects. However, why not understand the organization of mind, like mind itself, as being "graded" along its hierarchical levels? That could have nicely harmonized the duality between unification at higher levels and ubiquitous distribution at lower levels.

# 4. Mind is Enabled by a Multitude of Disparate Mechanisms

Related to the previous section, this third component of Franklin's theory concerns the homogeneity of mind processes. Speaking in Boolean terms, the mind agents could in principle be all the same or not. Which is correct? Being not fond of unified theories of cognition, Franklin favors the latter option, i.e., the modules of mind are distinct. The reasoning goes something like this. Although "the unified theory of cognition people" (p. 417) believe the mind to be a collection of identical modules, capable of performing different activities, such flexibility would be biologically expensive.

If we constrain ourselves to the lower hierarchies of the mind, and consider the anatomic-functional modules which have been identified by modern neuroscience, the question of whether such chunks of cortical tissue are equal can be discussed in a more down-to-earth way. Although we could sit and wait for the answer to be provided by forthcoming advances in functional neuroscience (and as far as neuroanatomy is concerned, the modules are indeed very similar), we can hypothesize that the modules are reasonably similar at the *anatomical* level (e.g. positions of the neuron bodies, types of cells, and distribution of synapses), but quite diverse regarding the *functional* level; a semantic rather than syntactic distinction. Another important lesson about this question can be gained by considering biological evidence, namely the process of neurogenesis and apoptosis. People lose about half of their neurons during their early life, a drastic phenomenon believed to be necessary for the proper tuning of the neural connections in terms of external and internal stimuli, since such connections can not be completely specified by the genetic code. Such reasoning can lead to the antithesis of Franklin's belief regarding agent variety. Let us be more explicit: neural development is ready to waste lots of cells and connections in order to economize genetic code. As a consequence, the modules are anatomically very similar at birth, only later becoming specialized, as the synaptic connections are tuned by the environment and internal mechanisms. Yet another indication supporting the homogeneity of hardware hypothesis is provided by the fact that neural functions can be, to some limited degree, relocated in the cortex. I would bet that "the unified theory of cognition people" are at least partially correct regarding this question.

## 5. The Overriding Task Of Mind Is To Produce The Next Action

Basically, this point is closely related to the reasons why agents act, the proposed answer having to do with drives defined by evolution or design. I do not think many people would argue about that, but the question does get more evolved at lower levels of abstraction, where neural subsystems may operate concurrently and/or competitively (e.g., Dennett's multiple draft theory). Yet, Franklin has found this simple assumption to be surprising and useful. It allows, for instance, questions such as "Do we humans have any built-in drives not found among our primate cousins, the great apes?" (p. 412).

## 6. Mind Operates on Sensations to Create Information for its Own Use

Franklin believes information is not out there, but is only created as a result of minds "processing information taken in from the environment through the senses" (p. 413). The same could be said about the world, leading to solipsism. As a matter of fact, the whole question depends entirely on what you mean by information. Consider the scientific definition of information in terms of entropy (is there any sensible alternative?): roughly speaking, entropy is a measure of information reflecting the number of possibilities of an outcome (Shannon & Weaver, 1949; Haken, 1977).

The information acquired by knowing the results of throwing a coin is smaller than that for throwing a die. Being dynamic and stochastic, the world is unquestionably full of information. So what was Franklin after with this point (chapters 8 and 16)? It seems to me that his belief arises from a confusion between awareness about information and information at large. Indeed, what Franklin does seem to be arguing is that our perception of the world is filtered by our senses--no two persons experiencing exactly the same world. I think that is indisputable. However, it is not that information only exists inside us, but only that the information we process is selected from that in the outside world. After all, there are other people out there, creating and storing all sorts of information. Information being ubiquitous, perhaps it is qualia that only start existing inside us.

# 7. Mind Uses Prior Information (Memories) To Produce Actions by a Reconstructive Process Rather Than by Retrieval

Are human memories stored in standard-computer-memory-like devices? If that is not so, what are the alternative mechanisms? Welcome to Franklin's sixth stop.

I would like to start discussing this question by recalling that, as cognitive science and neuroscience have clearly indicated, there are many kinds of memories, such as declarative, procedural, episodic, reflexive, etc. (Kandel, Schwartz & Jessel, 1991). Which one was Franklin assuming? Although not complete or very clear, such a classification of memory in several classes is relevant because it indicates that different mechanisms may be involved in each case, even though physically memory seems to be constrained, at the microscopic scale, to alterations in the synaptic connections (both in degree as well as in existence) and spatial densities of biochemical elements. Consequently, different memories would differ in terms of the place in the brain where they are implemented and by the mechanisms used to store and retrieve the respective information.

Franklin presents his ability to reconstruct much of the information about his living room just by heart (incidentally, it is said that most people cannot remember the exact color of their bedrooms or their underwear) as evidence that retrieval involves recognition, categories, and spatial relations, in which I believe he is completely correct. Association is an important mechanism in memory and mind (I would say it is the most important), one which should be receiving more attention from scientists. However, memory retrieval at lower hierarchical levels may involve different mechanisms, such as the effect of the current synaptic strength. No associations here, just the modulation of the electrotonic potential caused by the current status of the synaptic connection, a process not completely different from charge storing in dynamic computer memories. Strictly speaking, synaptic strength is very likely determined by association-based rules (e.g., Hebbian rules), but the "retrieval" of the synaptic value is direct and does not involve associations. By the way, biochemical gradients are also known to play an important role in storing and retrieving memories, adding a further important dimension not discussed in detail by Franklin (see Black, 1994, for additional information).

# 8. Mind, to Some Degree, is Implementable on Machines

Franklin and the author of the present lines seem to fully agree on this point, provided we restrict ourselves to "mechanical" minds. Indeed, perhaps his observation "to some degree" could even have something to do with qualia, but this is only speculation. It has always puzzled me that some people doubt the possibility of having computers behaving like *typical* humans. After all, humans are so often rather limited and prosaic creatures. While we are far from implementing robots which behave like typical humans, there is no compelling reason why this would not be achieved in the future. It all seems to be a question of technological progressions, not scientific quantum leaps. The more challenging problems with mechanical minds very likely lie elsewhere: qualia and creativity. As a matter of fact, Franklin has interesting things also to say about creativity, in chapter 13 (see also Sternberg & Davidson, 1995).

Another relevant aspect related to the question about the computational implementability of mind regards the means through which this could be accomplished in practice. The Second AI Debate, addressed in chapter 7, deals exactly with that. Although the issue is far from settled, there are many people, Franklin and myself included, who believe that connectionism is not inherently richer than symbolic AI. To begin with, it is virtually impossible to imagine a single situation involving observable human activities and thought which could not be implemented in a traditional Von Neumann architecture, which are "universal". Even parallel programs can be simulated in that way. But then the connectionist movement was not exclusively motivated by the inherently concurrent computation capabilities of neural networks, but also by fault tolerance and, mainly, the possibility of automated learning. However, little seems to have been accomplished even regarding such potentials, and there is a current trend to return to AI methods, especially reinforcement learning (see, for instance, Ballard, 1997).

Over the last decades, many computational paradigms have been disputing the privilege of being the exclusive means to implement minds, one of the latest trends being based on dynamic systems (see Globus, 1992, for instance). But, as Penrose has observed (Penrose 1997), all these paradigms can indeed be simulated in a standard personal computer. As a matter of fact, much of what is known about dynamic systems and chaos has been suggested and verified precisely through numerical simulations in pretty standard machines. What indeed transpired to be true is the fact that these different computational trends may in the end prove to be particularly effective for treating specific aspects of implementing mind in a machine. Neural networks, for instance, provide models for neurons at the intermediate spatial scale and allow the intense parallelism underlying the lower hierarchical stages of brain computations. Dynamic systems, in turn, can add interesting insights about microscopic neural phenomena (e.g., Guckenheimer and Labouriau, 1993; Fohlmeister and Miller, 1997), including chaos in neurons, time-evolving representations, and also to larger scale processes, such as cognition and selective attention (e.g.,Thelen & Smith, 1994; van Leeuwen, 1997; Kelso, 1997).

# 9. Concluding Remarks

Having commented specifically on each of the seven components in Franklin's new infant theory of mind, which I believe to be the real motivation behind his book, I would like to concluding this review by commenting on some other aspects of *Artificial Minds.*

In general, I found *Artificial Minds* to be informal and clearly written, incorporating a wide range of concepts and case-studies, ranging from AI to connectionism and artificial life. He seems to be very sincere, duly acknowledging his doubts and biases. In addition, though acknowledging the qualia issue, he adopts a most welcome physicalist approach throughout the whole work.

Although written with the lay-reader in mind, I believe the treatment of some of the concepts (such as dynamic systems) will not be enough to provide a really sound basis. But neither could any book with about 400 pages be expected to do so, for many of those concepts involve relatively complex mathematics and are generally accessible only to those with university level mathematics who are willing to read specific books. Nevertheless, the conceptual presentation of such topics is careful and fine, and should be enough to provide a good overall introduction to at least some of the concepts and jargon, facilitating further studies by the interested reader. Nevertheless, I think the book will be more profitable to those who are already involved in one or more of the included areas and want to know more and become up to date about related researches. One of the nice features of Franklin's book is the integrated review of many concepts and theories previously constrained to the specialized literature on artificial intelligence, associative neural networks, etc. The three AI debates - namely "can we expect computers to think in the sense that humans do" (First Debate); the "family squabble" between the AI and connectionist approaches (Second Debate); and the role of representations (Third Debate) are also nicely covered in specific chapters.

Regarding my personal experience with the book, I found it most interesting to have the possibility of touring in a short time such a baffling variety of different theories of mind which have been proposed. To me, this Babel provides a most definite proof that, in spite of the many scientific advances over the last three decades, mind is still an elusive concept. Yet, I do not believe such a variety of positions is necessarily bad. Rather, it seems to be the very important mechanism through which a more complete and realistic view of mind is slowly emerging. Not only have different positions allowed difficult problems to be treated relative to many distinct reference points, but information from the most varied areas have more and more been merged in a fabulous dialectical process of inter- and multidisciplinary integration. Variety is one of the essences of creativity and evolution, and should be highly prized in an era presenting globalizing tendencies such as ours. Incidentally, one of the few drawbacks in *Artificial Minds,* as already observed in T. Sejnowski's review (see "Reviews of *Artificial Minds*"), regards the fact that Franklin has paid relatively little attention to cognitive theory and neuroscience, being more restricted to mathematical and computational issues. However, I do believe that to proceed otherwise would have required the sacrifice of some of the interesting topics covered, and/or the undesirable extension of the book in length. Yet, the biological background has been playing such a special role in unveiling mind that it should have received a little more attention. As the reader may have noticed, many of the debates about Franklin's points in the present review relied strongly on biological evidence.

## Acknowledgments

The author is indebted to R. Poznanski for indicating Globus' paper, and to FAPESP (Procs. 94/3536-6, 94/4691-5, 96/05497-3) and CNPq (301422/92-3) for financial support.

## References

Ballard, D. H. (1997). *An introduction to natural computation* Cambridge, MA: MIT Press.

Black, I. B. (1994). *Information in the brain: A molecular perspective.* Cambridge, MA: MIT Press.

Chalmers, D. J. (1996). *The conscious mind.* Oxford: Oxford University Press.

Fohlmeister, J. H., & Miller, R. F. (1997). Impulse encoding mechanisms of ganglion cells in the tiger salamander retina. *Journal of Neurophysiology, 78,* 1935-1947.

Franklin, S. (1996). *Artificial minds.* Cambridge, MA: MIT Press.

Globus, G. G. (1992). Toward a noncomputational cognitive neuroscience. *Journal of Cognitive Neuroscience, 4(4),* 299-310.

Guckenheimer, J., & Labouriau, I. S. (1993). Bifurcation of the Hodgkin and Huxley equations: A new twist. *Bulletin of Mathematical Biology, 55(5),* 937-952.

Haken, H. (1997). *Synergetics: An introduction.* Springer-Verlag.

Julesz, B. (1995). *Dialogues on perception.* Cambridge, MA: MIT Press.

Kandel, E. R., Schwartz, J. H. & Jessel, T. M. (1991). *Principles of neural science* (3rd ed.). Appleton and Lange.

Kelso, J. A. S. (1997). *Dynamic patterns.* Cambridge, MA: MIT Press.

Kosko, B. (1997). *Fuzzy engineering.* Prentice Hall.

Muller, J. (1844). *Handbuch der physiologie des menschen* [Manual of human physiology]. J. Holscher.

Penrose, R. (1997). *The large, the small and the human mind.* Cambridge: Cambridge University Press.

Shannon, C. & Weaver, W., (1949) *A mathematical theory of communication* The University of Illinois Press.

Sternberg, R. J. & Davidson, J. E. (1995). *The nature of insight.* Cambridge, MA: MIT Press.

Thelen, E. & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action.* Cambridge, MA: MIT Press.

van Leeuwen, C. (1997). Stability and intermittency in large-scale coupled oscillator models for perceptual segmentation. *Journal of Mathematical Psychology, 41,* 319-344.

Zeki, S. (1993). *A vision of the brain.* Blackwell.