Book Review

Michael Arbib’s *The Metaphorical Brain 2: the sequel?* ☆

Stephen José Hanson 1

Psychology Department, Rutgers University, 101 Warren Street, Newark, NJ 07102, USA

Received 29 December 1994; received in revised form 31 May 1995

1. Sequels

Hollywood has a penchant for making sequels which are generally related to the box-office performance of the parent. Unfortunately, like a photo-copy that has been copied one too many times the plot, the characters and the general value of the movie begins to fade. As a good rule of thumb, it is usually advisable to avoid any movie, play, or book ending in a numeral. I thus, began my reading of Michael Arbib’s new book *The Metaphorical Brain 2* with some trepidation. I had at one time read *The Metaphorical Brain,* published now some 20 years ago (I read it in the 1980s), and have fond memories of Arbib’s ability to transcend so many fields and find some common ground on which neuroscience and AI could interact. In 1972, this was either a courageous or insane essay to write. In these years of the “AI summer” (1970–1981), this sort of blatant materialism could be met with years of rejection by funding agencies and journals. Besides Michael Arbib, Stephen Grossberg, Jim Anderson and Jack Cowan there were only a few others who were keeping the Neural Network candle burning during these dark years.

2. What kind of book is this?

In the mid-1980s, of course, Neural Networks flashed into a nuclear fire. And almost any book related to computational neuroscience or neural computation would find some audience lurking somewhere. *The Metaphorical Brain 2* is no exception. The question of course, nowadays is do these books deserve any audience? There are many distinctions to draw in neural network volumes. Introductory texts are a typical category that many authors are trying to target. Partly, because there has not yet been a coherent, clear voice

---


1 Email: jose@psychology.rutgers.edu.
in this market to congeal the masses of neural-net-cyber-punk-hackers and perhaps more critically there appears to be huge market which crosses at least 4-6 lucrative disciplines. Academic markets always being somewhat slim, it is important to be able to reach a market that is somewhat more popularized, you know, unions of various well-known genres like: “fuzzy neural networks”, “cybernetic neural brain systems”, or the ever popular, “artificial fuzzy neural life”. There are of course serious contributors to this category, I note in passing Hertz, Krogh and Palmer’s, little trendy text that provides an honest if not earnest view on the neural computation in the 1980s and 90s—unfortunately the book lacks a plot and tries to lure the reader through with just a chronology. Also I would be remiss if I did not mention Churchland’s and Sejnowski’s *Computational Brain*, this is a tightly written text reminiscent of Richard F. Thompson’s classic, *Physiological Psychology* spiced up with computational examples at just the right moments. I recommend both.

A second sort of the neural network text is omnibus neural network collection. This is an ever venerable genre which can be quite coherent if the chapters are co-authored by the same editor, or editor’s colleagues or students. Stephen Grossberg is well known for this genre and sets the pace by publishing such volumes once a year. Prototypical of this kind of text is his *Neural Networks and Natural Intelligence*. But the runaway classic in the genre is of course, the PDP volumes which represented some sort of zeitgeist arising from the southern California surf and sand, which of course set the stage for the ensuing neural net revolution we are now experiencing. Another volume of this sort which attempted to juxtapose computationalists with neuroscientists was *Connectionist Modeling and Brain Function*, edited by this reviewer and neuroscientist, Carl R. Olson. In this case any perceived coherence arises from fortuitous planning as opposed to chain-gang writing.

Finally, the third sort of text which presents a point of view and defining themes and underlying plot-lines is more polemical—but I mean that in a good sense! Really! Michael Arbib’s book falls more in this category. At times it seems to be an attempt at more of textbook genre, as there is considerable introductory material concerning, neuroscience, control systems, and even cognitive psychology. There is a conscious effort at juxtaposing neuroscience with computation and mathematics. Its almost as if when the familiarity of the biological concepts you are reading grows to some tolerable level, you are then drug, sometimes feet kicking and arms flailing into some seemingly coherent field of engineering or mathematics. Like some bad medicine you have to take in order to appreciate the subsequent material Arbib relentlessly shoves your nose into it as if he is saying “... but, this will be good for you! Just wait!” In this way, he migrates beyond simple textbook structures to a more personal view of field involving computation, brains and mathematics.

Because the book must be appreciated at a proper level of abstraction, one will not learn neuroscience or computer science in this book. Nonetheless, there are some basic themes that tie the chapters together and provide a backdrop to the diversity to which one is subjected. These themes revolve around “Schemas” and Frogs ... yes, Frogs.

3. Schemas

The lynch-pin of the book revolves around a concept of “Schema”. Arbib has clearly thought this concept through. And historically has some claim to its use in various
communities. Ironically, it was first used by Neuroscientists as a sort of intervening variable as in spatial maps for the body. It was recycled by Bartlett in his seminal work on memory and finally it entered the mainstream cognitive psychology with Neisser's use of it in a model of memory encoding/retrieval which first posited the critical relationship between perception and memory. Arbib has bigger and better plans for Schemas and begins a coherent account on pp. 36-45. Here we learn that schemas can serve many functions and provide the underlying constituent structure of the brain. Hence, neurons are too low a computational level and brain "centers" are too high. Worse yet, both kinds of structures are in fact static and provide no virtual computational level of analysis. Hence, we need Schemas. Schemas are the virtual lexicon of the brain, they are built up from features and are instantiated as specific instances—a type/token relationship.

Perhaps, though, lynch-pin is a bit too strong. Although true, that the book rests upon the notion of schema as some basic computational element, there are promissory notes throughout the early part of the text not paid off later. Instead we get a scholarly discussion of the topics and concepts, but typically without much resolution. For example, one might posit different types of schemas, ones perhaps that vary in abstractness or perhaps have some domain or topic dependency and then this begins to look like a vocabulary, dictionary or even simple taxonomy, unfortunately, nothing like this is spelled out in the book. So although there is a discussion of the "variety of schemas" it tends to be more historical or context based. We are given the analogy of computer programs, which unfortunately does not provide enough constraints to understand what the bounds on the "varieties of schemas" might be. Nor does it give us a hint on how schemas might be overall organized into some structure. Although I am not suggesting Arbib do what Schank did, and populate the world with catchy terms like "Scripts" "tops", "mops" and whatever. But rather what seems to be avoided here are serious issues of Generalization ... given a schema how do I calculate its similarity to other schemas and other input data? What determines the level of abstractness of a given schema? When does a schema have precedence over sensory data?

Schema acquisition might be an appropriate place to start understanding the structure of schemas and the nature of their organization for brain function. Afterall, the way schemas are learned and then subsequently guide learning might be diagnostic of the way they are used and organized by the brain. Surely, the subtle inter-play of perception and memory would be a key element of schema acquisition and consequent usage. And once again although schema acquisition is discussed we are given a coffee-table level account of Piaget, "habits" and other relevant but high-level concepts that don't lead either to a research question or a computational speculation of what Schema learning might look like.

I suspect Arbib can't be blamed here for taking on ideas that are tough, ones that are bound to be difficult to backfill with enough detail to make useful. Having the courage and vision to promote such ideas is its own reward. And in any case, there are many juicy vectors here, ones that will incite young researchers in good directions.

One structural concern with the book has to do with the conceptual level at which some material is presented and the hands on level of detail one gets with other material, Hence, the frog issue.
4. Frogs everywhere!

I exaggerate. Nonetheless, 10 references on frogs peppered throughout the text (starting on p. 14 and ending on p. 359—there is even a reference on p. 129 not in the index!) in a book on the brain and computation may seem excessive to some (Not to me!). Well what’s my point. I don’t really mean to disparage frogs here. I even think the examples of the frog’s visual system are presented in a relevant useful way. Nonetheless, the level of detail in these examples and many others, seems to be quite distant from the historical and contextual level presented elsewhere in the book. This leaves a gap in the material presented for support of the general thesis behind the book and the thesis itself which is not adequately detailed for further speculation or expansion. So although the frog’s tectum gets fair representation in the book, it is hard to struggle back to the schemas and computational backdrop that the book promotes.

5. ‘I’ll bee baack’—system T1000

So does number 2 begat a number 3? And so on. I think not. I think the “2” needs to be read as “also”. In many ways MB number 1 could be seen as the abridged version of MB 2. And as I mentioned earlier number 1 was great and there are probably diminishing returns on reading number 2 if you’ve already read number 1. Although the messages have been updated, very little about the coherence of neural networks receives much confirmation from the MB 2. The revolution in recent years, may have begat interest in the updating of the book, but book seems to be unaffected by the revolution. Indeed, in a late chapter on Memory and Learning we find references and discussions to PDP, Kohonen etc. almost like an afterthought. Perhaps it has to be there, but its relation to schemas and the like are mysterious at best.

Nonetheless, I suspect if there would be an MB 3 I would read that too. Not just because of the sheer breadth and underlying message of the book of which there is a lot to agree with. But mostly because I think still provides some hope for coherence and direction within this diverse, yeasty field we call neural networks.