The problems of cognitive dynamical models

Jean Petitot
EHESS, Mathematical Center, 54 bd. Raspail, 75 006, Paris, France and CREA, Ecole Polytechnique, 1 rue Descartes, 75 005, Paris, France.
petitot@poly.polytechnique.fr

Abstract: Amit's "Attractor Neural Network" perspective on cognition raises difficult technical problems already met by prior dynamical models. This commentary sketches briefly some of them concerning the internal topological structure of attractors, the constituency problem, the possibility of activating simultaneously several attractors, and the different kinds of dynamical structures one can use to model brain activity: point attractors, strange attractors, synchronized arrays of oscillators, synfire chains, and so forth.

The main idea of Amit's article is to identify the psychological concept of Hebbian reverberation with the dynamical concept of attractor and to ground a computational theory of mental representations on this basis. This ANN (Attractor Neural Network) perspective on *cognition* raises difficult technical problems.

1. A historical remark. In the context of a physicalist approach of connectionist networks Amit has been one of the first to emphasize the interest of a true dynamical perspective, but prior mathematical work already exists. As far as I know, it was Zeeman (1965; 1976) who in the late sixties introduced the cogent and seminal idea that one could use dynamics to bridge the gap between the small scale neural level and the large scale psychological one. According to him, mental contents could be modeled by attractors of neurally implemented dynamical systems and the temporal flux of mental representations by sequences of bifurcations of attractors.

The main limitation of these early dynamical models compared to current ones was that the effective neural dynamics were unknown. Deep theorems showing that there exist *universal* prototypes ("normal forms") for the relevant dynamical structures (e.g., universal unfolding of singularities of energy functions) nevertheless made it possible to work out dynamical cognitive models (e.g., for categorical perception or image processing; see Petitot 1989; 1995). These early models already showed the kinds of technical problems facing a dynamical cognitive theory. I wish to stress some of them here.

2. The internal structure of attractors. In an ANN perspective, what can be the internal structure of attractors? In the case of symmetric weights (Hopfield model) there exists an energy function and attractors are therefore point attractors. But in the case of asymmetric weights one can get topologically complex attractors; and routes towards chaos (such as the doubling period subharmonic cascade) are observable (see Sompolinsky et al. 1988, Dayon et al. 1993, or Renal & Rohwer 1990 work).

The fact that the topology of attractors is in general highly nontrivial could be essential for understanding the semantics of mental contents.

3. Attractor syntax and the constituency problem. If one identifies a mental representation with an attractor, then one must take up the challenge of modeling dynamically the syntactic constituent structures. In their replies to Smolensky's fundamental 1988 *BBS* paper "On the Proper Treatment of Connectionism," Fodor & Pylyshin (1988) and McLaughlin (1990) have shown dramatically that connectionism lacks any correct account of constituency and compositionality.

They were essentially right concerning a very weak (PDP) form of connectionism (see Petitot 1991), but this is no longer the case if one adopts a stronger form. Indeed, according to Thom (1980), it is possible to work out an "attractor syntax" using bifurcations (more precisely universal unfoldings of singularities) in an original way (see Petitot (1995). What is Amit's response to the constituency problem?

4. The simultaneous activation of several attractors. Another difficult problem concerns the simultaneous activation of several attractors. Indeed, a dynamical system can only be in a single asymptotic state at any time. If one uses attractors for explaining how several representations can be tagged by a stimulus and can be self-maintained in memory until further processing, then one faces a problem. One possible solution could be to activate attractors of a slow/fast dynamical system sequentially, but such an attractor-chaining is not satisfactory because in many cognitive tasks the co-activation must be done in parallel.

To tackle this difficulty, it seems that some sort of symbolic computing is unavoidable.

5. What kind of dynamics? My last point concerns the different kinds of dynamical structures one can use to model brain activity. Some solutions to the constituency problem (the so-called "binding problem") use results of experiments on cortical oscillations. Since the pioneering findings of Gray and Singer (see e.g., Engel et al. 1992) much work has shown that neural modules (e.g., orientation columns in the primary visual cortex) behave as oscillators and that their synchronization is stimulus-dependent and codes for the coherence of the stimuli. According to the "labeling hypothesis," the constituent structures of mental representations can be retrieved using as labels for the constituents the common phase of the synchronized oscillators they are implemented in. The problem of synchronizing of weakly coupled oscillators is a very difficult one which can be tackled only with sophisticated tools of statistical physics (Kuramoto and Nishikawa 1987: phase transitions, Daido 1990: renormalization group) or of qualitative dynamics (Kopell and Ermentrout 1990).

Even though these results are controversial (they can be significantly improved using pulse-coupled oscillators), they show that many kinds of dynamical structures can be relevant: point attractors, strange attractors, synchronized arrays of oscillators, and so forth.

It would also be interesting to see in what *exact* sense Miyashita's results confirm the ANN hypothesis. Indeed, similar experimental results can be interpreted in a different way. For example, Bienenstock (1994) uses the concept of *synfire chains,* that is, neural modules supporting wave-like patterns of activity. According to Abeles (1991), synfire chains can reverberate in different modes, depending on the context of their activation. They can also learn to recognize sequences of synchronized volleys and can dynamically bind with each other via synchronization. They might accordingly represent another major mechanism for local information processing in the cortex. What does Amit think of their links with ANN models?