Neurodynamics and the Revival of Associationism in Cognitive Science

Pim Haselager
Cognitive Science, NICI, Katholieke Universiteit Nijmegen
Montessorilaan 3, 6525 HR Nijmegen, The Netherlands
Email: w.haselager@nici.kun.nl

Abstract

The recent interest in Dynamical Systems Theory (DST) is part of a growing tendency within cognitive science to replace the structure-sensitive processing of representations by associative processes of a neurophysiological kind. More specifically, neurodynamical work on categorization can be located within an associationistic tradition because of its attempt to explain behavior on the basis of correlated processes in the brain. Although associationism is a tradition not favored by traditional cognitive science, a recognition of DST's associationistic tendency seems called for. This does not imply a weakening of DST's relevance for the study of cognition, but rather a strengthening of classic associationistic ideas.

Introduction

> "I shall not be able to execute, with any accuracy, what the reader might expect of this kind, in respect of the doctrines of vibrations and association, and their general laws, on account of the great intricacy, extensiveness, and novelty of the subject. However, I will attempt a sketch in the best manner I can, for the service of future inquirers." (Hartley 1749/1970, I, i, 1, p. 6).

Representations have been a central element of cognitive science since its beginning, but recently disagreements regarding the nature of and even the need for representations seem to flourish. The classical position is to consider information to be symbolically represented and to be processed on the basis of structure-sensitive rules (Fodor 1975; Newell 1981). Connectionists prefer to speak of the distributed representation of information by means of weights and activation patterns (Van Gelder 1991, 1992). Classicalists and connectionists continue to discuss the theoretical and empirical merits and deficits of their proposals. Over the last few years, however, a new participant has entered the debate. Proponents of the dynamical systems theory (DST) argue that cognitive models can be developed on the basis of a minimal or even no use of representation and computation (a.o. Brooks 1991; Edelman 1992; Thelen & Smith 1994; Van Gelder 1995). In this paper I will claim that DST intensifies the shift of attention, that started with connectionism, from structure-sensitive computational processes to neurophysiological associative processes. I will argue that the emergence of what I propose to call 'neurodynamics' (DST applied to the study of the brain) implies that cognitive science is developing in an associationistic direction, but that it is far from clear whether this is something to be regretted.

Going Beyond Classical Cognitive Science and Connectionism

The emphasis on the syntactic structure of representations and the structure-sensitive nature of information processing is an essential characteristic of classical cognitive science. Symbolic representation and structure-sensitive processing provide the primary computational tools for the construction of mechanisms operating on the functional level. Classical cognitive science posits this specific, separate level of analysis in between a neuroscientific (physiological) and a mentalistic (semantic or intentional) level.

Connectionism pays more attention to the neuroscientific level by advocating a 'brain style' type of modeling. Importantly, however, connectionism does not try to do without the notions of representation and computation. Analyzing cognition in terms of the internal processing of representations is still thought to be of essential explanatory value.

Although connectionists like Paul Churchland speak of prototypical partitionings of multi-dimensional activation spaces instead of syntactic symbol manipulation, the idea of 'organized libraries of internal representations' remains (Churchland 1989, p. 207). Because of the great applicability of the flexible pattern recognition ability resulting from the use of distributed representations, claims have been made that a basic neurocomputational mechanism behind many cognitive capacities has been found (Churchland 1989; 1995).

DST rejects this connectionist claim. Instead of being *the* alternative, connectionism is portrayed as a first, interesting but insufficient, step in the right direction. The classical PDP-approach of Rumelhart and McClelland is, for instance, rejected as "little more than an ill-fated attempt to find a halfway house between the two worldviews." (Van Gelder & Port 1995, p. 34). Van Gelder and Port (1995, p. 3, 32) argue that standard connectionist models are still too much attached to the classical ideas of representation and computation and merely substitute activation patterns and weights for symbolic processing. Similarly, Thelen & Smith (1994, p. 42) criticize connectionism for still trying to explain the stability of cognition, thereby concentrating too much on the representational aspects of its models.

Dynamical Systems Theory

Dynamical systems theory studies the *self-organization* of dissipative, complex systems. The components of the system interact in a nonlinear way resulting in the emergence of behavioral patterns of a higher order. The emerging higher order behavior is in its turn capable of 'enslaving' the components, which results in patterns that can, in many cases, be described in terms of relatively few dimensions. Importantly, system components do *not* communicate via task related representations but interact directly and physically. Typically, the system is driven through different behavioral patterns under the influence of a control parameter. These patterns arise because the system may be said to be *attracted* to a specific configuration out of many possible states. Attractors are relatively (but only relatively) stable behavioral modes. A system's total of behavioral options is represented in its state space which can harbor many different attractors, each with its own basin of attraction. Under the influence of the control parameter, the landscape of the state space may change: attractors can disappear or originate, and their basins of attraction can widen or narrow down.

Thus, the basic idea of DST is that: "simple behavioral patterns and considerable pattern complexity may arise from the process of self-organization, as emergent consequences of nonlinear interaction among active components." (Kelso 1995, p. 66–67). In general dynamical insights can be applied most easily to the analysis of movement, like finger oscillations (Haken, Kelso & Bunz 1985), infant walking (Thelen & Smith 1994) and bimanual tapping (Peper, Beek & Van Wieringen 1995). However, there are increasing attempts to subject cognitive phenomena to a dynamical analysis.

The work of Edelman (1987; 1992) provides a detailed, though somewhat idiosyncratic (Kelso 1995, p. 228), analysis of how conceptual behavior can emerge in a complex dynamic neurophysiological system. In Edelman's view, interconnected neuronal groups constitute local maps that relate to specific aspects of perceptual input or motor output. Local maps can be correlated and coordinated with other local maps via reentrant connections resulting in thoroughly intertwined sensorimotor activity. A multitude of reentrantly connected local maps make up a dynamical system (called a 'global mapping' by Edelman) whose functioning results in perceptual categorization. Basically, Edelman considers perceptual categorization to be instantiated in the *correlation* between the activities of the sensory and motor neuronal components. This correlated activity over time can be displayed as a trajectory in a phase space the dimensions of which are formed by the involved sensory and motor components. After repeated interaction with objects and via Hebbian learning, such trajectories can become attractors. Conceptual categorization involves one further step: the experience- and value-laden categorization and recombination of parts of perceptual categorizations. This further step is based again on the correlation of activities of the involved neuronal structures. On this account, briefly outlined above, categorization is ultimately not based on a symbolic structure or a partitioned activation space but on a very changeable, context-specific attracting trajectory of time-locked activation patterns of neuronal groups.

From such a perspective, it is easy to understand why the notion of representation is regarded as "dispensable or even a hindrance" by van Gelder (in press, p. 28): "Dynamics forms a powerful framework within which one can develop descriptions of behavior, including especially interactive behavior with external domains, which sidestep the notion of representations altogether." A system is not considered to react to incoming stimuli by computationally processing represented information. Indeed, Edelman (1987, p. 266) even rejects the notion of 'information' as being too static. From this, I think it can be seen that DST offers a conceptual repertoire that goes beyond connectionism in its rejection of the notions of computation and representation.

Neurodynamics and Associationism

For a proper assessment of this recent trend in cognitive science, I suggest that it is fruitful to take a historical perspective on the basic role that *associative principles* play in Edelman's neurodynamic analysis of categorization. It is contiguity based association, the correlation in real time, between different neuronal groups in response to a stimulus that provides the foundation of the organism's categorization capacities (Edelman 1987, p. 234). In their discussion of Edelman's work, Thelen and Smith (1994, p. 150) unequivocally stress the importance of the 'perfect real temporal association' between the activities of the interconnected maps.

Within cognitive science, however, associationistic principles are generally regarded as suspect. Especially Fodor (Bever, Fodor & Garrett 1968; Fodor & Pylyshyn 1988; Fodor 1994a; see also Levelt 1989; Anderson & Bower 1973) prefers to speak of associationism in a pejorative sense as a label underscoring the naivety of any approach to cognition that is related to it. Fodor claims that it is already well-known that associationism lacks the means to explain cognitive processes because the basic mechanisms it proposes are simply too limited.

In the remainder of this paper I will argue against this verdict. Although it is undeniable that neurodynamics can be easily located within an associationistic tradition, this need not imply a weakening of its relevance for cognitive science at all. Associationism is a complex research tradition that has appeared in many different guises throughout history (Warren 1921; Jorna & Haselager 1994). Using the term as a simplistic negative label is therefore unjustified. Moreover, it is not clear how limited associative mechanisms within a neurodynamic context really are. It is precisely because of the recent developments that we may find that the true potential of associationism is substantially greater than previously thought.

Neurophysiology and associationism, I suggest, do make 'natural allies'. Neurophysiological speculations and investigations can be found all throughout the history of associationism. Already Hume proposed a neurophysiological theory of association to account for errors (it is worth pointing out that even Aristotle used physicalistic terms in his treatment of association; Warren 1921, pp. 26, 50–51). Hume suggested that animal spirits that 'rummage' the cell which belongs to an idea may sometimes deviate from the right track and activate other, wrong, cells and ideas as well (Hume 1738/1978, I, ii, 5, p. 61). At about the same time as Hume, David Hartley attempted to produce a more general physiological theory of association. Simultaneous sensations will produce corresponding vibrations of small particles in the medullary substance that, after repetition, will tend to 'lean' toward each other (Hartley 1749/1970, I, i, 4, p. 9, and I, i, 11, p. 17). He explicitly intended his doctrine of vibrations as a sketch 'for the service of future inquirers' (Hartley 1749/1970, I, i, 1, p. 6). Warren (1920, p. 71) finds this neurophysiological tendency of associationism characteristic enough to mention Thomas Brown as the great exception, differing "on this point not only with Hartley but with nearly all later associationists."

Edelman's work fits quite comfortably in the associationistic tradition of explaining behavior on the basis of correlative processes in the brain. Indeed, DST may have no other option. After all, the whole point of the classical idea of syntactically structured mental representations was to allow structure-sensitive mechanisms in addition to (or even in replacement of) associative principles which were deemed to be insufficient (Fodor & Pylyshyn 1988). Connectionism rejected the classical proposal but tried to stay clear of charges of associationism by insisting that connectionism can represent and use structured information in an altogether different, but still computational and representational, way. This has led to the ongoing debate about systematicity and functional compositionality (see a.o. Van Gelder 1990; Chalmers 1993, Haselager 1997; Haselager & Van Rappard 1998). If neurodynamics wants to go beyond connectionism in minimizing or rejecting the notions of computation and representation, it is hard to see what alternative it has left besides proposing neurophysiological association as a basic mechanism. To put it bluntly; what *can* neurons do but correlate their activity?

The associationistic nature of neurodynamics need not diminish its relevance for cognitive science. Although traditionally associationism is regarded as being too simplistic to explain higher cognitive processes, it is far from clear whether associationism is 'too simple by necessity'. All throughout the history of associationism, there have been proposals for a stronger associative mechanism than the one based on pure contiguity. For example, Hume, J.S. Mill and Bain invoked a separate 'similarity principle' in their explanations. Still, it remained entirely unclear on what mechanism this principle could be based or how it could be derived from contiguity. Of course, merely *postulating* a similarity principle is rather more like a symptom than a cure, as Fodor (1994b, p.85) has been keen to point out. And it has to be acknowledged that the associationists that have attempted to explain how a more powerful principle could arise out of the principle of contiguity (most notably Hartley but also James Mill) have failed to be convincing.

In my view, the interest displayed by DST-proponents regarding the potential of simple learning mechanisms is well in line with the aspirations of their associationistic ancestors. More importantly, they may have greater chances to succeed. A recent example of standard intuitions regarding the limitations of contiguity based association being overthrown is provided by Rebotier & Elman (1996; see also Elman, Bates, Johnson, Karmiloff-Smith, Parisi & Plunket 1996, pp. 333–340). They focus on Hebbian learning, which has great neurophysiological plausibility and is a plausible candidate as a means to neurophysiological self-organization. However, Hebbian learning is generally thought to have important limitations such as the inability to learn the 'exclusive or' (XOR) relation (Elman, Bates et al., p. 339). Rebotier and Elman address this problem by modulating the Hebbian learning process of a neural network by means of a 'dynamic wave'. That is, the Hebbian learning occurs in a wave passing through the network, so that units in one part of the net learn at an earlier stage than units in another part of the net, instead of the learning occurring everywhere in the network at once. By incorporating a spatio-temporal aspect in the learning process, the unexpected result was that units that 'learned late' *are* able to compute XOR.

So, merely pointing to the associationistic nature of neurodynamics does not provide a sufficient reason for rejecting it. Given the recent ideas about the nature of the dynamic processes that take place in the brain, it is hard to see exactly what the limitations of 'correlated activity' really come down to. Put differently, the neurodynamic investigations indicate that the standard 'too simple' verdict on associationism is not correct. From a historical perspective, then, it is both a promise and a continuing challenge of neurodynamics to show how much can be accomplished on the basis of purely associationistic means.

Conclusion

In comparison to connectionism, DST constitutes a further retreat from the classical cognitivist notions of computation and representation. Relating the recent neurodynamical view to the old associationistic tradition in psychology provides, I suggest, an interesting perspective on this new trend. Recent developments may yield an explication in mathematical and neurophysiological, instead of computational and representational, terms of processes that account for complex behavior. The neurodynamical view offers new possibilities for the revival of associative principles that cognitive science generally deems to be overly simplistic. In this sense neurodynamics can be seen as a deepening of classic associationistic viewpoints. In my view, there is no need to warn against this development, as one should openly and seriously consider the option that this may constitute progress for cognitive science.

References

Anderson, J. R. & Bower, G. E. (1973) *Human associative memory.* Washington: Winston & Sons.

Bever, T. G., Fodor, J. A. & Garrett, M. (1969) A formal limitation of associationism. In Dixon, T. R. & Horton, D.L. (eds.). *Verbal behavior and general behavior theory.* New Jersey: Prentice-Hall.

Brooks, R. A. (1991) Intelligence without representation. *Artificial Intelligence* 47: 139–159.

Chalmers, D. J. (1993) Connectionism and compositionality: why Fodor and Pylyshyn were wrong. *Philosophical Psychology* 6 (3): 305–319.

Churchland, P. M. (1989) *A neurocomputational perspective: the nature of mind and the structure of science*. Cambridge, Massachussetts: MIT Press.

Churchland, P. M. (1995) *The engine of reason, the seat of the soul: a philosophical journey into the brain*, Cambridge, Massachussetts: MIT Press.

Edelman, G. M. (1987) *Neural Darwinism: the theory of neuronal group selection.* New York: Basic Books.

Edelman, G. M. (1992) *Bright air, brilliant fire.* New York: Basic Books.

Elman, J., Bates, E., Johnson, M., Karmiloff-Smith, A., Parisi D., & Plunkett, K. (1996) *Rethinking innateness: a connectionist perspective on development*. Cambridge, Massachussetts: MIT Press.

Fodor, J. A. (1975) *The language of thought*. Cambridge, Massachussetts: Harvard University Press.

Fodor, J. A. (1994a) Fodor, Jerry A. In Guttenplan, S. (ed.), *A companion to the philosophy of mind*. Cambridge, Massachussetts: Basil Blackwell, pp. 292– 300.

Fodor, J. A. (1994b) *The elm and the expert.* Cambridge, Massachussetts: MIT Press.

Fodor, J. A., & Pylyshyn, Z. W. (1988) Connectionism and cognitive architecture. *Cognition* 28: 3–71.

Haken, H., Kelso, J. A. S., & Bunz, H. (1985) A theoretical model of phase transitions in human hand movements. *Biological Cybernetics* 51: 347–356.

Hartley, D. (1749/1970) Observations on man, his frame, his duty and his expectations. In: Brown, R. (ed.) *Between Hume and Mill: an anthology of British philosophy, 1749–1843.* New York: The Modern Library.

Haselager, W. F. G. (1997) *Cognitive science and folk psychology: the right frame of mind.* London: Sage Publications.

Haselager, W. F. G. & Van Rappard, J. F. H. (1998). Connectionism, systematicity and the frame problem. *Minds and Machines* 8 (2): 161–179.

Hilgard, E. R. & Bower, G. H. (1975) *Theories of learning*. Englewood Cliffs: Prentice-Hall.

Hume, D. (1738/1978) *A treatise of human nature.* 2nd Selby-Bigge edition. Oxford: Clarendon Press.

Jorna, R. J., & Haselager, W. F. G. (1994) Associationism: not the cliff over which to push connectionism. *The Journal of Intelligent Systems* 4 (3–4): 279–309.

Kelso, J. A. S. (1995) *Dynamic patterns: the self-organization of brain and behavior.* Cambridge, Massachussetts: MIT Press.

Levelt, W. J. M. (1989) De connectionistische mode: symbolische en subsymbolische modellen van menselijk gedrag. In Brown, C., Hagoort, P. & Meijering, T. (eds.), *Vensters op de geest: cognitie op het snijvlak van filosofie en psychologie*. Utrecht: Stichting Grafiet, pp. 202–218.

Newell, A. (1981) Physical symbol systems. In Norman, D. A. (ed.), *Perspectives on cognitive science*. Norwood: Ablex Publishing Corporation, pp. 37–85.

Peper, C. E., Beek, P. J. & Van Wieringen, P. C. W. (1995) Multifrequency coordination in bimanual tapping: asymmetrical coupling and signs of supercriticality. *Journal of Experimental Psychology: Human Perception and Performance* 21 (5): 1117–1138.

Rebotier, T. & Elman, J. (1996) Explorations with the dynamic wave model. In Touretzky, D., Mozer, M. & Haselmo, H. (eds.), *Advances in neural information processing systems.* Cambridge, Massachussetts: MIT Press, pp. 549–556.

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

Van Gelder, T. (1990) Compositionality: a connectionist variation on a classical theme. *Cognitive Science* 14: 355–384.

Van Gelder, T. (1991) What is the “D” in “PDP”? A survey of the concept of distribution. In Ramsey, W., Stich, S. P. & Rumelhart, D. E. (eds.), *Philosophy and connectionist theory.* Hillsdale: Lawrence Erlbaum Associates, pp. 33–59.

Van Gelder, T. (1992) Defining ‘distributed representation’. *Connection Science* 4 (3&4): 175–191.

Van Gelder, T. (1995) What might cognition be, if not computation? *Journal of Philosophy* 91: 345–381.

Van Gelder, T. & Port, R. F. (1995) It’s about time: an overview of the dynamical approach to cognition. In Port, R. F. & Van Gelder, T. (eds.). *Mind as motion: explorations in the dynamics of cognition.* Cambridge, Massachussetts: MIT Press.

Van Gelder, T. (in press) The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences.*

Warren, H. C. (1921) *History of the association psychology.* New York: Scribner’s Sons.