P. Van Loocke

NEURAL NETS AND THEIR INTEGRATION WITH PRINCIPLES FROM OTHER DOMAINS. THE CASE OF QUANTUM THEORY AND FRACTALS

1. Introduction

Nowadays, neural nets are considered as a method that can be used on its own in a number of contexts, but that can be used in a larger domain when integrated with principles from other domains. Usually, one uses principles from other ai-domains (genetic algorithms, fuzzy logic, etc.) in order to enlarge the scope of a network. We briefly explain in this paper how such an integration can be enlarged to include principles from quantum theory and from fractals as well.

2. Connectionist models and the variable binding problem

Since the appearance of the three volumes of Parallel Distributed Processing (Rumelhart and Mc Clelland, 1986), connectionist models are at the focus of cognitive research. Although the precise value of these models is still a matter of dispute, it is fairly generally recognized that they have been a step forward in our understanding of a wide range of cognitive and perceptual processes. At the basis of their success is their capacity to perform fast pattern recognition, to work with soft constraints, and to learn from experience. However, most connectionist models suffer from the inability to solve the variable binding problem. In criticisms on the connectionist paradigm, this is generally a central point. In order to illustrate the problem, consider the sentence 'Mary receives a red shirt from John', and suppose that we represent this sentence in a connectionist network. More specifically, suppose that we use a local network in which every predicate, every object and every verb are represented by a single unit.

The problem with this representation is that also the sentences 'John receives a red shirt from Mary', 'Red John receives Mary from a shirt, and so on, have an identical representation. Hence, this type of representation is unable to represent patterns in which an object and a predicate are bound together, or patterns in which it must be clear if an object is the first or the second argument of a verb. More general, this sort of connectionist representation is not suited to represent constituent structure: when a pattern of activation is given, it is not possible to differentiate within this pattern sub-patterns that are bound together.

One way to solve this problem is to introduce auxiliary binding units, and to split a verb or a relation in a number of units that equals its number of arguments. According to this representation, the bindings between the units can be detected if their connections with the binding units are examined (Shastri and Feldman, 1986).

A person who is communicating or who is interacting with his environment has to represent continuously new structures with bound variables. If he would do so with help of binding units and binding connections, he constantly would have to generate and to skip connections. However, it is rather unlikely that there exists in the brain a mechanism that can support such widespread changes at a timescale of hundreds of milliseconds. Hence, a binding unit approach can not be considered as a solution for the variable binding problem of short term cognitive processes; at most, it may play a role in long term memory. Therefore, a solution that solves the binding problem with binding units and binding connections is called a solution with 'static' bindings. If we want to describe how new information is manipulated and how it is integrated in long term memory, we must look for a solution with 'dynamic' bindings.

3. The SHRUTI model and dynamic variable binding.

Recently, Shastri and Ajjanagadde (1993) proposed a new connectionist model with the aim to solve the variable binding problem. Like its authors, we will refer to this model as the SHRUTI network. Consider the example of section 1 again: 'Mary receives a red shirt from John'. For this sentence, the variable binding problem can be solved if the same mark is attached to constituent elements that belong together. For instance, 'red', 'shirt' and 'receive:object' may receive mark '1', 'Mary' and 'receive: subject' may receive mark '2', and so on. According to this approach, a unit must be able to support different markers instead of a single activation value.

The idea to extend the number of variables that is associated with a single neural unit is in a sense fairly natural. In real neurons, one often not only detects a degree of activation, but also temporal variables like a phase and a frequency. In principle, a frequency can be used as a marker that determines which units are bound together in a distributed pattern of activation. In effect, this suggestion has been made, for instance, in the domain of perception (Eckborn, Reitboeck et al., 1990; Engel, Koenig et al., 1991). Alternatively, a marker may refer to the phase of a neural unit. According to this perspective, neurons have the same marker if they are firing in synchrony. This possibility is defended in Shastri and Ajjanagadde (1993). We will not discuss how precisely markers may be represented in the brain. For our purpose, it is sufficient to notice that connectionist models that include markers can solve the variable binding problem. If we take a more abstract perspective, we can notice that the idea of using markers in networks has a predecessor: also in the marker passing system NETL (Fahlman, 1979), different markers are circulating in the network. NETL is able perform cognitive operations like class intersection and some types of rule following. However, unlike SHRUTI, NETL needs a central controller that is constantly supervising and directing the process of marker passing. In this sense, it is unlike a connectionist model.

We have noticed how markers allow to represent information in which variables are bound together. SHRUTI also represents rules. Consider the rules 'if person 1 receives an object from person 2, then person 1 owns the object', and 'if a person owns an object, then he can sell it'. In figure 3, 'to own' is represented by units corresponding to 'own:subject' and 'own:object'; similarly, 'can sell' is represented by 'can sell:subject' and 'can sell:object'. Each of these rules is represented by two links. For instance, the first rule is represented by a link from 'receive:subject' to 'own:subject', and a link from 'receive:object' to 'own:object'. These links pass the marker of the source unit to the destination unit. For instance, if 'Mary' has marker '2', then also 'own:subject', and, subsequently, 'can sell:subject' receive marker '2'. This way, the system can quickly infer a large number of facts every time an input pattern is presented. The specific pattern of connections of the net guarantees that variables are bound together in an appropriate way.

Along with rules and incoming information, SHRUTI also represents long term memories of facts. A fact is represented by a single unit. In order to be able to address facts, relational terms must be provided with query- and answer-units. Fact-units, query-units and answer-units are active or inactive; they do not have to generate markers. The query unit and an the answer unit corresponding with 'to receive' allow to ask if there are facts in long term memory in which objects or persons are related by the verb 'to receive'. Suppose, for instance, that 'Mary receives a shirt from John' is encoded as a long term memory item, and suppose that the system is asked 'Did Mary receive a shirt from John?'. Since the question involves the verb 'to receive', the query-unit of 'to receive' is put active. At the same moment, all units that are involved in the question are put active, and they are provided with appropriate markers. The incoming connections to the fact unit 'Mary receives a shirt from John' are wired in such a way that the fact unit becomes active when (i) its constituent elements are active and have appropriate markers and (ii) the query unit of 'to receive' is active. The activity of the fact unit is propagated to the answer unit to which it sends information. When the latter becomes active, the question is confirmed.

SHRUTI allows for a type of backward reasoning in which a question is converted a number of times in alternative questions. As a more detailed examination of this system shows, the evidence for these questions may be searched along various paths. If evidence is found for an alternative question, then the answer to the latter question is propagated until an answer to the original question is obtained. In sum, this model has a number of attractive features. It exploits the concept of marker passing without being dependent on a central processor. As a consequence, it is a connectionist model that is able to bind variables and to perform different types of inference types. However, it suffers from four drawbacks:

(i) Since facts are stored as single units, they are not content addressable. This entails that a crucial advantage of connectionist models over computational ones is given up.

(ii) The pattern of activation that results when a pattern is presented as an external input differs profoundly from the pattern of activation that is obtained when facts in long term memory are examined. As a consequence, reasoning with these different types of patters occurs in a different way: in case of a long term memory pattern, reasoning must be in terms of queries and answers, whereas input patterns can generate forward reasoning. This asymmetric situation is unpractical and has no psychological basis.

(iii) The network can generate a number of cognitive processes, but it is not so clear what it adds to our understanding of them. Ideally, a connectionist model is more than a brain-style implementation of a process that can be described equally well with help of more classical tools. The point that SHRUTI is wiring out operations that can also be described by classical models becomes more sharp when its details are examined further. It appears that the resulting pattern of connectivity is highly specific, and, more importantly, that a learning rule for the connections is not provided (Touretzky and Fahlman, 1993).

(iv) In a thought process, a concept or a property can enter in more than one relation. For instance, in the utterance 'The black dog has black puppies', the node corresponding to 'black' must be allowed to have two markers. Alternatively, different nodes that correspond to 'black' must be present in the network. Shastri and Ajjanagadde (1993) choose for the latter option. Due to the variety of inferences that the system is claimed to master, this entails that every concept- or property-node should be duplicated a number of times.

In the next section, we propose a model that includes the conception of markers, but that has an attractor dynamics, so that it does not suffer from problem (i). The pattern of activation that results if a fact is presented as an input is identical with the pattern of activation that results if a fact is retrieved from long term memory, so that problem (ii) does not occur. Rules are represented in an implicit way, and their effect can be understood in terms of the attractor dynamics of the network (see problem iii). The learning rule of the schema model is suited to store facts as well as implicit rules in this system. Finally, a single unit may have different markers at the same time, so that also problem (iv) does not apply.

4. QNET: A quantum mechanical neural network

QNET takes the pattern of connectivity of the so called schema-model as its point of departure, but it allows that the state of a unit i is characterized by a set of different values A_i^f (f=1 to n) instead of by a single value. We call the value A_i^f the 'amplitude' of the f-th frequency in unit i. This terminology is borrowed from physics; there, it would refer to a temporal pattern in i that can be decomposed in n discrete groundmodes. The fact that, in every unit, only a discrete number of groundmodes contribute to the state of the unit is a first point that reminds us of quantum theory. A detailed examination of QNET reveals several other such facts (note 1). The amplitudes of the frequencies in every unit have a function comparable to the one of markers. Let us consider a simple example. Suppose that n=3, which means that three different frequencies are allowed to contribute to the state of a unit. Then, every unit is characterized by a vector with three components. Consider the example of section 1: 'Mary receives a red shirt from John'. Suppose that the unit 'shirt' is activated in the first frequency, i.e. its activation state is characterized by (1,0,0). Then, the units which are bound to this unit are also activated in the first frequency. Similarly, if the state of activation of 'Mary' is (0,1,0), then also 'receive:subject' must be activated in the second frequency, and so on.

We allow that amplitudes vary continuously and that, in the same unit, more than one amplitude differs from zero. If the latter fact occurs, this means that a concept or a descriptor participates in more than one set of bounded variables. In QNET, the chance that a unit interacts with another one is determined by the square of amplitudes, whereas the information that is exchanged is proportional to the amplitudes themselves. It has been shown that this entails that the network can process distributed information concurrently in different frequencies (compare this with a quantum computer that computes in synchronously in 'different worlds'). Moreover, these frequencies interfere in such a way that the activity in one frequency prohibits the occurrence of a spurious solution in other frequencies. Since the occurrence of spurious solutions is a major problem in constraint satisfaction neural networks, it can be understood that QNET has proven to offer a viable connectionist solution for a wide class of constraint satisfaction problems.

QNET is a new type of connectionist model. It reconciles two approaches. On the one hand, a number of authors recently argued that the type of coherence that is found at a quantum level may be vital in brain processes and in cognitive processes too (Penrose, 1989, Lockwood, 1989, Deutsch 1985). On the other hand, neural networks have already acquired a central place in present day cognitive science. QNET reconciles these inspirations by the way of a neural net that integrates quantum principles. As a result, it allows to represent more internal structure and internal coherence than the schema model does; the additional coherence may be used, for instance, to solve the variable binding problem.

The integration of quantum principles in QNET bears no mysterious concepts. QNET can be straightforwardly simulated and even a hardware implementation would probably not be dramatically much harder than a hardware implementation of more classical recurrent neural nets. QNET is not dependent on the production of SQUIDS, or on any other quantum mechanical hardware components, in contradistinction with other proposals for quantum computers. Most importantly, we have shown that QNET has a number of fairly relevant practical applications in AI and cognitive science. Along with its new solution for the variable binding problem, it allows to search concurrently for different global solutions in constraint satisfaction problems. If different solutions are equally attracting a given network state, then QNET will give these different solutions; a more classical neural net, on the other hand, frequently gets trapped in a spurious state. In such conditions, QNET generates different solutions with high internal harmony instead of a single solution with unsatisfactory internal harmony. Therefore, we suggest that integrating principles from quantum theory is a relevant enrichment of neural network methodology. In the next sections, we argue that a similar point can be made with respect to fractal theory.

5. Neural nets: passive and constructive classifications

Neural nets and their applications have been studied extensively in cognitive contexts. At the same time, properties of wholes of sub-networks have been examined. Such wholes have been compared with the human brain, and extensive simulations have compared capacities of artificial systems with human cognitive capacities (Van Loocke, 1990; 1991; 1994a; 1994b; 1995; 1996a; 1996d; 1996e; 1995f; 1996g; 1997a; Vandamme en Van Loocke, 1993). Among the artificial systems, connectionist ones appear to be especially suited to create classes with help of which differentiations can be introduced for the surrounding world (Ripley, 1996). Recent connectionist systems show that different types of codes are used to encode different types of stimuli. Also, different codes are suited as input for different types of further processing (Kosslyn, 1993; Van Loocke, 1996c).

Recently, the distinction between passive and constructive classification came to the foreground. In a passive classification process, the stream of information is dominantly one-directional, and a stimulus is transformed in representations that are successively more abstract. In a constructive classification process, the classifying system generates in parallel a large number of codes, and it examines which codes fits best a (pre-processed) stimulus. In case of the human brain, visual processing is largely constructive (Kosslyn en Sussman, 1995; Edelman, 1989; Van Loocke, 1996e; 1997a). In order to simulate this fact, networks have been proposed that are active in different frequencies and that can generate a different constructions in every frequency. These constructions evolve in accordance with a typically connectionist algorithm, but also under the influence of a genetic algorithm (Van Loocke, 1996e; 1997a). In special when a (multi)-fractal code of a landscape has to be determined, this approach appears to be particularly promising.

The problem of finding the fractal code for a landscape can be expressed as an optimalisation problem that is not expressible by a bi-linear function. Genetic algorithms (but not neural networks) can construct solutions for such problems. By embedding the algorithm in a neural network, the memory capacities of neural networks can nevertheless be used in combination with he genetic approach (Van Loocke, 1996a; 1997a). This approach to the coding problem works well for natural landscapes, but it appears to be rather clumsy when human-made environments have to processed. For such stimuli, classical generative 3D codes (in the sense of D. Marr; see Marr, 1982; Biederman, 1987) appear to be better suited. Another class of stimuli that may demand for still another type of code concerns non-verbal communication by facial expression; such expressions appear to be representable efficiently by means of wavelets. The robustness of the division of visual stimuli in three classes is currently examined; preliminary results indeed suggest that there are (at least) three classes of visual stimuli to be differentiated.

Fractal codes are very compact. Hence, a connectionist system that classifies on basis of such codes can work with a relatively low number of input units. It is currently investigated if backpropagation neural networks do classify landscapes in aesthetical classes in the same way as humans do. In fact, humans classify landscapes in a remarkably uniform way in aesthetic classes (Kaplan en Kaplan, 1989), so that the question if a compact fractal code lends itself for such a classification is meaningful.

6. Ecological and political implications

When landscapes that an be coded in a fractal way are systematically destroyed in favor of man-made environments, then human systems specialized in fractal processing receive less exercise. As a consequence, the skills to make aesthetical appreciations of such landscapes may be damaged. As a result, the opposition against the destruction of human nature becomes less strong. This way, a circular process results, during which both the external environment as well as the human cognitive skills are changed. Man risks to loose a skill that played a vital role since his appearance on earth. Moreover, since the fractal processing of stimuli is embedded in a complex web of other cognitive operations, the elimination of the corresponding type of stimuli may have consequences for other operations also. Current research examines if also man-made environments can be constructed in such a way that the fractal type of processing can be (partially) used. This type of work is of obvious relevance for landscape politics.

References

Biederman, K. (1987), Recognition by components: a theory of human image understanding, Psychological Review, 94, 115-147

Deutsch, D. (1985), Quantum theory, the Church-Turing principle and the universal quantum computer, Proceedings of the Royal Society, A400, 97-117

Eckhorn, R., Reitboeck, H., Arndt, M., Dicke, P. (1990), Feature linking via synchronization among distributed assemblies: simulations of results from cat visual cortex, Neural computation, 2, 293-307

Edelman, G. (1989), The remembered present,New York: Basic

Engel, A., Koenig, P., Kreiter, A., Gray, C. Singer, W. (1991), Temporal coding by coherent oscillations as a potential solution to the binding problem: physiological evidence, in: Schuster, H., Singer, W. (eds.), Nonlinear dynamics and neural networks, Weinheim

Fahlman, S.E. (1979), NETL: A system for representing real-world knowledge, MIT Press

Kaplan, S. , Kaplan, R. (1989), The experience of nature: a psychological perspective, Cambridge: Cambridge University Press

Kossyn, S., Chabris, C. (1992), Categorical versus coordinate spatial relations, Journal of experimental psychology: human perception and performance, 18, 562-577

Kosslyn, S., Sussman, A. (1995), Roles of imagery in peception, in M. Gazzaniga (ed.), The cognitive neurosciences,1035-1042

Lockwood, M. (1989), Mind, brain and the quantum, Blackwell

Marr, D. (1982), Vision, San Francisco: Freeman

Musgrave, K. (1994), Fractal landscapes in context, in K. Musgrave (ed.), ACM Interactions, 385-396

Penrose, R. (1989), The emperor's new mind, Oxford Press

Petrosky, T., Prigogine, I. (1995), Poincarй resonances and the extension of classical dynamics, Chaos, solitons and fractals, 7, 441-497

Ripley, B. (1996), Pattern recognition and neural networks, Cambridge: Cambridge University Press

Rumelhart, D., Smolensky, P., Mc Clelland, J., Hinton, G. (1986), Schemata and sequential thought processes in PDP models, in : Mc Clelland, J., Rumelhart, D., Parallel distributed processing: explorations in the microstructure of cognition, vol 2, MIT Press

Shastri, L., Feldman, J. (1986), Semantic nets, neural nets and routines, in: Sharkey, N. (ed.): Advance in cognitive science, Ellis Harwood/Wiley

Shastri, L., Ajjanagadde, V. (1993), From simple associations to systematic reasoning: a connectionist representation of rules, variables and dynamic bindings using temporal synchrony, Behavioral and Brain Sciences, 16, 417-494

Vandamme, F., Van Loocke, Ph. (eds.) New trends in neural networks, Ghent: Communication and Cognition

Van Loocke, Ph. (1990), Neural networks and higher mental processes: providing a Boltzmann network with meta-layers, Cognitive systems, 2-4, 1990, p. 383-409

Van Loocke, Ph. (1991), Study of a neural network with a meta-layer, Connection Science, 4, 1991, p. 367-379

Van Loocke, Ph. (1993), Neural nets: black boxes or interpretable models? Proceedings of the 18-th International Conference on Information Technologies and Programming, Sofia, p. 76-100

Van Loocke, Ph. (1994a), The dynamics of concepts. A connectionist model, Berlin/ Heidelberg/ New York: Springer Verlag

Van Loocke, Ph. (1994b), Properties of the conceptual database of the right hemisphere, Cognitive Systems, 4, 75-91

Van Loocke, Ph. (1994c), Connectionisme en natuurfilosofie. Een onderzoek naar enkele mogelijkheden tot concrete interactie, Gent: Communication and Cognition

Van Loocke, Ph. (1995a), QNET: A quantum mechanical neural network. A new connectionist architecture and its relevance for variable binding and constraint satisfaction problems, Cybernetica 38(1), 85-106

Van Loocke, Ph. (1995b), Nieuwe types van concepten in de Prigoginiaanse natuurfilosofie, in: Vandamme, F., Van Loocke, Ph.: Capita Selecta uit de Toegepaste Epistemologie, Volume IV, Gent: Communication and Cognition

Van Loocke, Ph. (1996a), The introduction of principles from elementary quantum theory in a Hopfield-style network and the solution of constraint satisfaction and variable binding problems, in: Dagli, C., Akay, M., Chen, P., Fernandez, B., Ghosh, J. (eds.), Smart engineering systems: neural networks, fuzzy logic and evolutionary programming, ASME Press, pp. 129-136

Van Loocke, Ph. (1996b), Een relatie tussen actuele epistemologische thema's en ecologische thema's, in Vromen, J. (ed.), Acta van de 18-de Vlaams-Nederlandse Filosofiedag, pp. 141-147

Van Loocke, Ph. (1996c), On the properties of a network that associates on basis of share of components and its relation with the conceptual memory of the right hemispere, in: Van Loocke, Ph. (ed.), Concepts in human cognitive activities, Communication of Cognition, 29(3-4)

Van Loocke, Ph. (1996d), Concepts: Representations and evolution, in Van Loocke, Ph. (ed.): Concepts: representation and evolution, Philosophica, 57

Van Loocke, Ph. (1996e), in van Loocke, Ph. (ed.): The connectionist model QNET and its combination with genetic algoritms, in Van Loocke, Ph., Concepts: representation and evolution, Philosophica, 57

Van Loocke, Ph (1996f) Complex aesthetic patterns in a recurrent neural network and their relevance for widespread communication, in: Van Loocke, Ph. (ed.), Concepts and Complex systems, CCAI: the Journal for the Integrated Study of Artificial Intelligence, Cognitive Science and Applied Epistemology, 13(3)

Van Loocke, Ph. (1996g) (ed.), Concepts and complex systems, CCai: the Journal for the Integrated Study of Artificial Intelligence, Cognitive Science and Applied Epistemology, 13(3)

Van Loocke, Ph. (1997a) (ed.), The nature and evolution of concepts, New York: Routledge (appears in winter 1997)

Van Loocke, Ph. (1997b), De metafysica van het bewustzijn. Een historisch perspectief, In Vandamme, F., Van Loocke, Ph. (eds.), Capita selecta uit de Toegepaste Epistemologie VI, Gent: Communication and Cognition

Van Loocke, Ph. (1997c), Kennisleer en milieufilosofie. Verschuivingen in cognitieve operaties in historisch perspectief, in Vandamme, F. Van Loocke, Ph. (eds.), Capita Selecta uit de Toegepast Epistemologie VI, Gent: Communication and Cognition.