BOOK REVIEWS

Natural and Artificial Intelligence

Misconceptions about Brains and Neural Networks

by Armand M. de Callataÿ

NORTH-HOLLAND, Amsterdam, 1992, 560p.

ISBN 0-444-895027

Angela Ionita was born in Romania, in 1955. She received M.Sc. in Mathematics -Computer Science from the University of Bucharest in 1978. Since then she has worked as programmer and senior researcher at the Research Institute for Informatics, Bucharest. She has been with the team designing and implementing software products on DBMS, DSS, Software Engineering. Since 1992 she has been head of GeMaSOFT Laboratory of which main objectives are basic and applied research, technology transfer and applied systems development in the field of non-industrial environment:GIS/LIS(distributed, homogeneous, multimedia) Databases for strategic planning of the societal informatization in Romania, DSS for agricultural management, marketing, public administration, environmental analysis. Currently, her executive activities cover R &D and Pilot Projects on Applied IT in public sphere, Multimedia Education, GIS/LIS. Her research activities focus on DSS for natural systems, neural networks, genetic algorithms, image processing. She is a Ph D candidate in Applied Informatics.

A first version of the book described an improbable brain model but did not show why any other paths were impracticable. These explanations might have helped researchers not to repeat errors of the past.

This is the second version of the book. The book topic is "large scale architecture" as defined in Minsky (1990).

Modelling brains is like designing a processing plant. An engineer must connect many devices in an efficient way. The brain designer has a large choice of components in the machinery of regulation systems, instrumentation, neural networks, and artificial intelligence. The design task is to reproduce natural intelligence in a robot by integrating all these devices.

Discovering how brains work is a detective job. A motto of this book could be: "When you have eliminated the impossible, whatever remains, although improbable, must be the truth" (Sherlock Holmes, Conan Doyle, 1884).

This is the proposed method here.

In the Prolog of this book, the author tries to dispute many current ideas, and its contents explains why the response to all the following questions would be "no".

- Should thought have an unifying principle?
- Should neural networks be homogeneous?
- Are all-or-none switches biologically impossible?
- Must memory disappear in biological organisms continuously modified?
- Is the plasticity of neural maps incompatible with computer models?
- Are the neural computations mostly based on inhibition and excitation?

- Is the topographical organization of brains necessary everywhere?
- Do we forget our previous habits and beliefs after having changed them?
- Are our behaviours frequently perturbed by noise?
- Is the large variance of our movements incompatible with digital systems?
- Does machine behaviour have less variance than that of animals?
- Does adaptive learning exclude irreversible memorization?
- Is symbolic computation incompatible with approximate reasoning?
- Is combinatorial explosion of cases unavoidable in symbolic systems?
- Is information distribution more reliable than redundancy?
- Do Lashley's experiments prove that brains memory must be distributed?
- Is addition of learned events impossible in conventional neural networks?
- Is instant learning rare?
- Are generalized rules necessary for reasoning?
- Is AI a rigid logic method?
- Is rate learning incompatible with understanding?
- Does intelligent classification need a teacher?
- Can classification be a continuous operation?
- Can a rhythmic system be non-oscillatory?
- Is the variance of reaction time incompatible with a rhythmic control?
- Does natural intelligence infer instead of deciding?
- Is decision-making a complex algorithm?
- Does decision depend on the way the quantum have collapses?
- Must we find mathematical formulae explaining thought?
- Is document retrieval a complex algorithm?
- Is reasoning by analogy a complex algorithm?
- Is intention a mental concept not implementable in hardware?
- Is consciousness not understandable as a mechanism?
- Do the limitations of expert systems prove that AI is an inadequate model?
- Do Goedel's and Tusing's limitations of mathematics and computers prevent natural intelligence by machines?
- Are the primary elements of brains knowledge permanent things?
- Must we first find how similarities are completed?
- Is long-term memorization preceded by short-term memorization?
- Can parsing be a continuous operation?
- Is human behaviour frequently efficient?
- Is human behaviour optimized?
- Is it impossible that simple mechanisms, if combined, produce intelligence?

There are contributions, like Marvin Minsky's, which keep close to Callatay's model. Callatay uses such concepts in his model as: semantic information processing in a network of knowledge, perceptions for adaptive learning and not for discriminations, frame processing, replacement of registers by activations in the memory cells and instantiation of activity lines, processing by transmissions along a pyramided network and learning by adding a very large store of facts.

The model's symbol manipulation originates from: the memory organization with pointers and lists as with Newell, Shaw and Simon; Newell's production rule system with its short term memory and its phased processing; Newell's physical symbol system.

Instead of such sophisticated methods as self-organizing networks, Callatay's model puts forward a combination of old fashioned and at hand methods:

- simple information retrieval methods in order to find the best document answering a query;
- classifiers producing no more than one response to a limited number of possible categories, thus reducing the total amount of information processed during a period;
- transformation of fuzzy information into definite symbols by classifiers;
- a grandmother neuron for each local decision or choice;
- relative independence of administrative, executive and evolutive functions;
- hierarchical control of operations with strict assignment of responsibilities;
- fast, all-or-one, permanent, irreversible record of facts, which are afterwards time stamped as episodes of a history;
- selection of a knowledge subset based on facts having occurred under similar circumstances;
- adaptive knowledge for urging the next intention;
- parsing of temporal sequences by the stack technique;
- substitution logic for inferences;
- goal-directed problem -solving to reduce combinatorial explosion in the search space;
- best choice in a breadth-first search and automatic backtracking;
- weighting of hypotheses as to processing uncertain or inconsistent information;
- proof methods of robot controllers with two stages: continuous and symbolic;
- sequential processing of sensory data,first in analog devices for sensory data reduction, and second, in symbolic devices for associative retrieval of meaning;
- conventional servo mechanisms for execution of symbolic processing decisions.

Next these methods will be reviewed. Most of them are as old as ten years, or even older ,and have been dropped by now for a variety of reasons, viz.:

- no single method can explain intelligence, as cybernetics or associative memories were supposed to do;
- no physical bases could be found for relevant, presumably mental processes such as logic of inference or intentionality;
- most of the methods were considered to be problem props, as though they had been using homunculi or social organization as metaphors rather than explanations;
- some of them have been given up because of the system under analysis not being implemented in the brain as hypothesized: e.g. grandmother neurons; computer gates.

Callatay's model is an extension of, or differs from, the published models in the following :

- Its main system is not **a model with storage modifications**, like Rosenblatt's (1959), Marr's (1969, 1971) and Albus's (1971) but an all-or-none model (Callatay, 1969);
- It uses learning **by conjunction of activity** as Hebb's (1949), Marr's (1969), Albus's (1971), Eccles's (1972), but it shifts the time of the previous state so as to associate it with the result of trial in a single rule (Callatay, 1969a, 1969b);
- It is an **associative model** like Willshaw's et al (1969), Kohonen's (1980), Kohonen's et al (1981), but with a **rhythmic processing and symbolic all-or-none** function (Callatay, 1969a). Its information retrieval system allows fuzzy information processing;
- It includes a **content addressable memory** as in Miller et al (1960), Lee and Paul (1963), Pribram (1971), but the RCAM is a symbolic computation unit and not an analog device;
- It matches input patterns like the **Pandemonium** of Selfridge (1959);

- It is a **connectionist model** like Cajal's, Feldman's (1981, 1982b), Fahlman's (1979, 1981) and Hillis's (1981, 1985), but processing gets manageable by a generalized use of inverted pointers, content addressable memories and rhythmic processing (Callatay, 1969);

- It is not a general-purpose computer as von Neumann machine is, although it can solve any problem (Kowalski, 1979). Depending on the size of the direct network processor system (DNPS), problems requiring large knowledge can at once be solved by insight. Puzzles or simple calculations on computer are almost strange to paper-and pencil-abiding persons ;

- It includes Post's **production rule system** (as described in Minsky, 1967), and that of Newell and Simon (1972) and Anderson (1982), but an inverted file organization allows memorizing and logic processing with no need for compilation (Anderson, 1982);

- It is a **mark-passing processor** like Quillian's (1968), Minsky's (1977), Fahlman's (1981), Hillis's(1981) and Sowa's (1982), but with a path control (Callatay, 1969a) or more precisely, with a DNPS organization (Callatay, 1982a);

- It is **neither analog nor symbolic**, but it **integrates** both types of antagonistic methods at many stages in its hybrid processing;

- It does not use the **reverberating circuits** of Lorente de No, but includes long loops. Reverberating circuits may be used for arbitration control (Callatay, 1969a), as in Kohonen et al (1981) and Hirai (1983);

- It does not **self-organize an unorganized neural network** but uses an organized network with potential logic capabilities;

- It stores **memory units or tables** like the robots of Ivakhnenko (1964) and Raibert (1978). However it does not use numbers and nor requires arithmetic units either;

- It has **two main types of learning** as in Estes (1982), Anderson and Bower (1973) and Wickelgren (1979), but they are **not defined as horizontal and vertical memory associations**. It uses permanent recording of simultaneous symbols and modifiable learning of the choice of the subsequent intention (Callatay, 1969a);

- It defines **automatic behaviour**, and **effort behaviour** as Pribram and McGuiness (1975) and Schiffrin and Schneider (1977) do. Both are based on pre-constructed sensory motor schemes. There are no recorded events in automatic behaviour. Effort behaviour is performed against possible distraction risks from more interesting events. Events unrelated to the present intentional sensory-motor scheme, are dismissed on decision-making, although they may be stored among contextual events recorded in the episode;

- **Short-term memory (STM) and long -term memory (LTM)** (Callatay, 1981a, 1981b) are defined as computerized functions. These definitions are compatible with those of Schneider and Shiffrin (1977). Callatay's model differs from the psychological models of James (1891), Atkinson and Shiffrin (1984), Newell and Simon (1972) and Normann and Bobrow (1976) in the following respects: **Nothing enters the STM which has not been stored yet in LTM**. In letter recognition paradigms, what is stored in the searched LTM is not a letter, but an association "letter - recent episode". The first keyword of the query is the letter (itself produced by the LTM which stores the letter's visual features). The second keyword is the intentional command enacting events related to the recent episode of the letters being learnt. Such STM and LTM functions were used, but not defined, in push-pull basal ganglia memory (Callatay, 1969b);

- The **iconic memory** has a feature detector or RCAM output. Because of the rhythmic processing, the classified symbol persists for one period;

- Long-term memory is **additive** as in the model of Raaijmakers and Shiffrin (1981);

- The **information retrieval algorithm for fuzzy sets** applied to LTM recognition is compatible with the SAM model of

Raaijmakers and Shiffrin (1981) and of Gillund and Shiffrin (1984), but adds a learning method based on formation of associations between symbols which are simultaneously present in the brain during the same period (STM and feature extractors). This storing method provides a computer-defined algorithm which distinguishes modalities of learning items (data in context) and learning sequences of items under the same category (sentence repetition):

- Long-term memory cannot be characterized by such words as **parallel or serial**. It is a complex device called RCAM. It includes an arbitration device which makes relaxation computations in parallel. This latter device hardly requires more processing time if more neurons are compared in order to find the maximum activation;
- The data element stored in LTM is a conjunction of relations which are represented by a symbol;
- STM is instantiated rather than memorized. The STM content is not memorized, as it is there in the LTM. What there is stored are the new symbols recording the event, its being related with instantiated symbols in other areas of STM;
- Relations may be built in LTM between the STM contents and the just instantiated symbols and the episode identifier. LTM links an episode with the STM contents;

- Callatay's model uses two methods for in-depth data processing:
  1) A processing loop: the number of iterations and, implicitly the reaction time are task-dependent. This process resembles the depth of processing described in Craik and Lockhart (1972);
  2) A processing through the various levels of the cortical hierarchy. Such processing may explain the increase of period duration with age. Mind is unable to limit its processing to the first levels, even though a higher processing is not required. This feature takes notice of infrequent threatening events and ensures common sense;

- **Remembering** is a problem-solving activity as in Neisser (1967), and Callatay's model explains it by the execution of Prolog like programs;
- Processing indexed by **the P300 evoked potential wave** recorded on the scalp (Donchin, 1979) is not a continuation of period processing, but a re-start of the same process using new data in STM. P300 is observed, for example, when unexpected events,which are temporarily stored in STM , must be combined to build a new event in LTM;
- Callatay's model uses Piaget's learning scheme, except for **assimilation** which is not modifiable, and **accommodation** which is only used for choice of intention. Behaviour transformations observed during the stages of child growth, do not imply memory modification, but a new upper set of intention commands to control lower stage schemes, just as metalevel rules can alter behaviour of Prolog clauses at a lower level (Kowalski, 1979; Bundy et al, 1980);
- Callatay's model has **no learning feature extractors**, as cybernetic models do. Learning happens in the symbolic classifiers;
- As in Milner (1962) and Arnold (1984), the **hippocampal circuits** are used to control **in which cortical area the new information must be stored in LTM**. This computation algorithm is now defined and has a neural mapping. It is based on the following rule: the data are stored in the same areas as the most similar events are;
- The **appraisal function** (Arnold, 1984) is located in the cingulate cortex (event storage). The association mechanism is a result of links existing between the event symbol and its experienced emotive reaction. In Callatay's model, the septal area has a major role in commanding the learning of the emotional responses;
- **Imagination**, the possibility of generating images not directly evoked by a present sensation, is controlled by the amygdala as in Arnold (1984). The imagination function (required for creativity or dominant

behaviour: Pribram, 1971) allows the brain to modify an object of thought by inserting additional instantiations which are part of a well-connected subnetwork. In Callatay's model, the basolateral amygdala and the claustrum can reload a stored situation, and thereby transform the present state. An old event is then mixed up with the current state, giving an imaginative interpretation of the present situation;

- The **"affective memory"** of Arnold (1960, 1984) consists of such connections as those from episodes and events to autonomic commands. These links provoke emotive reactions against the reconstruction of the event environment (Callatay, 1969b, 1971);
- Callatay's model includes the **motivation control** of Pavlov (1926), Mowrer (1960), Miller (1971), Gray (1975) and Olds (1977), but its predictive system, based on "token" neurons (neurons kept activated over many periods), better defines the system computation;
- System decisions depend on **reward and punishment variations** more than on their absolute value as in Gray (1975).

Conclusions

The brain model may contribute to four open directions of research:

1) **Finding whether the suggested methods of the brain model may be extended to complex behaviour control or not.** Large scale simulations are well-known methods. They should include connections to a real robot.

2) **Finding theories capable of proving self-learning from experience.** Callatay has not found explanations to the conventional theories saying how logic may be based on permanent functions instead of permanent objects, and why rules extracted from examples are correct. Another difficulty must be coped with: man reasons with negations; probably, animals have similar deductive capabilities, but developments are needed to find how negations can be processed in an environment of functions.

3) **Designing neurophysiological and psychological experiments that confirm or deny the brain model predictions on the neural hardware.** Callatay once thought that; the results of known experiments would permit to prove brain fast rhythms: this is not the case, because results have been averaged and primitive rhythms have partly disappeared in the process, if ever there. Searching for the infrequent event of a burst in an awaken brain requires new experimental paradigms,by means of which the exceptional and not the frequent is to be studied.

4) **Finding alternative models that, more likely, have neurophysiological grounds, but which have the same computational capabilities.** New computer architectures are being studied, due to new business opportunities. Computers, like connection machines, database tree machines, hypercubes, flow processors, vector processors, systolic machines, RP3, DNPS, are all of a different kind. Their study may show common principles of parallel processing. New robust algorithms applicable to the brain may be discovered as a by-product.

What are the applications of the brain model? The discovery of natural reasoning methods will help in programming new computers with easy-to-use and partly learned procedures: some methods providing a human-like interface are to be used in business applications. DNPS shows interesting algorithms of parallel processing. Callatay ends up with two research applications of the brain model: in neurology and psychology, both are speculative, and both present new views on traditional problems.

The book ranks high among IT professionals and deserves larger circulation among philosophers and medical staff, who will be rewarded the effort of understanding it.

The book is available from Elsevier. North Holland. Excerpta Medica (Postal Address: P.O.Box 1991, 1000 BZ Amsterdam, The Netherlands), or in the USA and Canada from Elsevier Science Publishing Co, Inc. (P.O. Box 882, Madison Square Station, New York, NY 10159, U.S.A.).

Angela Ionita