ESANN'2002 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-26 April 2002, d-side publi., ISBN 2-930307-02-1, pp. 71-82

Neural networks for modeling memory: Case studies.

Hélène Paugam-Moisy, Didier Puzenat,
Emanuelle Reynaud, Jean-Philippe Magué

Institut des Sciences Cognitives, UMR CNRS 5015,
67 boulevard Pinel, F-69675 Bron cedex, France

Abstract. First, neural networks have been inspired by cognitive processes [30, 16, 41]. Second, they were proved to be very efficient computing tools for engineering, financial and medical applications [15, 9, 18, 10]. In this article we point out that there is still a great interest, for both engineering and cognitive science, to explore more deeply the links between natural and artificial neural systems. On the one hand: how to define more complex learning rules adapted to heterogeneous neural networks and how to build modular multi-network systems for modeling cognitive processes. On the other hand: how to derive new interesting learning paradigms back, for artificial neural networks, and how to design more performant systems than classical basic connectionist models. After a short survey of connectionist models for modeling memory, we develop two case studies. The first is a model for a multimodal associative memory and the second is a model for more deeply understanding the mechanisms of spatial cognition[1].

1 Contribution of models to cognitive science

Artificial neural networks are born from the idea that the brain is the most efficient computational device for solving complex problems. Engineers inspired themselves from brain processing and used, with the success that we know, neural computation to build efficient algorithmical solutions for classification, regression, prediction, pattern recognition, etc. Nevertheless, artificial neural networks also have the power to help us investigating the mysteries of human brain mechanisms. Cognitive science has the aim to understanding how the brain works. This young science proposes theories, most of the time coming from experiments on human or animal. Building and running a connectionist model can help in many ways. First, translating a theory made of words, or of boxes and arrows, into a model that can be simulated by a machine forces the theory to be very specific, and to detail every point. This makes the theory

[1]Part of this work has been funded by the project COG-72 of the ACI "Cognitique".

ESANN'2002 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-26 April 2002, d-side publi., ISBN 2-930307-02-1, pp. 71-82

clearer and more complete. Second, if the model cannot tell us that things work in a certain way, it leads to eliminate impossible solutions, and helps to test the "computational feasability" of a theory. Third, it can highlights which part of the input information is the most crucial to realize a cognitive function. Fourth, a model allows to perform complex experiments, that cannot be realized in real world. For example, connections between neurons can be cut to simulate a precise brain lesion, hard to find in natural neural networks and impossible to make in reality. All the properties observed on the models can drive scientists to reconsider their theories, or to build new ones. As instance of model bringing new theorethical ideas, let us cite the classical work by Hinton and Shallice [19] on acquired dyslexia which is exemplary now. They showed that their lesioned network exhibited an error pattern resembling deep dyslexia, and that this pattern could come from a lesion anywhere in their system, which was an innovative idea, questioning the mainstream theory, which is now abandoned.

2 Connectionist models for modeling memory

Since memory can be considered as a central point in many brain processes, it is the object of numerous studies and models in cognitive science. The idea of understanding memory as an heterogeneous system is present since the nineteenth century [8]. Several decompositions have been proposed in the last forty years [3, 1, 11, 42, 45, 49, 43]. The computational approach of cognitive psychology aims to define sub-systems contributing to realize each processing step required by a system (e.g. visual system) for achieving a given task [29, 2], but most of the time, this "computational approach" is not confirmed by simulation of models on computers. Moreover, models usually handle symbolic representations and are very far from the neurophysiological level of cognitive processing. Connectionist models of the literature (e.g. Hopfield networks, Multilayer perceptrons, Kohonen maps) are too simple for modeling memory (even specific aspects of memory), but more complex learning rules can be developed or modular multi-network systems can be built for modeling cognitive processes. In [48], Tiberghien proposed a classification of connectionist models for modeling human memory:

- Connectionist models (*i.e.* neural networks), such as distributed auto-associative memories (recurrent networks) for simulating the other race effect in face recognition [36, 51] or multilayer perceptrons, with adapted architecture and learning rule, for face identification in context [44] ;
- "Neo-connectionist" models, based on convolution and correlation, for simulating the failure in recognizing informations, even if they can be recalled, without hidden units [14, 34] or with hidden units [20] ;
- Hybrid models, or "symbolico-connectionist" models, that integrate neural and symbolic processes, in different ways [47], taking into account the different levels of processing (fast and automatic recognition vs. slow and controlled information processing), such as using knowledge in text understanding [22].

ESANN'2002 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-26 April 2002, d-side publi., ISBN 2-930307-02-1, pp. 71-82

As a last example, Miikkulainen [31, 33] modelized the episodic memory, *i.e.* the memory of autobiographical episodes (like a first day at school or a wedding day). They adapted classical Kohonen map algorithm, and built a hierarchical model, from several feature maps representing memories to be encoded and stocked memories. The model exhibits a behavior facing "memory overload" similar to the human one : old memories fade away, but singular ones still remain. The model is an associative memory, and recall of the memory can be obtained from a partial sensory clue (eg: Proust's madeleine). The model also allows to make hypotheses on the order of magnitude of the number of real neurons and connections needed to store memories over a human lifetime in the hippocampal memory system.

After this short survey, next sections develop two case-studies. The first is a model of multimodal associative memory, binding sensory modalities, and the second is a model for more deeply understanding the mechanisms of spatial cognition.

3 Modeling a multimodal associative memory

3.1 Memory model, in cognitive psychology

In real life, both for human and animal, the perception of objects, in their environment, often involves more than one sensory channel. It has been proved that human ability for pattern recognition takes advantage of fusioning data perceived from different modalities, e.g. speech perception is enhanced by the vision of speaker's face [46].

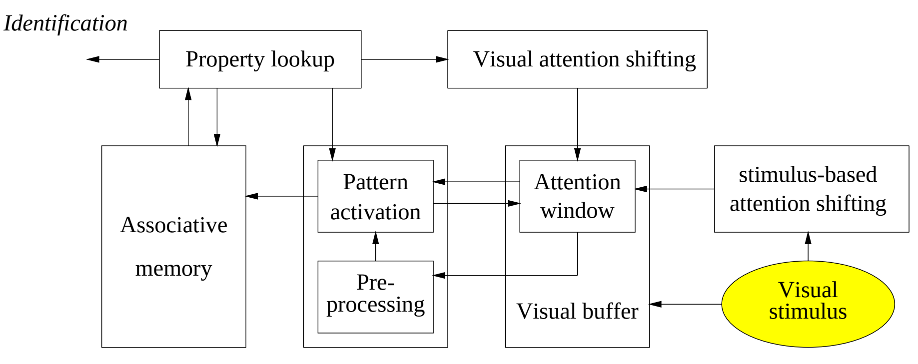

Figure 1: Simplified diagram of a functional architecture for visual perception

This section presents a model for data fusion, based on cognitive psychology. Starting from a functional architecture for high-level visual perception proposed by Kosslyn and Koenig [26] (figure 1), we assume, with psychologists, that similar architectures hold for other sensory modalities. In such a functional architecture, the treatment starts by a temporary storage of the stimulus in a *memory buffer* in which an *attentional window* selects a part of

ESANN'2002 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-26 April 2002, d-side publi., ISBN 2-930307-02-1, pp. 71-82

the input. Then, the stimulus is pre-processed and compared to representations previously met and stored in a *pattern activation subsystem*. In case of good matching between the stimulus and a stored pattern, the input is **recognized**. This low-level part of the system is specific to each sensory modality. Afterwards, a global *associative memory* receives all the outputs coming from the different sensory perceptive subsystems, and associates them to an abstract representation. If this representation is sufficiently close to a representation stored in memory, then the object is **identified**. This high-level part of the system is multimodal. In case of failure of the identification step, feedback loops allow top-down processes from associative memory back to modality-specific modules, and help to solve remaining ambiguities.

3.2 Connectionist modular network

A connectionist modular network has been built, from several basic bricks, for modeling this multimodal architecture of associative memory [13]. Several samples of an artificial "Incremental neural Classifier" [4, 38] simulate both pattern activation subsystem (InC1 and InC2, one modality-specific classifier for each modality) and identification (InC, an output classifier). The InC model has the property to create prototypes during its learning phase, according to two hyperparameters (a confidence threshold and a confusion threshold). In generalization phase, InC is able to evaluate the confidence of its response, and to give a *none* answer in case of doubt. The multi-modal *associative memory* is simulated by a "multiple Bidirectional Associative Memory (m-BAM)", a recurrent network derived from Kosko's Bidirectional Associative Memory [23], with several input vectors, the low layer being divided in sub-layers. The m-BAM learning rule has been adapted from Oh and Kothari PRLAB algorithm [35].

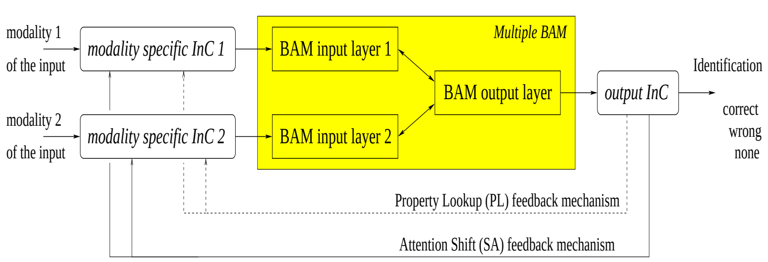

Figure 2: A two-modalities implementation of the system, with feedback loops

The whole model is presented in its two-modalities version on figure 2. The input object is taken as a set of modal inputs (*e.g.* its image and its noise, for a bi-modal object). Each modal input is sent to a modality-specific classifier to be recognized. Classifiers are trained to associate several representative "prototypes" to each category. Thus, each classifier outputs the prototype that best matches its input to the corresponding input sub-layer of the m-BAM;

ESANN'2002 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-26 April 2002, d-side publi., ISBN 2-930307-02-1, pp. 71-82

however, if a classifier doubts, a random pattern is sent to the sub-layer. When m-BAM inputs are fed, either with a prototype or a random pattern, the m-BAM reaches a stable state after a finite number of iterations, and an abstract representation of the object is obtained on the BAM output layer. Then, the output classifier tries to identify the object. If the output classifier doubts, the object is *un-identified*. Depending on the source of the problem, *e.g.* one or several modalities un-recognized by one or several modality-specific classifiers, a suitable feedback mechanism re-processes the object in order to simulate either the *property lookup loop* or the *attention shift loop* of the psychological model (cf. figure 1). The feedback algorithms have been implemented by *ad-hoc* algorithms, in the model.

3.3 Discussion

First, the computational feasability of the model has been successfully tested on objects with two modalities. Both modalities were character images, letters for one modality and numbers for the other one [39]. The relevance of the multi-modal architecture and of feedback mechanisms has been checked: the data fusion realized by the multimodal associative memory get a measurable performance gain for the model, feedback mechanisms significantly improve the success of identification, the model is able to simulate the phenomenon of mental image evocation [13]. Second, the model has been tested with three modalities and more realistic inputs (visual, auditory and tactile - via Braille's coding - perception, for the first five vowels in the french alphabet). Experiments proved the robustness of the model, even when one or two modalities are missing in input [37]. Third, the model has been placed in a virtual robotic environment. Thanks to the specificity of each modality, the virtual robot correctly identifies object, even at large distance. When a modality is missing (*e.g.* at night) or too noised (the object is too far) the system still shows good identification performance. Furthermore, it can identify the object while one of the modalities is incorrectly recognized, *e.g.* seeing an object while hearing another one. This example shows that a multisensory robot can take advantage of a two-step model based on cognitive theories [40]. Moreover, the model can be useful for engineers, as a model, based on local learning and data fusion, able to process multi-instance data, even in case of missing data (cf. [12]).

4 Model for understanding spatial cognition

Within the occipito-parietal pathway of the visual system, whose role is to infer the position of objects, two subsystems have been identified: The *categorical* encoding subsystem and the *coordinate* encoding subsystem. While the former places objects in space with categories such as "above/below" or "inside/outside", the latter uses precise measures to encode the position of objects [50]. This section describes how a mixture of experts has been used to model the interaction of the two subsystems, CAT and COO.

ESANN'2002 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-26 April 2002, d-side publi., ISBN 2-930307-02-1, pp. 71-82

4.1 Hemispheric specialization for CAT and COO tasks

Kosslyn [24], carrying out a computational analysis of high level vision, postulated the existence of two types of spatial relations, the categorical and the coordinate spatial relations, and thus two distinct subsystems to encode them. To test this hypothesis, different paradigms have been designed. One of them, proposed by Hellige and Michimata [17] consists in presenting to the subjects an horizontal bar and a dot, which can be at either one of three positions above the bar or one of three positions below the bar. With such stimuli, two tasks are possible. In the *categorical task* (CAT task), the subjects are asked to determine whether the dot is above or below the bar, thus requiring a treatment of the categorical subsystem. The *coordinate task* (COO task) consists in asking the subjects whether the dot is closer or further than a certain distance ($3mm$) from the bar, thus requiring a treatment of the coordinate subsystem. The stimuli are shown either in the right visual field (hence processed by the left cerebral hemisphere, LH) or in the left visual field (processed by the right visual field, RH). Analysis of the results showed that the CAT task is realized faster when the stimuli are presented in the right visual field (processed by LH) than when they are presented in the left visual field (processed by RH). Results obtained from the COO task showed the opposite pattern. Other experiments [27, 5] gave similar results, showing a predisposition for the left hemisphere to process the categorical spatial relations and a predisposition for the right hemisphere to process the coordinate spatial relations. It has been argued that this difference could be due to the use of neurons with different sizes of receptive field (RF) [25]. The RF of a neuron is defined by the area of the visual field in which the presence of a stimulus triggers a reaction of the neuron. Neurons with small non-overlapping receptive fields may be used to define the categories, while neurons with large overlapping receptive fields provide a coarse-coding that can precisely encode distances. This precise point has been confirmed by a connectionist simulation [28]. In multilayer perceptrons, we included the tuning of receptive fields in the parameters to be learned by back-propagation. The value of this parameter σ converges towards two different values, according to the task to be learned by the network: $\sigma = 0.763$ for the CAT task and $\sigma = 1.56$ for the COO task.

Moreover, many experiments (*e.g.* fMRI records) have shown that the predisposition of the right hemisphere for processing the coordinate task may disappear while the left hemisphere becomes better and better. It has been proposed that this practice effect is due to the development of new categories (such as "near/far" in the case of bar and the dot) by the left hemisphere [27, 7, 5], which becomes able to process the coordinate task in a categorical way. Next section proposes a model which helps to validate this explanation to the practice effect observed on human subjects.

ESANN'2002 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-26 April 2002, d-side publi., ISBN 2-930307-02-1, pp. 71-82

4.2 Mixture of experts for modeling the practice effect

4.2.1 Architecture of the model

For modeling the practice effect, the mixture of experts, proposed by Jacobs *et al.* [21], appeared to be the most suitable connectionist model. A mixture of experts is composed of several expert modules and a gating node. Expert modules are in competition, each of them receiving the same input and proposing an output. The gating node receives the same input also and gives as output one probability to each expert module. These probabilities are interpreted as measure of confidence in the outputs of the expert modules. For a given input, the output of the expert which receives the higher probability is chosen to be the response of the mixture of experts. The learning rule proposed by Jacobs *et al.* is adapted from backpropagation. At the end of the learning procedure, each expert becomes specialized in a part of the set of inputs. For modeling the competition between CAT and COO processing, we have implemented a mixture of experts, with two expert modules, the *CAT module* being designed to model the categorical subsystem, and *COO module* being designed to model the coordinate subsystem (see Figure 3).

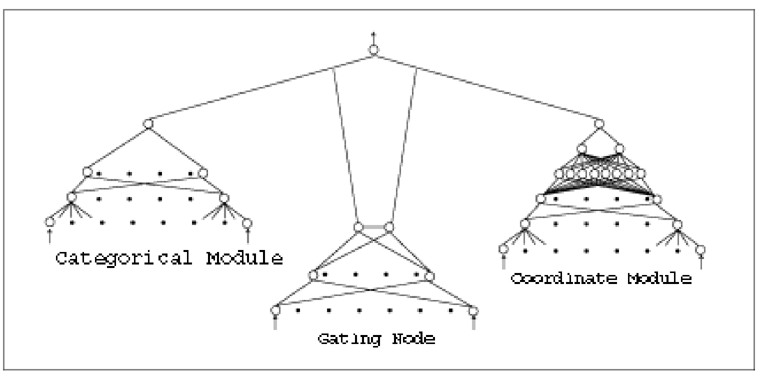

Figure 3: Architecture of the mixture of experts for modeling the practice effect.

The CAT module is a multilayer perceptron, with two hidden layers and small receptive fields. The weight between the i*th* neuron of the input layer and the j*th* neuron of the first hidden layer is determined by $w_{i,j} = \kappa \cdot e^{-\frac{(i-2\cdot j)^2}{\sigma^2}}$, centering each neuron of the hidden layer on one neuron of the input layer, with decreasing influence of its neighbors. The decrease rate is fixed by σ: The greater σ, the larger the receptive field. For the CAT module, σ has been fixed to 0.8. The COO module includes a similar architecture, with larger receptive fields, since σ has been fixed to 1.5. Moreover, in order to take into account the fact that people performing a coordinate task use the distance evaluation skill they have acquired all along their life, the coordinate module has been provided with a similar initial knowledge. Therefore, a third and a fourth hid-

ESANN'2002 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-26 April 2002, d-side publi., ISBN 2-930307-02-1, pp. 71-82

den layers have been added to the COO module. The upper subnetwork has been previously trained by backpropagation to measure the distance between the bar and the dot. The output layer of this subnetwork acts as a thermometer : the number of activated neurons indicates the distance between the bar and the dot. During the learning phase of the mixture of experts, weights of connections of this subnetwork remain fix.

The gating node is a multilayer network with one hidden leayer and two output units, giving the probability for each of the expert modules to be chosen as the output of the mixture of experts. The two outputs of the gating node must sum to 1. Therefore, each neuron calculates its activation (a_{cat} and a_{coo}) and communicates it to the other (there are lateral connections between the output neurons). Then each neuron calculates its output (p_{cat} and p_{coo}) from both activations [32] :

$$p_{cat} = \frac{e^{a_{cat}}}{a_{cat} + a_{coo}} \quad \text{and} \quad p_{coo} = \frac{e^{a_{coo}}}{a_{cat} + a_{coo}}$$

4.2.2 Experiments and results

Previous multilayer network simulations [25, 6] for studying the categorical and coordinate subsystems have based their learning set on the bar and dot paradigm of Hellige and Michimata. The bar is represented by 3 adjacent active neurons, and the dot by one active neuron. The dot can be place between 2 and 9 neurons from the bar, above or below (see Figure 4). To avoid an effect of correlation, the bar can take 29 different positions. Hence the input layer has 49 neurons. For each position of the bar the dot can take 16 different positions, resulting in a database of 464 input vectors. The COO task consists in deciding whether the dot is placed at more or less than 5 positions from the bar. Hence the expected output associated to an input vector is +1 if the dot is at more than 5 neurons from the bar, -1 otherwise.

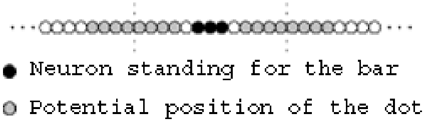

Figure 4: Input patterns for the mixture of experts model

The learning rule proposed by Jacobs *et al.* is not suitable for our model. The purpose of the model is no longer to partition the input space, specializing each expert via the gating node decision. The phenomenon to be explained is the practice effect of the COO task on the activities of the CAT and COO subsystems. We start from the hypothesis that the CAT subsystem develops new categories from examples and that the task is processed mainly by the COO subsystem until this new categories become sufficiently well defined. Under this assumption, what we expect to observe is a dynamic in the attribution of the probability by the gating network that reflects the development of new categories by the CAT module, transfering the preference towards the CAT module, to the detriment of the COO module, when the task is repeated. Therefore, the two experts modules are trained by a classical backpropagation, in order to make them learn the task (*i.e.* learn the critical distance, for the COO module, and develop new categories for the CAT module). The gating network is trained by backpropagation as well, but the error to be minimized depends on the progress of the development of new categories. The function to be minimized by the algorithm is E_{cat} for the CAT output and E_{coo} for the COO output:

$$E_{cat} = \frac{1}{2} \cdot \left(a_{cat} - \frac{1}{\epsilon + (S_d - S_{cat})^2} \right)^2 \qquad E_{coo} = \frac{1}{2} \cdot \left(a_{coo} - \frac{\kappa \cdot (S_d - S_{cat})^2}{\epsilon + (S_d - S_{coo})^2} \right)^2$$

where S_d is the expected output of the model, S_{coo} and S_{cat} are respectively the outputs of the coordinate and the categorical modules, and ϵ and κ are constants.

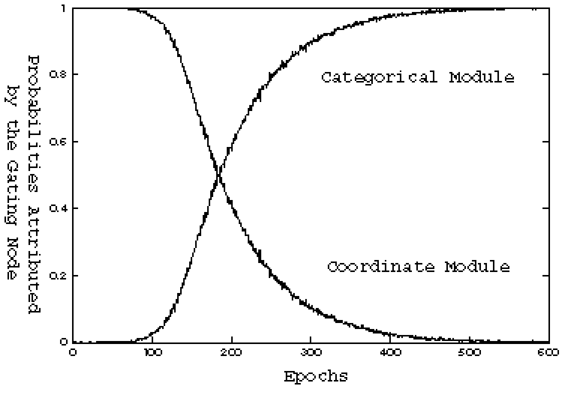

Figure 5: Evolution of probabilities attributed by the gating node to the CAT and COO modules

A mean of the tests performed on 20 networks, with a training procedure of 600 epochs, is shown in Figure 5. The model behavior is similar to the observations of psychologists. At the beginning of the procedure, the output of the COO module, due to its initial knowledge, is prefered by the gating node to give the answer of the model. But as long as the CAT module develops new categories, the gating node attribute larger and larger probabilities to its output, reflecting the practice effect observed in cognitive psychology experiments. This result argues in favor of validating the hypothesis that the reason why this practice effect is observed is that the CAT subsystem develops new categories. Although the model has proved the functional plausibility of this hypothesis, it would remain to develop more biologically plausible models to more deeply investigate the understanding of the cognitive process of spatial cognition.

ESANN'2002 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-26 April 2002, d-side publi., ISBN 2-930307-02-1, pp. 71-82

5 Conclusion

This article has proposed a short survey and two case studies illustrating how neural networks can be usefull for modeling cognitive processes and how inspiration coming from human brain can still improve connectionist models for engineering. It is clearly admitted that modular connectionist networks and hybrid models are more powerfull than classical neural networks, but their design is still complex and even hazardous. Starting from theoretical models proposed by psychologists and neurobiologists can help in many ways. However, the task remains hard, since the developer has to study the connectionist literature in order to find the best suitable neural networks as basic bricks, to adapt their learning rules to the situation or to the phenomenon to be modeled, to define and to implement cooperation algorithms between the modules. All these points have to be meticulously examined, otherwise the models are not functional and cannot be efficiently simulated on computers. In conclusion, involving neural networks in cognitive science is a challenging research area and can bring many advantages both to cognitive science and to connectionist models.

References

[1] J. R. Anderson. *The architecture of cognition.* Cambridge University Press, 1983.

[2] D. Andler. *Introduction aux sciences cognitives.* Collection Folio Essais. Gallimard, 1992.

[3] R.C. Atkinson and R.M. Shiffrin. Human memory: a proposed system and its control processes. In K.W. Spence and J.T. Spence, editors, *The psychology of learning and motivation*, volume 2, pages 89–195. New York : Academic Press, 1968.

[4] A. Azcarraga and A. Giacometti. A prototype-based incremental network model for classification tasks. In *Proceedings of Neuro-Nîmes*, pages 121–134, November 1991.

[5] M. Baciu, O. Koenig, M. P. Vernier, N. Bedoin, C. Rubin, and C. Segebath. Categorical and coordinate spatial relations: fMRI evidence for hemispheric specialization. *Neuroreport*, 10:1373–1378, 1999.

[6] D. P. Baker, C. F. Chabris, and S. M. Kosslyn. Encoding categorical and coordinate spatial relations without input-output correlations: new simulation models. *Cognitive Neuroscience*, 23:33–51, 1999.

[7] M. T. Banich and K. D. Federmeier. Categorical and metric spatial processes distinguished by demands and practice. *Journal of Cognitive Neuroscience*, 11:153–166, 1999.

[8] H. Bergson. *Matière et mémoire.* Paris : P.U.F. (4ème édition Quadrige, 1993), 1896.

[9] C.M. Bishop. *Neural networks for pattern recognition.* Oxford University Press, New York, 1995.

[10] F. Blayo and M. Verleysen. *Les réseaux de neurones artificiels.* Que sais-je ? Presses Universitaires de France, Paris, 1996.

[11] N. Cohen and L. Squire. Preserved learning and retention of pattern-analysing skill in amnesia: dissociation of knowing how and knowing that. *Science*, 210:207–210, 1980.

[12] A. Crépet, H. Paugam-Moisy, D. Puzenat, and E. Reynaud. Mémoire associative multimodale et mécanismes de retour de l'information pour la discrimination. In C. de la Higuera, editor, *Proceedings of CAP'2000.* Hermès, 2000.

[13] A. Crépet, H. Paugam-Moisy, E. Reynaud, and D. Puzenat. A modular neural model for binding several modalities. In H. R. Arabnia, editor, *The International Conference on Artificial Intelligence*, pages 921–928, 2000.

[14] J.M. Eich-Metcalfe. Levels of processing, encoding specificity, elaboration and charm. *Psychological Review*, 92:1–38, 1985.

[15] J. Freeman and D. Skapura. *Neural networks: Algorithms, Applications and Programming Techniques*. 1991.

[16] D. O. Hebb. *The Organization of Behaviour*. Wiley, 1949.

[17] J. B. Hellige and C. Michimata. Categorization versus distance: hemispheric differences for processing spatial information. *Memory and Cognition*, 17:770–776, 1989.

[18] J. Hérault and C. Jutten. *Réseaux neuronaux et traitement du signal*. Hermès, Paris, 1994.

[19] G.E. Hinton and T. Shallice. Lesioning an attractor network: investigations of acquired dyslexia. *Psychological Review*, 98(1):74–95, 1991.

[20] D.L. Hintzman. Recognition and recall in minerva 2: Analysis of the recognition failure paradigm. In P. Morris, editor, *Modelling cognition*, pages 215–229. Wiley, 1987.

[21] R. A. Jacobs, M. I. Jordan, S. J. Nowlan, and G. E. Hinton. Adaptative mixture of local experts. *Neural Computations*, 3:79–87, 1991.

[22] W. Kintsch. The use of knowledge in discourse processing: A construction-integration model. *Psychological Review*, 95:163–182, 1988.

[23] B. Kosko. Bidirectional associative memory. *IEEE Systems, Man and Cybernectics*, 18:42–60, 1988.

[24] S. M. Kosslyn. Seeing and imagining in the cerebral hemispheres: a computational approach. *Psychological Review*, 9:147–175, 1987.

[25] S. M. Kosslyn, C. F. Chabris, C. J. Marsolek, and O. Koenig. Categorical versus coordinate spatial relations: computational analyses and computer simulations. *Journal of Experimental Psychology: Human Perception and Performance*, 18:562–577, 1992.

[26] S. M. Kosslyn and O. Koenig. *Wet mind: the new cognitive neuroscience (2nd edition)*. The Free Press, 1995.

[27] S. M. Kosslyn, O. Koenig, A. Barrett, B. C. Cave, J. Tang, and J. D. E. Gabrieli. Evidence for two types of spatial representations: hemispheric specialization for categorical and coordinate relations. *Journal of Experimental Psychology: Human Perception and Performance*, 15:723–735, 1989.

[28] J. P. Magué. Modèles connexionnistes pour l'étude de la cognition spatiale : codages catégoriels et coordonnés, 2001. Mémoire de DEA de sciences cognitives.

[29] D. Marr. *Vision: A computational investigation into the human representation and processing of visual information*. W. H. Freeman, New York, 1982.

[30] W.S. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5:115–133, 1943.

[31] R. Miikkulainen. Trace feature map: a model of episodic associative memory. *Biological Cybernetics*, (66):273–282, 1992.

[32] P. Moerland. Some methods for training mixtures of experts. Technical report, IDIAP, 1997.

[33] M. Moll and R. Miikkulainen. Convergence-zone episodic memory: Analysis and simulations. *Neural Networks*, 10:1017–1036, 1997.

[34] B.B. Murdock. Todam2: A model for the storage and retreival of item and associative information. *Psychological Review*, 100:183–203, 1993.

[35] H. Oh and S.C. Kothari. Adaptation of the relaxation method for learning in bidirectional associative memory. *IEEE Transactions on Neural Networks*, 5(4):576–583, 1994.

ESANN'2002 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-26 April 2002, d-side publi., ISBN 2-930307-02-1, pp. 71-82

[36] A.J. O'Toole, K.A. Deffenbacher, H. Abdi, and J.C. Bartlett. Simulating the other race effect as a problem of perceptual learning. *Connection Science*, 3:163–178, 1991.

[37] H. Paugam-Moisy and E. Reynaud. Multi-network system for sensory integration. In *Proceedings of IJCNN 2001, Washington, July 2001*, pages 2343–2348, 2001.

[38] D. Puzenat. Priming an artificial neural classifier. In *Proc. IWANN'95, From Natural to Artificial Neural Computation*, 1995.

[39] E. Reynaud, A. Crépet, H. Paugam-Moisy, and D. Puzenat. A modular neural network model for a multimodal associative memory. In *Proc. of 4th International Conference on Cognitive and Neural systems, ICCNS'2000*. Boston University, US, 2000.

[40] E. Reynaud and D. Puzenat. A multisensory identification system for robotics. In *Proceedings of IJCNN 2001, Washington, July 2001*, pages 2924–2929, 2001.

[41] D. E. Rumelhart and J. L. McClelland. *Parallel Distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations.* The MIT press, Cambridge, 1986.

[42] D. Schacter. Implicit memory: history and current status. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 13:501–518, 1987.

[43] D. Schacter and E. Tulving. *Systèmes de mémoire chez l'homme et chez l'animal.* Marseille : Solal, 1996. (traduction de B. Deweer).

[44] A.-C. Schreiber, S. Rousset, and G. Tiberghien. Facenet: A connectionist model of face identification in context. *The European Journal of Cognitive Psychology*, 3:177–198, 1991.

[45] L. R. Squire. Memory, hippocampus and brain systems. In M. S. Gazzaniga, editor, *The cognitive neuroscience*, pages 825–837. MIT Press, 1992.

[46] W. H. Sumby and I. Pollack. Visual contributions to speech intelligibility in noise. *Journal of the Acoustical Society of America*, 26:212–215, 1954.

[47] R. Sun and L.A. Bookman. Integrating neural and symbolic processes: The cognitive dimension. In *Procceedings of AAAI'92*, San Jose CA, USA, 1992.

[48] G. Tiberghien. *La mémoire oubliée.* Liège : Mardaga, 1997.

[49] E. Tulving. Organisation of memory : Quo vadis ? In M.S. Gazzaniga, editor, *The cognitive neuroscience*, pages 839–847. Cambridge, Massachussets : MIT Press, 1995.

[50] L. G. Ungerleider and M. Mishkin. *Two cortical Visual Systems.* MIT Press, Cambridge, 1982.

[51] D. Valentin, H. Abdi, A.J. O'Toole, and G.W. Cottrell. Connectionist models of face processing: A survey. *Pattern Recognition*, 27:1209–1230, 1994.