

Why Topological Maps Are Useful for Learning in an Autonomous Agent

Stephane Zrehen*, Philippe Gaussier**,

*Swiss Federal Institute of Technology

Microcomputing laboratory

EPFL, CH-1015 Lausanne

** ENSEA-ETIS, 6 Av. du Poinceau,

F-95014 Cergy-Pontoise Cedex

Email: zrehen@di.epfl.ch, gaussier@ensea.fr

Abstract

In this paper, we discuss the usefulness of topology preservation in an on-line learning neural control system. We discuss biological and information processing arguments. Then, we propose an experiment performed on a mobile robot that shows that with a Probabilistic Topological Map (PTM) much less information needs to be learned than using a Winner Take All.

Keywords

Neural Networks, Topology Preservation, Internal Representation, On-Line Learning

1. Introduction

Self-organized feature maps [KOHONEN82] have been widely publicized both for their modelization of feature extractors commonly found in mammals nervous systems and for their ability to allow a relevant dimension reduction of the input space that preserves certain topology relations. Most research on applications of Kohonen maps has focused on establishing categories at the level of the map itself, [KOHONEN88],[RITTER89], defining them as connex zones in the array. This vision thus requires a supervisor at some stage of the processing to define the categories' frontiers.

The enactivist paradigm [MATURANA87], [BOURGINE91], [STEWART91] is an alternative to the classical "cognitivist" paradigm [LAKOFF87] which sees cognition as a manipulation of formal symbols. According to the enactivist school of thought, cognition is an emergent phenomenon, proper to autonomous systems, defined as a *coupling of perceptions and actions via the external world that allows survival*. In this view, supervised learning should be excluded from the design of learning procedures for an autonomous system. The only abilities that can be given to the system are general and non-prescriptive, i.e., they are not the result of the computation of a pre-given function. Therefore, meaning cannot be given per se to any neural cell or group of cells. It has to emerge as the consequence of an association between a sensorial state and a motor action.

However, direct coupling of actions and perceptions leads immediately to the dead-ends well identified in behaviorist psychology. Indeed, such an architecture supposes implicitly that input data are non-ambiguous, and that close data should always lead to close actions, except over the boundaries defined in advance in the input set. This is done for instance in a Braitenberg type architecture for obstacle avoidance [BRAITENBERG84]. On the other hand, by taking into consideration the output neurons, they must be activated by similar input, if no "hidden" layer is present. This is due to the fact that neurons should not have an activation function with several maxima.

Gestalt psychology has shown that data can be essentially ambiguous. For instance, fig. 1 can be interpreted in two different ways, either as a vase or as two human profiles. However, human beings can "choose" one interpretation. The ambiguity can only be lifted by acting on the environment in conformity with a memorized sequence of perceptions. This suggests that an internal representation between the input and the output of the system is necessary in those cases, even though this does not imply that this memory is equivalent to that of a computer. However, if the data can be categorized in a single way, such an intermediary step is useless [VERSCHURE92].

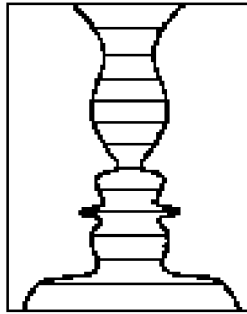


Figure 1: This image can be seen as representing two faces or a vase

Our research is in general concerned with on-line learning of an interesting behavior by an autonomous robot. In our networks, there is no separation between a learning and a utilization phase: learning continues all the time. Therefore, in order to learn boundaries of perceptual categories or to be able to lift ambiguities, an internal representation of perceptual situation is needed. In classical Neural Network (N.N) terminology, that is to say that we need hidden units between the input and the output of the network. Otherwise, it would be impossible to put into question the categories established thus far without performing learning all over again.

In this paper, we focus mainly on the on-line construction of an internal representation of the sensorial situations that preserves topology of its input data. We will discuss why this topology preservation is useful in an on-line learning context, as compared to a traditional Winner Take All (WTA) group. Next we will remind basic principles of the Probabilistic Topological Map (PTM) which is a fast-learning algorithm with topology preservation properties [GAUSSIÉ93]. In the end, we propose a "conditioning" experiment performed on a mobile robot, aimed at illustrating the performances of the PTM.

2. Topological Maps

Accepting the necessity of internal representations does not tell how they should be made. Nevertheless, biological evidence suggests that a) topographic organization is found in most parts of mammalian nervous systems, b) features are extracted on the earlier stages of treatment, c) precise analogical data is not available after those early processing areas [BURNOD89].

In the following, we will first expose arguments for using a topological map and then present a well-known diffusion mechanism for processing both the input and output data of the PTM. The basic qualitative aspects of functioning of the PTM will be described. Then, we will propose results of an experiment implemented on a Khepera robot [MONDADA93], that illustrates the performances of the PTM in function of the importance given to topology information.

In this section, we will review both the evidence for topological organization in biological nervous systems, their usefulness for building internal representations, and artificial models.

Topographic organization is very common in all of the peripheral nervous system of mammals: two close stimuli activate adjacent neurons. Empirical evidence suggests that the same kind of organization can be found in the cortex where different maps have been isolated in several parts of the brain, among which the auditory cortex, the visual cortex and the motor cortex (Brodmann areas). This arrangement of cells and connections is the result of the ontogenesis of animals [PENFIELD50].

Some maps are anatomically connected either to input (sensory maps) or to output (motor maps) of the animal [GEORGOPOULOS88], [SHAMBES78], [GHEZ85]. Moreover, a number of associative maps have also been found in mammalian brains, connected to both a sensory and a motor map [MOUNTCASTLE75].

Ordered maps thus seem to be a fundamental element of the nervous system. However, reconstructed maps show that there is not always continuity in the topology preservation, i.e. the organization is only local. Input patterns tend to be represented in clusters, and certain short shifts in the map correspond to large differences in the input [EDELMAN84]. Therefore, it does not seem necessary to obtain a global organization as with Kohonen maps to obtain features observed in natural brains.

2.1. The interest of a topological map

Durbin and Mitchison [DURBIN90] showed that topological self-organizing maps can be helpful in two respects: they allow a relevant dimension reduction of the input space and they also minimize the needed wiring required for local operations.

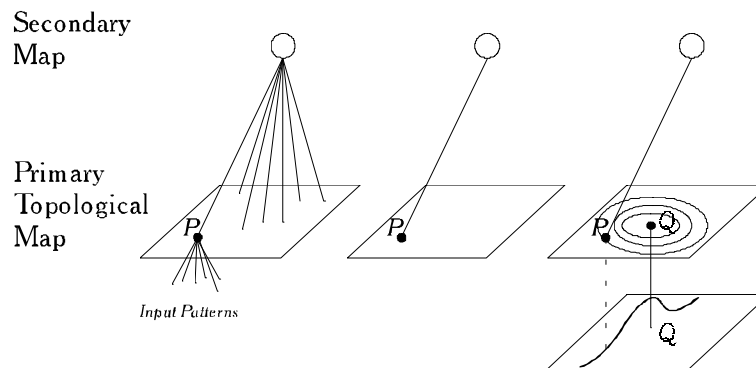


Figure 2: Usage of a topological map. a) A pattern is coded on the map on neuron P. b) A learning mechanism modifies the link between P and the connected neuron of a further group. c) A modification of the learned pattern is presented to the map. I is the new winner, and P responds in function of the distance between P and I, according to the diffusion law depicted below.

Dimension reduction is essential, as the cerebral cortex, where most cognitive activity seems to take place, is a two-dimensional array. Most cortical ordered maps represent two-dimensional spaces, such as the skin or the retina projection. But they get access to the information from the sensors through a space that has as many dimensions as sensors [OBERMAYER90]. Statistical methods face hard problems to find in such cases the non-planar 2-D subspace where the data is concentrated. That was one of the most promising aspects of Kohonen maps: they seem to be able to organize themselves onto the subspace defined by the input distribution without other knowledge about the input.

On the other hand, computer simulations have no particular problems with localization precision. But memory resources are always limited, and it is therefore fruitful to implement neural networks that minimize the number of links to memorize or to compute.

For most recognition tasks, it should in principle be possible to use a Winner-Takes-All (WTA) group of neurons [RUMELHART86]. But in the absence of topology preservation, all possible associations should be learned. Moreover, it would be impossible to use information coming from the neurons that have not yet learned something. Whereas, with topology preservation, all neighbors of previous winner cells will respond in function of their distance to the winner (fig. 2). This response thus provides a distance measure between the activity generated by the exact input corresponding to something already learned and that generated by a distortion of this input. Generalization capacities are enhanced by such features of coding.

Let us now imagine that we want to train a network to respond differently to different intervals of a given signal, for instance a sound frequency. If one intends to do that with a WTA, then two choices are possible: either the neurons are very selective, or they are not. If they are not, then the first winner is likely to win for any other frequency. Therefore, they should be made rather selective, once they have learned.

Let us suppose that we first present a 1000Hz sound, and the neurons have a 400Hz span (fig. 3a). Then 1000Hz is learned by a neuron N, and any sound in the 800Hz-1200Hz interval provokes a good response in N. Now let us present a 1600Hz sound. As this value is higher than N's upper bound, a new neuron, say M, must learn this new sound. Then it is impossible to generalize for a sound around 1290Hz: if the desired response for this sound is the one associated to N, then a new link must be learned. Introducing overlapping between neurons would allow the generalization, but the it could not be put into question.

With a topological map, the generalization from the 1000Hz can be extended to any value inside the bubble that can be chosen as wide as desired. Thus, if 1000Hz and 1600Hz are presented consecutively, the whole interval 0-∞Hz is covered, and separated in two categories (fig. 3b). Now, if a sound around 1290Hz is presented, its winner is K, closer to N than to M (fig. 3c). If 1290Hz is to be associated to the same category as A, then nothing must be done: N's activity is higher than M's as it is equal to the diffusion of K on N. On the other hand, if 1290Hz must be associated to another category, then the repartition among neurons N, M and K is transformed.

Actually, with the PTM, the functioning is even richer: if two patterns are successively and they have the worst possible matching, then the second one is coded at the border of the first winner's bubble [GAUSSIER94]. Thus, in a mono-dimensional case, all frequencies will yield a response in the first winner, even very low.

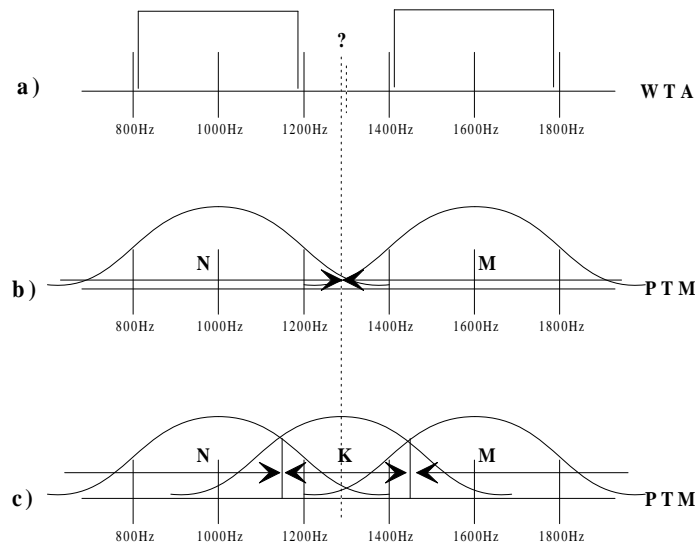


Figure 3: Category formation with a WTA and a PTM.

One final argument in favor of topological maps is that they provide a systematic tool to transmit information from one level to a level further in an information processing chain. Indeed, a diffusion mechanism can also be applied to the input itself, and close inputs then give rise to similar activity patterns. This will be explained in detail in the paragraph devoted to the PTM.

2.2. The Probabilistic Topological Map

Most neural networks, such as Kohonen maps or multi-layer perceptrons rely on the presentation of a large number of data, and on a separation between utilization and learning. This makes them unsuitable candidates for on-line learning of a limited number of processed input vectors in an autonomous system. Indeed, the most active neuron for a given input may change with time during learning. Moreover, for an autonomous system, the decision of stopping when to learn cannot be made from the outside. Therefore, an ART1 algorithm may be suitable, but it provides no means of preserving the topology. However, the plasticity/stability dilemma put forward by Carpenter & Grossberg [CARPENTER85] remains present for that type of internal representation of met perceptive situations.

We have proposed an algorithm for a topology preserving neural map that fits in a counter-propagation type neural architecture [HECHT-NIELSEN88]. The Probabilistic Topological Map (PTM) algorithm is based on the mechanism of cooperation-competition, where the weights are binary, and their modifications obey a probabilistic law.

The Probabilistic Topological Map (PTM) is an attempt to bring together features of a fast-learning algorithm and of a topology preserving map. It has been presented in detail elsewhere, therefore we shall only recall its main functioning principles:

1. Weights are binary. Therefore, the PTM should be utilized with already pre-processed input.
2. When a winner is elected (competition), its activity is diffused on its neighbors inside a bubble of a given size (cooperation). Then the weights of all the neurons inside the bubble, including the winner, are transformed. The magnitude of the probabilistic weight transformation decreases with the distance to the winner. Thus the winner's is transformed in the input and its neighbors' weights are probabilistically brought closer to the input. This weight transformation rule ensures that for large dimensional input spaces, an activity bubble is formed under presentation of the same input at a further stage.
3. After a weight transformation, the selectivity of all modified neurons grows higher, thus making them eligible as winners only for inputs very similar to the one they have learned. The selectivity implemented by a Gaussian activation function, decreases with the distance to the winner.
4. A neuron's weight vector can be modified only if the neuron is closer to the actual winner than ever.
5. Learning is performed if the selectivity of the proposed winner is lower than a vigilance parameter. This helps to avoid a saturation of the map, if the regulation of the vigilance is properly done.
6. Once a neuron has coded a pattern, the winner for a similar pattern is in the neighborhood of the winner, at a distance between winners that decreases with the matching between their two associated patterns. This feature of the PTM, which it shares with Kohonen maps, is the most important, because it somehow allows to implement prototype theory from psychology. Indeed, a category can be defined as a subset of the input space which is associated to the same response. By using the PTM, one assumes that the first input of a category is a good prototype,

a priori extending its properties to its neighbors in the activity bubble. Should one neuron in a bubble be associated to another category, a learning mechanism linking the map to a further group should allow to learn this frontier. The simple choice we made is presented in the next paragraph.

The diffusion mechanism can also be applied to the input: if the input pattern is viewed as the result of a competitive mechanism, as a contour extraction or more simply the quantization of a continuous value, then the maximum can also be diffused on its neighbors [SEIBERT89]. This also helps to preserve topology, as identical patterns displaced by one pixel should be seen as more similar than if they are shifted by two pixels. The general situation is illustrated on fig. 4.

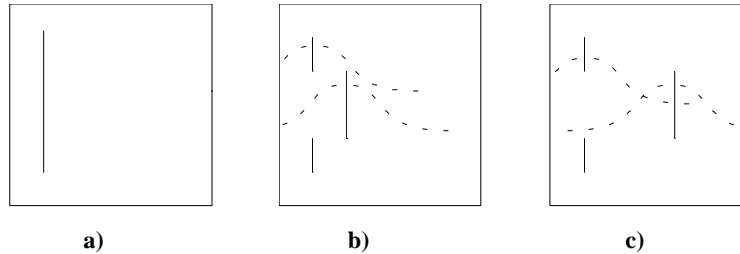


Figure 4: Deformation of a bar pattern. The matching between a) and b) should be higher than between a) and c). This is obtained by using a diffusion in one direction on the active pixels of the input image.

The PTM provides a local topological organization, as the map is constructed step by step, with a memory that is fixed, i.e., the neurons which have coded something cannot learn something else instead afterwards. Therefore, it cannot provide a global organization over the whole data set, contrary to Kohonen maps. If the whole data set is visited, then very different patterns are also likely to be coded on close bubbles, because that is where there is room left. This kind of organization has been reported about rats' hippocampus by J. O'Keefe: "These neurons bear no necessary spatial relation to one another; that is they are as likely to represent distant patches of an environment as to represent close ones" [O'KEEFE89].

3. A simple experiment on a mobile robot.

In order to illustrate the behavior and usefulness of the PTM, we implemented in a neural architecture playing the role of a "brain" for a Khepera™ mobile robot. The experiment consists in having the robot learn not to collide in walls and not gyrate endlessly. We decide that either collisions or turning when there is nothing on the front sensors cause pain. The goal is to learn this task with the minimal number of time steps and of stored information. Obviously, using a PTM for such a trivial task may seem unnecessary, but we should stress that we are not interested in the task itself. We took this experiment first because it is real and second because it shows that if one decides to use an internal representation of sensory states, then one had better use topology. This is a general statement, not dependent on this task whatsoever.

A 20x20 cell PTM is used as a support for internal representation of the sensorial situations met by the robot. Actually, this task is very much similar to any classification with the help of a self-organizing map, such as Kohonen's phoneme map [KOHONEN88], except that here we insist on learning both the map and the classification on line. On such previous applications, there is always a separation between learning of the map, and of further associations: the classification phase, usually

performed with a perceptron-type rule, is done after the map is completely organized. Our approach requires a regulation of learning, i.e., a mechanism to decide when to learn. The architecture we propose is a very simple realization of that.

The eight sensors provide the input to the PTM, and a diffusion mechanism is applied to them. Three basic actions are used: turn 35° left, turn 35° right or move forward by about 1.5cm. Such small movements are required because the range of the infrared sensors is about 3 cms [MONDADA93].

The global network is depicted on fig. 5.

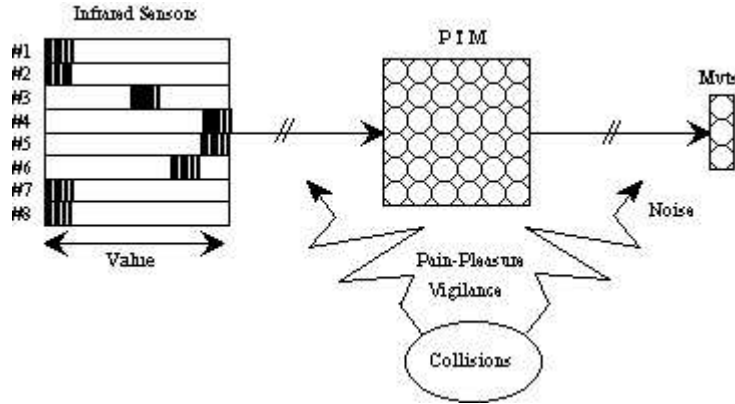


Figure 5: The control neural network for learning obstacle avoidance. Input is coded as an array of eight vectors, each corresponding to an infrared sensor. For each sensor, the value is quantized and diffused. Sensorial situations are coded on a PTM. A WTA group codes basic movements. A pain-pleasure signal (a function of the collisions) decides when to learn on the PTM and between the PTM and the movements (mvts) WTA.

At each time step, the sensors values are collected, and proposed to the PTM which decides to learn this pattern or not, in function of the value of the vigilance. In any case, a winner is found on the PTM and an activity bubble is formed by the diffusion of the winner's activity. Then the activity of neurons in the Mvts group is computed and a winner is chosen, which results in performing the corresponding movement (L, R or S).

The activity of these latter neurons is computed as follows:

$$Act(N_j) = Max(W_{ij}D_i) + noise$$

where N_j is the j^{th} neuron in the mvts group, W_{ij} the weight of connection to cell i in the PTM, and D_i the value of the winner's issued diffusion (see fig. 6).

After the step-movement is performed, W weights are adapted according to the following rule, corresponding to an extremely simple form of reinforcement learning [BARTO83]:

$$\begin{array}{ll} \text{Pain} = 1 & \text{Pain} = -1 \\ W'_{i*j} = W_{i*j} - \epsilon & W'_{i*j} = W_{i*j} + \epsilon' \end{array}$$

where ϵ and ϵ' are "learning rates" with $\epsilon' > \epsilon$, and weights are kept in $[0, 1/2]$ by a hard threshold function mechanism. When $\text{pain} > -1$, we also adapt the vigilance to a minimal value that allows the coding of patterns very different from the ones already learned. We also set the noise value very high, which allows the robot to escape the painful situation, by providing an activity in all the mvts neurons that can be higher than that of the winner. On the other hand, when $\text{pain} = -1$, we set the vigilance very high, allowing the coding of almost all shapes (that is the actual sensorial situation) and the noise

to a very low level. The value chosen for the lower bound of the vigilance is such that learning is possible on the PTM, even in the absence of a pain signal.

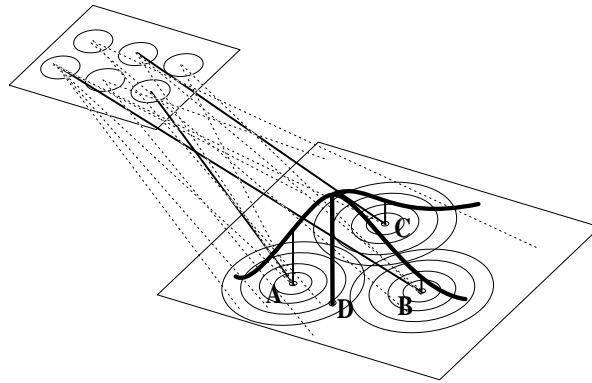


Figure 6. Usage of a PTM for connections with a further neuron group. Three different input patterns have been coded on neurons A, B and C on the PTM. Associations from these cells to the corresponding cells in the further group have also been learned. A new input is presented and its winner is cell D which happens to be closer from A than from B and C. The diffusion emanating from is D is therefore higher in A than in B and C. As a consequence, the most activated neuron in the next group is that previously associated to A.

We have $\text{pain} = 1$ when the last movement provoked a collision, and $\text{pain} = -1$ when the robot moved from a collision position to a position where the collision value (the sum of the value of the front sensors) is lower. This is measured by the number of saturated sensors on the front. In all other cases, $\text{pain} = 0$.

The functioning of the network as expressed above thus enables learning of a new shape only when $\text{pain} = -1$. This particular choice to learn only about the "good things" and not the "bad" appears to be judicious as will be seen in the results section. On the other hand, when $\text{pain} = 1$, we propose to add noise to the movement neurons in order to propose a movement different from the one that causes pain. It is in those cases that the categories boundaries are found: the first proposed movement is the one associated to the closest winner. If it is wrong, then a new pattern is learned, and its association to another movement. This results in the local separation of the set of neurons into two subsets corresponding to two different categories.

Actually the inhibition which consists in lowering a weight value is not really needed, as the movements group is competitive, i.e., it already contains inhibiting lateral links. It was introduced here only to learn faster, as the random process for finding an appropriate solution may take long. Unfortunately, that same process makes it extremely difficult to measure the performances of the robot objectively, according to the parameters. Performances over a given time span always depends on the length taken to find the first solution to a painful situation, as well as on the direction choices at important times. This fact is important, because it illustrates the difficulty of measuring the performances of a single component in a complete N.N architecture for on-line learning on a real robot. In that respect, it is completely different from traditional N.N which can always be tested on well known data distributions and compared to the desired results a posteriori.

Nevertheless, we devised the following experiment: the robot is put in the middle of a somewhat round arena made of pieces of wood which present in fact a lot of different pattern on the robot sensors. We noticed that very soon, it chooses a particular behavior, for instance to turn left while keeping constantly the wall on its right. Then after forty steps, we turned it around and let it learn the other direction. The experiment was performed with several sizes of diffusion, with the same tuning parameters for the map.

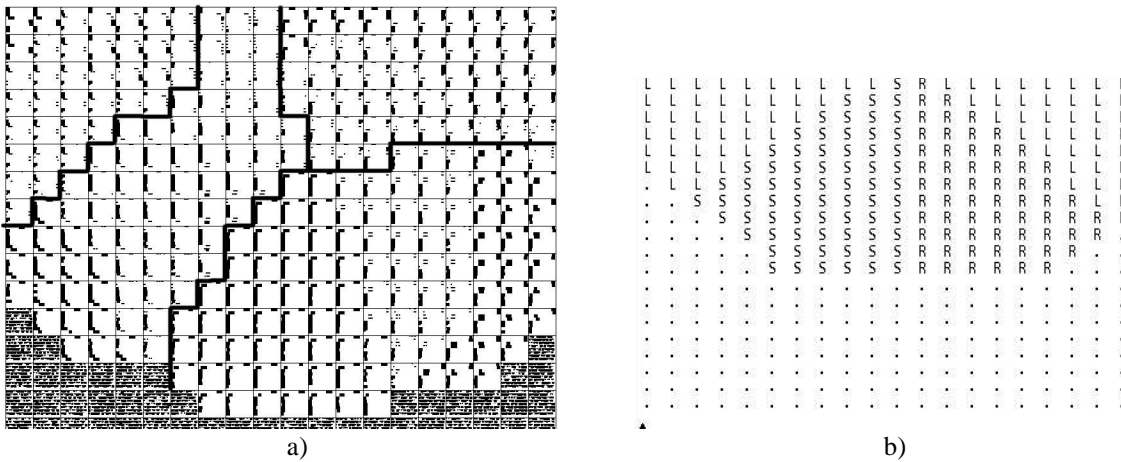


Figure 7: Results of the learning experiment on the PTM. a) The weight vectors of the map's neurons, i.e. the input pattern they code. Thick lines are the frontiers of the subjective interpretation of categories by S. Zrehen. b) The movements associated to each neuron of the map: S straight ahead, L left, R right.

A typical result of learning is depicted on fig. 7. Fig. 7a represents the patterns learned by the PTM. In the square located at the place of the corresponding neuron, the weight vector is represented in the same fashion as for the input. It is therefore possible to see what sensor configuration has been learned by each neuron. On this drawing, we added by hand a subjective interpretation of the categories of sensorial situations. Four categories can be identified by taking into consideration the sensors with a high value. This interpretation is similar to the one performed when analyzing Kohonen maps organized over sets of complex data, such as economic data [BLAYO91]. It should be noted that analyzing the map alone requires a large degree of subjectivity, or at least background knowledge about the data's meaning. In the case of our research, we exclude that subjectivity completely, and concentrate only on the interpretations made by the autonomous agent itself. In the present case, we represented on fig. 7b the categories as seen from the mvts WTA. We see that also four connex zones are present, but it should be noted that their boundaries are not the same as on fig. 7a. This result should also cast questions about interpreting activation of cortical maps in biology: it appears that analyzing only one part of the information processing chain as isolated could lead to mistakes.

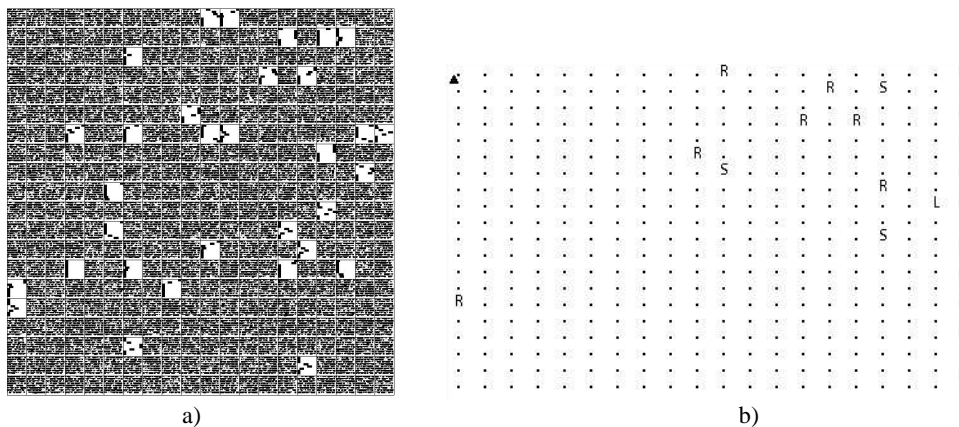


Figure 8: Results of the learning experiment on the PTM with no diffusion, after 80 steps. Many patterns need to be learned, as well as the category they belong to. a) the PTM's neurons weights. b) The movement associated to the neurons. The mvts neurons with a low activity are represented with a dot.

On fig. 8, we represented the patterns learned by the PTM with a null diffusion radius, that is, a WTA. It is clear that categories cannot be defined in that case, even subjectively. Moreover, there is no possible generalization from previous experience when a new pattern is learned: the association with the right movement is proposed completely randomly. Therefore, learning of obstacle avoidance takes much longer than with using topology.

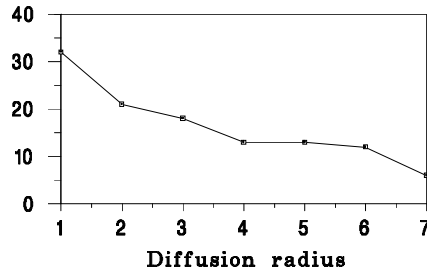


Figure 9: The number of learned shapes in 80 setps, in function of the diffusion radius expressed in distance between neurons.

As we mentioned earlier, it is almost impossible to measure the real performances of the robot over a given time span, because chance plays too big a role. However, one criterion appears to be important, and to decrease with the size of the bubble: the number of learned patterns in an experiment. Indeed, generalization capacities grow with the size of the bubble: the diffusion is extended to more levels. We can see on fig. 9 a typical curve representing the number of learned patterns in function of the bubble size.

The number of patterns learned by the map decreases with the diffusion width, and we have seen that in average, the robot stops colliding extremely fast, with a best performance for a radius of 7. This shows how topological maps help to learn as few patterns as possible for a given task, if parameters are well tuned. Thus, a lot of space is left for learning other patterns. It seems there is an optimum in a given experiment. Nevertheless, it is impossible to know in advance all the possible positions that the robot can meet. Therefore it is important to leave learning possible permanently.

4. Conclusion

We have exposed the advantages of topology preservation for on-line learning. In this respect, we have shown that the PTM can be successfully, and that performances increase with the importance given to topology: less patterns need to be learned to allow a good generalization.

However, the N.N we have proposed is far from complete. A lot of mechanisms could be added to enhance its capabilities. For instance, we used diffusion only in one direction, i.e., for each sensor independently. It should be possible to diffuse from one sensor to the next, thus extracting more topological information from the raw input. On another level, we took a particular activation and weight transformation rule for the WTA to speed up learning. But traditional sum of products activation rules, with the same kind of reinforcement can also be used, and allow many kinds of reinforcement learning: with or without pain or pleasure. At this stage, associations can only be made between events that take place in the same time step. Adding a time integration possibility would allow associations between things that happen at different times, with or without pain or pleasure, thus implementing models of conditioning experiments in psychology, classical or operant. Our choice of binary events: pain, pleasure, collisions is motivated by simplicity, as the main topic of this paper is to illustrate the behavior of a topological map, and not of a complete system. Obviously a

continuous pain with a quenching function and a modulation of pain and stress should lead to richer behaviors.

One point appears important: the complexity of learnable associations. Topology preservation is possible only relatively, as one projects high-dimensional spaces onto a two-dimensional discrete space. However, locally, it is possible if small movements correspond to changes of a little number of variables. This is the case with the coding we chose for Khepera's infrared sensors, and it is the case with any optical flow [GIBSON79]: a small movement can be identified by important changes of a limited number of variables in the flow. If the input are of that nature, then local topology preservation is possible and can be used for identification of subjective categories on the robot's part.

5. Acknowledgments

This research is supported by the Swiss National Funds PNR 23 program.

We would like to thank Francesco Mondada for designing the robot, and for his precious help during this research, Dario Floreano for his precious comments on the manuscript and the French Ministry of Foreign Affairs who allowed Ph. Gaussier to make his civil service at the LAMI.

6. Bibliography

- [BARTO83] Barto A.G., Sutton R.S., & Anderson C.W. (1983) Neuronlike Adaptive elements that can solve difficult learning control problems. IEEE Transactions on Systems, Man, and Cybernetics **SMC-13**, 5, pp 834-846.
- [BLAYO91] Blayo F., Demartines P. (1991) Data Analysis: How to compare Kohonen neural networks to other techniques, Proceedings of the International Workshop on Artificial Neural Networks, Granada, Sept. 17-19.
- [BOURGINE91] Bourgin P., Varela F., 1991, Towards a Practice of Autonomous Systems, Proceedings of the SAB 1991, Bourgin P. Varela F. eds, Paris, MIT Press .
- [BRAITENBERG84] Braitenberg, V. (1984). Vehicles: Experiments in Synthetic Psychology. Cambridge, MA: MIT Press-Bradford Books.
- [BURNOD89] Burnod Y. (1989). An adaptive neural network : The cerebral cortex, Masson, Paris.
- [CARPENTER85] Carpenter G. & Grossberg S. (1985) A Massively Parallel Architecture for a Self-Organizing Neural Pattern Recognition Machine. CGVIP **37**, 54-115.
- [DURBIN89] Durbin R., Mitchison G., (1989), A dimension reduction framework for understanding cortical maps, Nature 343, pp. 644-647.
- [EDELMAN84] Edelman, G., Finkel, L.H. (1984) Neuronal group selection in the cerebral cortex: dynamic aspects of neocortical function, in Neuro-Computing 2, Anderson J.A. Pellionisz A. and Rosenfeld Ed. eds, 1990, pp308-334.
- [GAUSSIER93] Gaussier P. (1993) Simulation d'un système visuel comprenant plusieurs aires corticales, application à l'analyse de scènes, Doctoral Thesis, Orsay Paris XI.
- [GAUSSIER93] Gaussier P., Zrehen S.(1994), The Probabilistic Topological Map: A Self-Organizing and Fast Learning Neural Map that Preserves Topology, LAMI internal report R93.51P.
- [GEORGOPOULOS88] Georgopoulos, A. (1988) Neural interpretation of movement: role of motor cortex in reaching. F.A.S.E.B Journal, **13**, pp. 2846-2857.

- [GHEZ85] Ghez, C. (1985) Voluntary movement, in Principles of Neural Science, Kandel E., Schwartz, J. eds, 2nd edition , Elsevier, Amsterdam, pp. 487-501.
- [GIBSON79] Gibson J.J (1979) The Ecological Approach to Visual Perception. Houghton-Mifflin, Boston
- [HECHT-NIELSEN87] Hecht-Nielsen R. (1987) Counterpropagation Networks. Applied Optics **26**, 23, pp4979-4984.
- [KOHONEN82] Kohonen T. (1982). Self-organized formation of topologically correct feature maps. Biological Cybernetics **43**, 59-69.
- [KOHONEN88] Kohonen T. (1988) The "neural" phonetic typewriter. Computer **21**(3),11-22.
- [LAKOFF87] Lakoff G. (1987) Women, Fire and Dangerous Things: What Categories Reveal about the Mind, The University of Chicago Press, Chicago.
- [MATURANA87] Maturana H., Varela F. (1987) The Tree of Knowledge, The Biological Roots of Human Understanding, Shambhala, Boston
- [MONDADA93] Mondada F., Franzi, Ed., Ienne P., (1993) Mobile Robot Miniaturisation: A Tool For Investigation In Control Algorithms, Proceedings of the Third International Symposium on Experimental Robotics, Kyoto, Oct. 28-30.
- [MOUNTCASTLE75] Mountcastle, V.B., Lynch, J.C., Georgopoulos, A., Sakata, H. and Acuna, C., (1975), Posterior parietal association cortex of the monkey: Command, functions for operations within extrapersonal space, J. Neurophysiol., 38, pp 871-908.
- [OBERMAYER90] Obermayer, K. , Ritter H., Schulten, K. (1990) Large-Scale Simulations of Self-Organizing Neural Networks on Parallel Computers: Application to Biological Modelling, Par. Comp., 14, pp 381-404.
- [O'KEEFE89] O'Keefe (1989) Computations the hippocampus might perform, in Mental Computation, *In Neural Connections, Mental Computation, ed Nadel et al.*, MIT Press, p 225-284.
- [PENFIELD50] Penfield, W., Rasmussen, T. (1950) The Cerebral Cortex of Man: a Clinical Study of Localization of Function, Mc-Millan, New-York.
- [RITTER89] Ritter H. Kohonen T. (1989) Self-Organizing Semantic Maps, Biol. Cyb. **61**, pp 241-254
- [RUMELHART86] Rumelhart D.E. et al. (1986). Parallel Distributed Processing,.. MIT-Press : Cambridge.
- [SEIBERT89] Seibert M., Waxman A., Spreading activation layers, visual saccades and invariant representations for neural pattern recognition systems., Neural Networks 2, pp 9-21.
- [SHAMBES78] Shambes, G.M., Gibson, J.M., Welker, W., 1978, Fractured somatotopy in granule cell tactile areas of rat cerebellar hemispheres revealed by micromapping, Brain. Behav. Evol., 15, pp 94-140.
- [STEWART91] Stewart (1991): Cognition = Life: The Epistemological And Ontological Significance Of Artificial Life, Proceedings of the SAB 1991, Bourguine P. Varela F. eds, Paris, pp.475-483,MIT Press.
- [VERSCHURE92] Verschure P., Krose B., Pfeifer R. (1992). Distributed Adaptive Control: The self-organization of structured behavior. Robotics and Autonomous Agents, 9, pp 181-196.