

A coincidence detection network for spatio-temporal coding: application to nominal composition

Christian JACQUEMIN
I.R.I.N / Université de Nantes
3, rue du Maréchal Joffre
F-44041 NANTES Cedex 01 France
Email jacmin@naiut.dnet.circe.fr

Abstract

This paper describes an architecture for semantic interpretation and classification of nominal phrases. The network, based on the coincidence detection of incoming events, shows an ability to extract relevant linguistic features. Through a combinatorial study, we justify that the topology is rich enough to extract the desired feature associations. It is designed to handle and to combine several linguistic problems. Two particular applications are reported which account for a computational approach of compositionality with implicit knowledge representation:

- The interpretation of locative noun phrases 'cooking+Prep+NP' is inferred from a set of associations of prepositions and noun phrases.
- The semantics of 'Noun+à+Noun' compounds is linked to the acceptabilities of linguistic tests.

1 Introduction

The electrophysiology of the cerebral cortex reports an ability of the neurons to perform spatio-temporal coding. For example, [Abeles, 1982] shows that the firing of a neuron can elicit a spike in another one a few milliseconds later. This time-locking of activity between neurons leads to chains of synchronous firing units: *synfire chains*. This ability to detect event coincidence has already been implemented in formal neural networks, for example [Roques and Beroule, 1991] handles the application of *guided propagation* to sentence parsing.

For symbolic processing a fine observation of the network is necessary to deduce the emergent representations and to give their cognitive counterpart (cf. [Harris, 1990]). Therefore, the network presented here implements coincidence detection rather than correlation. The spikes in our model can be assimilated to bursts of spikes in one using correlation. The learning procedure *naturally* associated with units reacting to synchrony is the Hebbian rule [Little and Shaw, 1975] from [Hebb, 1949]. It enhances the links between synchronously firing units and brings forth the paths detecting spatio-temporal patterns of activity.

The application of a formal network to subtle linguistic problems like categorisation [Lakoff, 1987] and well-

formedness calls for compositional dynamical representations. The topic of nominal composition is chosen for it mixes compositionality and idioms [Jacquemin, 1991].

Some connectionist networks have been applied successfully to the problems raised by noun phrase studies:

- [Jacquemin, 1992] is a network for robust parsing of compound nouns and their modifications.
- [Wermter and Lenhart, 1992] is a hybrid system for handling complex prepositional phrase attachments.
- [Wermter, 1992] is a hybrid model for scanning understanding by assignment of an interpretation.

These models use back-propagation as supervised learning mode. The first one relies on a strong topology with a diffusion of activation in the hidden layer. The connectionist part of the last one is inspired from [Elman, 1988]. The model presented here is a classifier associating each entry with the closest one from the training set

2 The neural network

2.1 The unit

The formal neuron is designed to detect the coincidence of incoming spikes. Incoming connections are excitatory and the unit reacts to the summation of the Excitatory Post-Synaptic Potentials (EPSP). In order to allow for interaction by overlapping, spikes are shaped like the half of a Gaussian curve. The time clicks are close enough to have several updatings during a spike. This ensures a tolerance to small time discrepancies between EPSPs. If the sum of the EPSPs is higher than the threshold, the unit sends a spike to the efferent units (Figure 1). Each triggering is followed by a *refractory period* of low excitability [Abeles, 1982].

The parameters controlling the shape of the spikes are the duration T and the coefficients k, λ ($T = 2, k = \lambda = 1.0$):

$$\begin{aligned} t_0 \leq t \leq t_0 + T & \quad f(t) = k \exp(-\lambda (t - t_0)^2) \\ t_0 + T + 1 \leq t \leq t_0 + 2T & \quad f(t) = 0 \end{aligned}$$

To allow for learning, the amplitude of each EPSP depends by a multiplicative factor on the weight of the connection. Connections have temporal delays so that the temporal topology does not coincide with the spatial one. Hence the units can detect sequences of non simultaneous events (Figure 2).

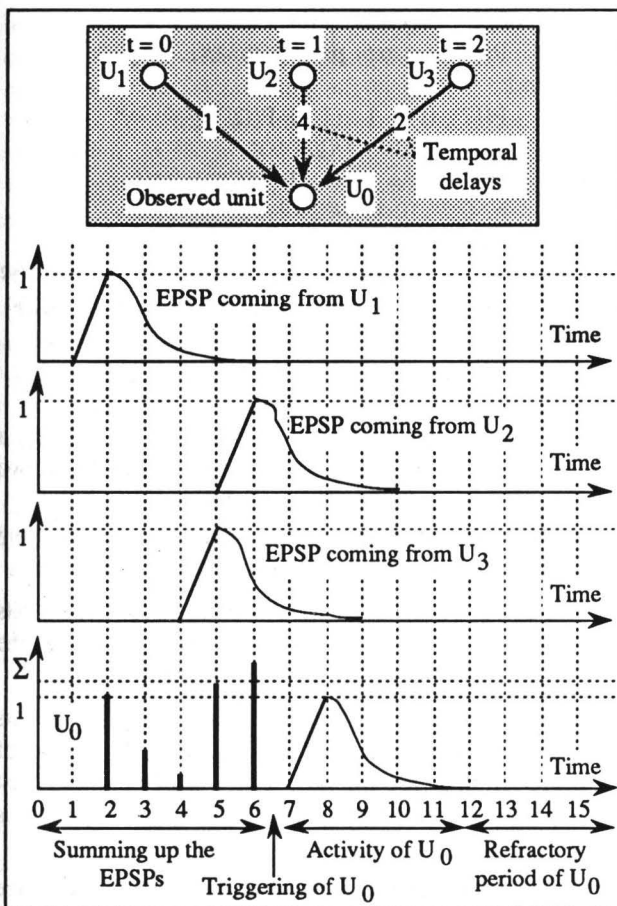


Figure 1. At time 6, when the sum of the induced EPSPs reaches its threshold, U_0 detects the triggering of U_2 and U_3 at times 1 and 2. The connections have a weight of 1.0.

2.2 The topology

With a recurrent topology, repeating spatio-temporal patterns corresponding to cycles of *synfire chains* can be observed [Abeles, 1992]. To understand the basic phenomena, we have chosen a feed-forward organisation with non-overlapping events. The detection of fixed point or limit cycles is replaced by an observation and a comparison of output activities (Figure 2).

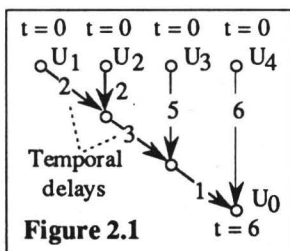


Figure 2.1

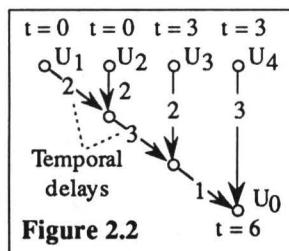


Figure 2.2

Figure 2.1. The simultaneous activities on input units U_1 , U_2 , U_3 and U_4 is detected 6 time steps later by U_0 .

Figure 2.2. The network of Figure 2.1 with different delays detects a sequence of non-simultaneous activities.

The stimuli patterns are fed into the input layer where they trigger the corresponding units. Then, a wave of activity shifts towards the output units. A subnetwork N with one output unit is characteristic of a sequence S of events if the introduction of S leads to a firing of the output unit. The firing should not be provoked by any subsequence of S or with any subnetwork of N .

2.3 The construction

The initial network is built accordingly to the parameter values defining its size and its connectivity. It is composed of C columns built from L layers of R rows (Figure 3).

Each unit which is not an output unit possesses a number of efferent links randomly chosen between C_1 and C_2 . Each link has an initial strength w_0 ($w_0 < 1.0$) and a temporal delay uniformly chosen between C_3 and C_4 . Each link from a unit in the n^{th} layer reaches a unit whose layer is a random number between $n+1$ and L . $(100 \cdot C_5)\%$ of the connections are interior to a column. The other ones reach another column which is chosen randomly.

For the numerical application of 2.4, the values used for the parameters are: $R = 12$, $L = 7$, $C_1 = 5$, $C_2 = 13$, $C_3 = 1$, $C_4 = 5$ and $C_5 = 0.7$. C has two values: 2 for *cuisson*+Prep+NP (used in 2.4) and 3 for Noun+à+Noun.

The optimisation of these parameters has been obtained with a simple genetic algorithm. The performance of a network is the evaluation of the quality of its answers on a training set composed of linguistic and logic examples. The genetic operator is the generation of a new network through small random variations of the parameters from their values in the network with the best performance. The absence of intrinsic parallelism in this algorithm does not prevent it from staying in a local maximum.

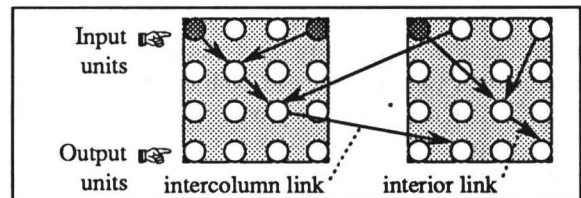


Figure 3. Example of a network composed of 2 columns of 4 layers of 4 units, fed with the pattern 1 0 0 1 1 0 0 0.

2.4 A combinatorial study

This section is intended to check that the ranges of the construction parameters give a rich enough combinatorial potentiality. The general problem being too complex, the question is limited to the detection of a sequence of two activities of input units. Such a sequence is given by i_1 and i_2 the input units and by 0 and δt the times when the activities occur. We calculate the probability of finding a subnetwork characteristic of this particular coincidence. The weights are equal to 0.5 and the thresholds are equal to 1.0 so that a postsynaptic neuron can only be triggered under the synchronous excitation of at least two neurons.

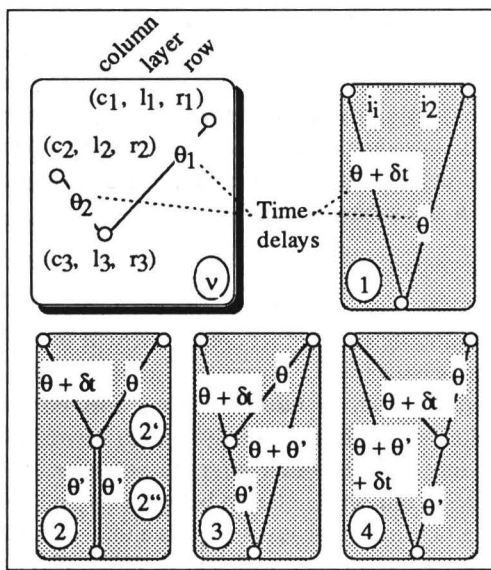


Figure 4. The four networks characteristic of the sequence of activities of input cells i_1 and i_2 at times $t = 0$ and $t = \Delta t$, having at most three layers.

The different subnetworks of two or three layers characteristic of the two events are given by Figure 4. Let us find the probability of having the general subnetwork v :

$$(i) \quad P((c_1, l_1, r_1), (c_2, l_2, r_2), (c_3, l_3, r_3), \theta_1, \theta_2) = 1 - (1 - p_1 p_2)^n \approx n p_1 p_2$$

where p_1 is the probability of finding two efferent links from u_1 and u_2 reaching the same unit u_3 :

if $c_1 = c_2 = c_3$

$$p_1 = (C_5)^2 (1/R^2) 1/(L - l_1)^2$$

if $(c_1 = c_2 \neq c_3)$ or $(c_1 \neq c_2 \text{ and } c_1 \neq c_3 \text{ and } c_2 \neq c_3)$

$$p_1 = (1 - C_5)^2 1/(C - 1)^2 (1/R^2) 1/(L - l_1)^2$$

if $c_1 \neq c_2$ and $(c_1 \text{ or } c_2 = c_3)$

$$p_1 = C_5 (1 - C_5) 1/(C - 1) (1/R^2) 1/(L - l_1)^2$$

where p_2 is the probability that those links have the expected delays (we suppose that both delays belong to the interval $[C_3, C_4]$, otherwise the probability would be null):

$$p_2 = 1/(C_4 - C_3 + 1)^2$$

n is the average number of combinations of efferent links:

$$n = ((C_2 - C_1 + 1)/2)^2.$$

From (i), we deduce the probability of finding a network like (1) (Figure 4):

$$P((1), \theta, \theta + \Delta t) = P((c_1, l_1, r_1), (c_2, l_2, r_2), (c_3, l_3, r_3), \theta, \theta + \Delta t)$$

$$\text{if } c_1 = c_2 \quad P((1), \theta, \theta + \Delta t) = P \approx 9.06 \cdot 10^{-5}$$

$$\text{if } c_1 \neq c_2 \quad P((1), \theta, \theta + \Delta t) = P' \approx 4.72 \cdot 10^{-5}$$

If $\Delta t = 3$, θ can take any value between 1 and 2, c_3 any value between 1 and C and r_3 any value between 1 and R . Hence:

$$\text{if } c_1 = c_2 \quad P((1)) = 1 - (1 - P)^{2 \times C \times R} \approx 2 C R P \approx 4.35 \cdot 10^{-3}$$

$$\text{if } c_1 \neq c_2 \quad P((1)) = 1 - (1 - P')^{2 \times C \times R} \approx 2.27 \cdot 10^{-3}$$

The probabilities of the networks (2), (3) and (4) (Figure 4) are calculated similarly. Each network (x) is divided in two subnetworks, the upper one (x') and the lower one (x'') For a given layer of the hidden unit (l_3), the probability of having a subnetwork (2') is equal to $P((1), \theta, \theta + \Delta t)$. The

probability of (2'') depends on the value of l_3 :

$$P((2'')) \approx n p_1 p_2 \approx 0.0634/(L - l_3)^2$$

For a given l_3 :

$$\text{if } c_1 = c_2 \quad P((2), l_3) = 1 - (1 - P P((2'')))^{2 \times 5 \times C \times R}$$

$$\approx 1.38 \cdot 10^{-3} \cdot 1/(L - l_3)^2$$

$$\text{if } c_1 \neq c_2 \quad P((2), l_3) = 1 - (1 - P' P((2'')))^{2 \times 5 \times C \times R}$$

$$\approx 0.71 \cdot 10^{-3} \cdot 1/(L - l_3)^2$$

When l_3 varies from 2 to $L - 1$:

$$\text{if } c_1 = c_2 \quad P((2)) \approx 1.38 \cdot 10^{-3} \cdot R C \sum (1/(L - l_3)^2) \approx 4.8 \cdot 10^{-2}$$

$$\text{if } c_1 \neq c_2 \quad P((2)) \approx 0.71 \cdot 10^{-3} \cdot R C \sum (1/(L - l_3)^2) \approx 2.5 \cdot 10^{-2}$$

Similarly,

$$\text{if } c_1 = c_2 \quad P((3)) \approx 8.6 \cdot 10^{-3} \text{ and } P((4)) \approx 1.23 \cdot 10^{-3}$$

$$\text{if } c_1 \neq c_2 \quad P((3)) \approx 4.45 \cdot 10^{-3} \text{ and } P((4)) \approx 0.64 \cdot 10^{-3}$$

These results show that the probability of finding a topology characteristic of two events separated by Δt , in two different columns, is 0.025. This value seems low but:

- Taking into consideration the subnetworks of more than 3 layers would slightly increase the probabilities.
- The average number of pairs of features for two 12-bits patterns is 36. The probability of detecting one of the 36 pairs is about 0.6, in the preceding conditions.
- The probability of detecting simultaneous events ($\Delta t = 0$) in the same column is higher: 0.11 instead of 0.025. In this case the probability of detecting at least one cooccurrence of two features in a 12-bits pattern is 0.97.
- The problem can be extended to cooccurrences of three or more events. The subnetworks can also be organised into a hierarchy and detect coincidences of coincidences like the one given in Figure 8, section 3.2.

Nevertheless, the quality of the results depends on the match between the stimuli and the available topology.

2.5 The learning rule

The aim of the learning procedure is to select topologies (and the associated links) which are characteristic of relevant feature cooccurrences. Therefore, the Hebbian learning rule is chosen to enhance the strength of the connections in case of transmission success and reduce them in case of transmission failure. More concretely: let us assume that neuron N_j is linked to N_i by a connection with a delay δ_{ij} and a strength w_{ij} . If N_i triggers δ_{ij} time steps after N_j :

$$w_{ij}(t + 1) = \min(k_1 w_{ij}(t), w_{\max}) \quad k_1 > 1.0$$

and if N_i triggers without activity of N_j δ_{ij} time steps earlier or N_j triggers without activating N_i δ_{ij} time steps later¹:

$$w_{ij}(t + 1) = \max(k_2 w_{ij}(t), w_{\min}) \quad 0 < k_2 < 1.0$$

The value of the weight modification is independent from the distance δ_{ij} between both neurons. It is considered as a local law modifying the weight according to the correlation of an incoming spike and a triggering of the unit without consideration for the distance.

To avoid overactivity, the connections must have an

¹ In [Little and Shaw, 75], transmission failures only occurs when the presynaptic firing does not induce a postsynaptic one.

initial strength w_0 which is low enough to prevent the spike of a presynaptic neuron from activating a postsynaptic one on its own. The initial connectivity is rich, but only a few links are enhanced during training. For example, in the experiment on Noun+à+Noun, out of 1,600 initial links, 1,000 are left unchanged, 450 vanish and 150 are enhanced. In this case only 9% of the links are significant.

2.6 Learning and prediction

The learning is done by introducing repeatedly the training set into the network. 10 epochs are sufficient because the values chosen for the coefficients k_1 and k_2 are high.

The learning procedure enhances the afferent links corresponding to the coincidence of at least two events. After learning, the same coincidence can be detected if only one of both signals is present. This ability to detect partial information can be used for prediction on novel data.

When learning is achieved, the activity of the output units is recorded after each introduction of an input pattern (or of a sequence of patterns). The prediction of the network for a new stimulus relies on a comparison between the recordings of the output activities. The response is the stimulus from the training set which gave the closest activity on output units. The closeness is the number of common activities decreased by the number of distinct activities. Hence, the output units are not the output of the system as seen by the user but an intermediate representation. The final output is obtained through a classic comparison algorithm.

3 Application to nominal composition

3.1 Meaning and acceptabilities of 'N+à+N'

The Noun+à+Noun compounds are particularly interesting because they offer a wide range of interpretations. Some are strongly compositional like *boîte à (chaussures + clous)* [(shoe + nail) box], when others have a conventional and idiomatic meaning like *pied-à-terre*[*pied-à-terre*]. Whatever the compositionality is, the meaning of the whole depends on the meanings of its parts and on the operation performed by the preposition [Cadiot, 1989]. Similar phenomena have been modelled with the help of connectionist networks. For example, [Victorri, 1988] describes the interpretation of the polysemous adverb *encore* [yet, still, again] with a perceptron. It accounts for the different linguistic observations, their continuity and the various ambiguities.

The coincidence detection network can account for the correlations observed between the meaning of a compound noun and the acceptability values of different tests [Cadiot, 1992] (Figure 5).

Each array of acceptability values is fed into the input layer, and the activities of the output units are recorded for 30 time steps. The answer of the network to the novel compound noun is interpreted as its most plausible interpretation(s) (see Figure 6). When fed with the acceptabilities of *verre à bordure dorée*[glass with a golden rim], the network assigns correctly the interpretation *non removable part*. It gives the highest score to *verre à pied*[stemmed glass], the compound from the training set

24 acceptabilities of syntactic modifications + 32 meanings	N1 à N 2 compounds																																		
	Avec N2, j'ai un N1 Pour N2, j'ai un N1 Un N2-ier Un N1 dit à N2 Un N2 N1 à N2 Adj Anaphora of N2 N1 à Dét N2 N1 à Dét0 N2 Overcomposition with/for (N1 à N2 à N'2) for/with (N1 à N'2 à N2) same sense (N1 à N2 à N'2) same sense Un N1 habituellement à N2 Deverbal Le (N1 de Max) à N2 Celui à N2 Le N2 du N1 (unstable) Le N2 du N1 (stable) A N1 à N2 is a sort of N1 A N2, j'ai des N1 J'ai des N1, mais à N2 C'est à N2 que Max veut un N1 Ce N1 est à N2 Locative Instrumental Fuel Incorporation Non removable part Partie/tout non amovible Adverbial of place/time Make N2 from N1 Creation of an object Put N1 in N2 Use N1 for Pred(N2) Put N2 in N1																																		
verre à eau	0	1	0	1	0	0	0	0	1	1	1	0	0	1	0	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	
verre à dents	0	1	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
verre à machine	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
verre à vitre	0	1	0	0	0	0	0	0	1	1	1	0	0	1	0	1	1	0	0	0	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0
verre à pied	1	0	0	0	0	1	1	0	1	1	1	1	1	0	0	0	0	1	0	1	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0
verre au plomb	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	1	1	1	0	0	0	0	0	0	0	0
four à gaz	0	0	0	0	1	1	0	1	1	0	1	0	0	0	0	0	1	1	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0
four à pain	0	1	0	1	0	0	1	1	0	0	1	0	0	1	0	0	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
pied-à-terre	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
arbre à cames	0	0	0	1	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
saut à la perche	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
personne à risque	0	0	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
passage à niveau	0	0	0	1	0	0	0	1	1	0	1	0	1	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0

Figure 5. The acceptabilities (first 24 columns) and the meaning (last 12 columns) of the compound nouns composing the training set. A "1" in the second column of the first row means that *verre à eau*[water glass] accepts the construction *pour l'eau, j'ai un verre*[for water, I have a glass]. The translations are given in Figure 6.

having the same head noun and the same interpretation. However, *four à gaz*[gas oven] obtains a positive score because it has several common acceptabilities with *verre à pied* (and *verre à bordure dorée*). This confirms that the quality of the response depends on the available topology and on the orthogonality of the different codes.

Score	Compound	Interpretation
56	verre à pied	stemmed glass
17	four à gaz	gas oven
-2	verre au plomb	lead glass
-5	verre à vitre	window glass
-11	verre à eau	water glass
-58	four à pain	baker's oven
-60	saut à la perche	pole vaulting
-69	passage à niveau	level crossing
-115	arbre à cames	camshaft
-122	personne à risque	high-risk person
-124	verre à dents	tooth glass
-126	verre à machine	machine glass
-172	pied-à-terre	pied-à-terre

Figure 6. Scores of plausibility of the compound nouns of the training set when *verre à bordure dorée*[glass with golden rim] is introduced with the acceptabilities (1 0 0 0 0 1 1 1 1 1 1 1 0 0 1 1 0 1 0 1 1 1 1).

3.2 Prepositional phrase understanding

The second application concerns the spatial interpretation of *cuisson*[cooking]+Prep+NP (see examples Figure 7). For the sake of clarity, the determiners are not taken into consideration in this experiment, they can be ignored in a first approximation.

Prep	Noun phrase	Scheme
dans	la cocotte	in the casserole
à	feu vif	over a brisk heat
dans	le gril	in the grill
à	l'eau bouillante	in boiling water
à	la cocotte	in the casserole
à	four chaud	in a hot oven
dans	l'eau bouillante	in boiling water
à	la poêle	in the frying pan
dans	le four	in the oven
sur	feu vif	over a brisk heat
dans	3 C.A.S. d'huile	in three tsf of oil
sur	la poêle	in the frying pan
dans	3 C.A.S. d'huile	in three tsf of oil
à	le gril	on the grill
sur	le gril	on the grill
dans	la poêle	in the frying pan
à	le four	in the oven
sous	le gril	under the grill

Figure 7. Spatial schemes associated with each combination Prep+NP. C.A.S.[tsf] is used for *cuillère à soupe* [tablespoonful].

The interpretation of the phrase (Figure 7) is a *spatial schema* [Langacker, 1987] characterised by the relative position of a *trajector*, the thing to cook, and a *landmark*, the place where it is cooked (see also [Vandeloise, 1986]).

The network has to infer the composition rules from a set of representative constructions using a rich enough variety of preposition / noun phrase combinations.

A sequence of words and an interpretation must be processed simultaneously. Such a dynamic binding can be found in works following Elman's recurrent feed forward network [Elman, 1988]. For example, [Jain, 1989] presents a network intended for parsing sentences with embedded clauses where words are represented by semantic micro-features. In other works, the dynamic semantic role assignment is done by a localist network (one unit per word or per symbol) [Cottrell and Small, 1983].

In order to avoid localist or micro-localist representations, the data fed into the networks are the random codes of the constituents (the preposition and the noun phrase) and the meaning (the spatial schema). The coincidence detection allows for direct association between form and meaning [Langacker, 1987]. The topology is composed of two columns of 12 units. Each code is a random binary number between 1 and $2^{12} - 1$. At time step t_0 the code of the preposition is fed into the first layer of the first column. At $t_0 + \delta t$ ($\delta t = 3$), the code of the noun phrase is fed into the first column simultaneously with the code of the spatial schema into the second column. Whatever codes are chosen, provided they are orthogonal enough, the network has to extract the relevant associations between the preposition, the noun phrase and the schema.

Figure 8 shows how the reinforced links of the network represent the direct association between the prepositions *sur* [on] and *sous* [under] and the associated semantics ON and UNDER. Such *coding* prepositions force their semantics to be the interpretation of the whole structure. Conversely *neutral* prepositions like *à* [at] call for an inference mechanism to calculate the meaning from the predicate

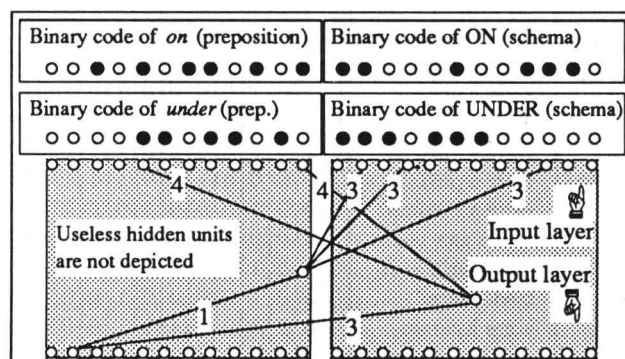


Figure 8. The links of the subnetwork characteristic of the association between a *coding* preposition like *on* or *under* and the associated spatial schema ON or UNDER. The output unit triggers 7 time steps after the introduction of the preposition and 4 time steps after the schema.

(*cuisson* [cooking]) and the argument (the noun phrase) [Cadiot, 1989].

The network can also be fed with incomplete data like structures without semantic interpretation. When fed with the novel compound *sur la cocotte* [in the casserole], it associates correctly the most plausible interpretation *ON* (Figure 9). The three first scores of *sur[on]* are slightly different. This is due to the existence of spurious coincidences between *sur* and the three noun phrases *le gril*[the grill], *la poêle*[the frying pan] and *feu vif*[a brisk heat]. These coincidences may be activated by *la cocotte*[the casserole] depending on the similarity of its code with the ones of the three other noun phrases.

Score	Prep	Noun phrase	Schema
25	sur	le gril	ON
22	sur	la poêle	ON
15	sur	feu vif	ON
-11	sous	le gril	UNDER
-37	à	la cocotte	INSIDE
-37	à	four chaud	INSIDE
-38	à	le four	INSIDE
-42	à	la poêle	ON
-42	à	le gril	ON
-46	à	feu vif	ON
-54	dans	la cocotte	INSIDE
-55	à	l'eau bouillante	INTO
-57	dans	3 C.A.S. d'huile	ON
-61	dans	la poêle	ON
-64	dans	3 C.A.S. d'huile	INTO
-66	dans	l'eau bouillante	INTO
-70	dans	le gril	INSIDE
-75	dans	le four	INSIDE

Figure 9. Scores of plausibility on the training set when *sur la cocotte* [in the casserole] is introduced.

4 Conclusion

This study presents a mechanism for processing complex patterns or sequences of patterns. A rich connectivity and a detection of event coincidences are associated with the Hebbian learning rule. When linguistic data are introduced into the network, the subnetworks corresponding to the relevant features and features associations have their links enhanced. The observation of these subnetworks confirms the predictions of linguistic studies for the modes of composition between noun predicates and arguments. Such a system can be used for complex semantic interpretations depending on interacting linguistic clues. It can help in extracting implicit categorisation from sets of well-formed expressions.

References

[Abeles, 1982] Moshe Abeles. *Local cortical circuits*. Springer Verlag, Berlin, 1982.
[Abeles, 1992] Moshe Abeles. *Corticonics*, Cambridge

University Press, 1992.
[Cadiot, 1989] Pierre Cadiot. La préposition: interprétation par codage et interprétation par inférence. *Cahiers de Grammaire* 14. Toulouse Le Mirail, 1989.
[Cadiot, 1992] Pierre. Cadiot. «à» entre deux noms: vers la composition nominale. *Lexique* 11, P.U.L., Lille, 1992.
[Cottrell and Small, 1983] Garrisson W. Cottrell, S. L. Small A connectionist scheme for modelling word sense desambiguation, *Cognition and Brain Theory*, 6(1), 89-120, 1983.
[Elman, 1988] Jeffrey L Elman. Finding structure in time. (TR 8801 Centre for Research on Language. UCSD, and *Cognitive Science*, 14, 1990), 179-211, 1988.
[Harris, 1990] Cathy Harris. Connectionism and Cognitive Linguistics. *Connection Science* 2(1&2), 7-33, 1990.
[Hebb, 1949] Donald O. Hebb, *The organization of behavior*, Wiley, New York, 1949.
[Jacquemin, 1991] Christian Jacquemin. Une grammaire d'unification des noms composés contrôlée par l'acceptabilité. *Cahiers de Grammaire* 16. Université de Toulouse Le Mirail, 51-71, 1991.
[Jacquemin, 1992] Christian Jacquemin. Activation Diffusion: A Connectionist Network for Robust Parsing. *Proceedings of ECAI 92*. Wiley, 183-187, 1992.
[Jain, 1989] Ajay N. Jain. A Connectionist Architecture for Sequential Symbolic Domains. (TR CMU-CS 89-187). Carnegie Mellon University, 1989.
[Lakoff, 1987] George Lakoff. *Women, Fire and Dangerous Things*. Chicago University Press, 1987.
[Langacker, 1987] Ronald W. Langacker. *Foundations of Cognitive Grammar. Vol I. Theoretical Prerequisites*. Stanford University Press, 1987.
[Little and Shaw, 1975] W.A. Little and Gordon L. Shaw Analytic study of the memory storage capacity of a neural network, *Math. Biosci.* 39, 1975.
[Roques and Beroule, 1991] Muriel Roques and Dominique Béroule. Strategies of Unsupervised Learning for a Parallel Parsing Architecture. In *Proceedings of the IJCNN*, Seattle, 1991.
[Vandeloise, 1986] Claude Vandeloise. *L'espace en français: sémantique des prépositions spatiales*, Seuil, 1986.
[Victorri, 1988] Bernard Victorri. Modéliser la polysémie. *TA informations* 1988 N° 1-2. ATALA, Paris, 1988.
[Wermter, 1992] Stephen Wermter, A hybrid and connectionist architecture for a SCANNing understanding. *Proceedings of ECAI 92*, Wiley, 187-192, 1992.
[Wermter and Lenhart, 1992] Stephen Wermter and Wendy G. Lenhart A Hybrid Symbolic, Connectionist Model for Noun Phrase Understanding. In *Connectionist Natural Language Processing* (Readings from Connection Science), Noel Sharkey, editor, Kluwer Academic Publishers, 1992.