1028

LOGICAL SYNTHESIS OF REGULATORY MODELS.

D. Thieffry and R. Thomas.

Université Libre de Bruxelles - Laboratoire de Génétique.
Rue des chevaux, 67, B-1640 Rhode Saint Genèse, Belgium.
e-mail : dthieffr@ulb.ac.be

Abstract

During these last years, our group has developed a powerful qualitative method which formalises the interactions between the elements of a network in terms of logical (n-level) variables, functions and parameters. This method emphasises the role of feedback loops and has been fully automatized. Until now, it has been mostly used to analyse models for the regulation of viral expression and of the immune response.

In this paper, we discuss the inductive (or synthetic) use of the new concepts of "feedback loop efficiency" and "feedback loop characteristic state" to determine the basic features necessary to obtain defined spatial and temporal patterns. By "inductive" we mean that instead of proceeding from a model to its dynamics we start from the behaviour and try to rationally infer models. We will illustrate our methodology with two simplified but representative examples taken from Neurobiology and Genetics of Development.

INTRODUCTION.

Classically, regulatory networks are treated in either of two diametrically opposed ways : the purely verbal description and the fully quantitative description in terms of (non-linear) differential equations. Intermediate or "logical" approaches have been proposed several times, notably by Rashevsky (1948), Sugita (1961) and Kauffman (1969), in a biological context.

Starting in the seventies, our group has developed a logical formalism which avoids some of the unrealistic aspects of the classical logical description. This work led to the "generalized logical method" which formalises the interactions between the elements of a network in terms of discrete (binary or multi-level) variables, functions and parameters. The salient features of this method are an emphasis on the role and properties of feedback loops, the possibility for the logical variables and functions to take threshold values and a fully asynchronous treatment. Since the generalized logical formalism and method are described in detail elsewhere (Thomas, R. & D'Ari, R., 1990; Thomas, R., 1991), we recall here only the points necessary to understand the contents of this paper.

Feedback loops and their characteristic states

The generalized logical method emphasises the role of the **feedback loops** as fundamental dynamical determinants. In a feedback loop, each element exerts a direct effect only on the following element of the loop but an indirect effect on all elements of the loop, including itself. In a given loop, either all elements exert on themselves a positive effect or all elements exert on themselves a negative effect. For this reason, we distinguish **positive** and **negative** loops ; which one depends on the parity of the number of negative interactions. Negative loops can generate **homeostasis** whereas positive loops can produce **multistationarity**. We say that a loop is "effective", "efficient" or "functional" when it actually generates homeostasis (negative loop) or multistationarity (positive loop).

In a given feedback loop, each variable acts above a characteristic threshold ; one can thus define a loop-characteristic singular state[1] formed by the intersection of the different thresholds involved (for example, if x acts on y above its 2nd threshold ^2s and y on x above its 3rd threshold ^3s, the characteristic state of the loop is simply $^2s^3s$, i. e. $x = {}^2s$, $y = {}^3s$). The main interest of the notion of loop-characteristic state reside in a reciprocal link with the property of loop efficiency : the parametric conditions for a loop to be efficient (i. e. to produce multistationarity or homeostasis) coincide with the condition for which the corresponding characteristic state is steady.

Moreover, it was discovered (see Thomas 1991) and later formally proven (Snoussi and Thomas, 1992) that among the logical states located on one or more threshold ("singular" states), only those which are loop-characteristic can be steady. This enormously simplifies the analysis of complex systems because, in order to identify all the steady states, instead of being obliged to scan all the logical states for steadiness, one just has to consider the regular states and those (usually few) singular states which are loop-characteristic. In fact, as developed below, the analysis focuses on the identification of the feedback loops and their characteristic state, and on the conditions (set of parameter values) which render a state steady (and the characteristic loop efficient).

We have developed **computer programs**, written in *Pascal* and available on *Macintosh, DEC* and *CDC* computers, which automatize the dynamical analysis of complex networks (Thieffry *et al.*, 1993). So far, we have applied our logical method to the analysis of the regulatory network responsible for the lysis/lysogeny choice in bacteriophage λ, as well as to the regulation of HIV and HTLV expression. Another important application deals with the modelisation of the regulation of the immune response (Kaufman & Thomas, 1987 ; Muraille and Kaufman, in preparation).

---

[1]We use "singular states" for the states located on one or more threshold(s) *vs* "regular states" for the others.

Analytic vs synthetic approaches

In the applications mentioned so far, one starts from the model (graph of interactions, itself inspired from experimental data) and the logical work proceeds from the model toward its implications (which in turn will have to be checked with new experiments) ; this is an **analytical** (or deductive) approach.

In this paper, we want discuss another approach, called **inductive** or **synthetic** approach, which proceeds in the opposite way. Starting from a given behaviour, one uses the logical tool to propose models (as simple as possible) which permit or impose this behaviour. This approach has been already discussed by one of us several years before, in the context of a pure Boolean formalism (Thomas, 1979; Thomas and D'Ari, 1990; see also Richelle, in preparation). In the context of two examples taken from Neurobiology and Developmental Biology, we briefly recall the method developed earlier and discuss a recent development based on the knowledge of the properties of feedback loop and on the notion of characteristic steady states.

EXAMPLES OF MODEL SYNTHESIS

A) Modelising an oscillating neural network

This first example comes from well known experimental and theoretical work on neural networks (Friesen & Stent, 1977). In the context of this paper, we have chosen to discuss a simple **oscillating network** comprising four neurons. Let us first recall the treatment proposed initially (Thomas, 1979; Thomas & D'Ari, 1990). If we translate into logical terms the **experimental** "phase diagram" for this four neuron system, we obtain the following logical cycle[2] :

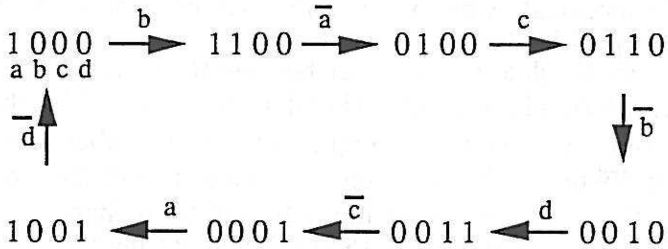


Figure 1. Each group of four digits represents a state of the system. For example, 1000 means that neuron a is on and neurons b, c and d are off. The series of state transitions is closed on itself, constituting a logical cycle.

Given this cycle, one first wishes to find the models in which the system is **forced** to follow the cycle indefinitely once one of the above states is reached. This amounts to saying that for each of the eight states in the cycle, the only possible follower is the next state in the cycle. This gives us the table 1.

In this table, the dashes correspond to unspecified values. This means that for each of the 32 dashes one can arbitrary use a 0 or a 1; each of these 2^{32} models will behave as expected. Among the numerous ways to construct connections that would impose this behaviour, one can find the **simplest** circuits using subtables for each functions (Florine, 1964). For this purpose, one method consists of drawing a two-dimensional truth table for each function, as in table 2.

---

[2]Since the experimental phase diagram mentions a neuron simply "on" or "off", we need here only binary variables.

a b c d	A B C D
0 0 0 0	- - - -
0 0 0 1	1 0 0 1
0 0 1 1	0 0 0 1
0 0 1 0	0 0 1 1
0 1 1 0	0 0 1 0
0 1 1 1	- - - -
0 1 0 1	- - - -
0 1 0 0	0 1 1 0

a b c d	A B C D
1 1 0 0	0 1 0 0
1 1 0 1	- - - -
1 1 1 1	- - - -
1 1 1 0	- - - -
1 0 1 0	- - - -
1 0 1 1	- - - -
1 0 0 1	1 0 0 0
1 0 0 0	1 1 0 0

Table 1. Incomplete state table corresponding to logical cycle of figure 1. The vector formed by the variable a b c and d is called the "state vector" ; the functions A, B, C and D are the prospective values of the corresponding variables and form the "image vector". When the value of the function vector differs with the corresponding value of the state vector, there is a transition order on the corresponding variable(s).

A (cd \ ab)	00	01	11	10
00	-	0	0	1
01	1	-	-	1
11	0	-	-	-
10	0	0	-	-

B (cd \ ab)	00	01	11	10
00	-	1	1	1
01	0	-	-	0
11	0	-	-	-
10	0	0	-	-

C (cd \ ab)	00	01	11	10
00	-	1	0	0
01	0	-	-	0
11	0	-	-	-
10	1	1	-	-

D (cd \ ab)	00	01	11	10
00	-	0	0	0
01	1	-	-	0
11	1	-	-	-
10	1	0	-	-

Table 2.Truth subtables corresponding to the four functions. The frontiers separating 1's and 0's correspond to the most compact logical expressions.

Chosing the most compact logical expression for each function, we obtain the following equations :

Boolean equations :

$A = \overline{b} . \overline{c}$

$B = \overline{c} . \overline{d}$

$C = \overline{a} . \overline{d}$

$D = \overline{a} . \overline{b}$

in which A, B, C and D are Boolean functions; a, b, c and d are Boolean variables; "-" and "." are the Boolean operators "not" and "and".

Generalized logical equations :

$A = d_a(k_1 + k_{1.2}\overline{b} + k_{1.3}\overline{c})$

$B = d_b(k_2 + k_{2.3}\overline{c} + k_{2.4}\overline{d})$

$C = d_c(k_3 + k_{3.1}\overline{a} + k_{3.4}\overline{d})$

$D = d_d(k_4 + k_{4.1}\overline{a} + k_{4.2}\overline{b})$

in which, in addition, d_i represent the operators of discretisation and $k_{i.j}$. the real parameters introduced by Snoussi; here, the "+" represents the arithmetical sum.

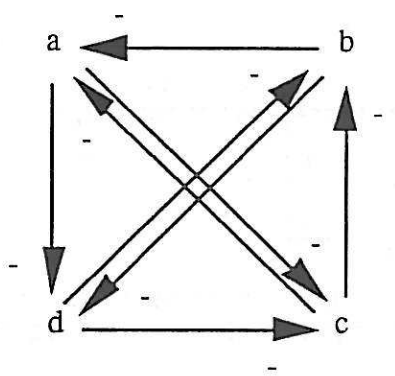
One can easily verify that these equations (corresponding to the figure 2) impose the expected behaviour (logical cycle of figure 1).

In this example, we were looking for models which **impose** a given behaviour. One might also ask what logical circuits will simply **permit** a given behaviour. In fact, one can proceed in a similar way, starting from a less restrictive state table (i. e. a table containing more dashes). This has already been done in the case of the same oscillatory network by Thomas and d'Ari (1990).

Let us now discuss how the notions of feedback loops and characteristic states help us to find the basic features necessary to obtain the same logical cycle. In the center of the logical cycle, one expects to find a singular steady state corresponding to the "focus" of the differential description. In order to have a cycle involving transitions of all four variables, the steady state must be located at the four thresholds, i. e. precisely at $^1s^1s^1s^1s$.

The easiest way to obtain such a steady state, located in the centre of a logical cycle, would be provided by a negative loop involving the four elements. But this doesn't work here, because Friesen and Stent believe, for experimental reasons, that their system involves only **inhibitory** interactions, and we know that a loop composed by four negative interactions has to be positive.

In this context, the simplest way to obtain a logical cycle symmetrically involving all four variable, consists in using **four** three-element negative loops, which, taken together, have a characteristic state in $^1s^1s^1s^1s$ (the corresponding graph and matrix of interactions are given below). For parameter values which render this characteristic state steady and the four negative feedback loop efficient (see appendix), the system behaves exactly as in figure 1. In fact, the graph of interactions corresponds to the generalised logical equations given above, whereas the parametric constraints corresponds to the more restrictive set of Boolean equations.



$$\begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} a & b & c & d \end{array} \\ \begin{pmatrix} 0 & \overline{1} & \overline{1} & 0 \\ 0 & 0 & \overline{1} & \overline{1} \\ \overline{1} & 0 & 0 & \overline{1} \\ \overline{1} & \overline{1} & 0 & 0 \end{pmatrix} \end{array}$$

(A) (B)

Figure 2. Graph (A) and matrix (B) of interactions corresponding to the logical cycle of figure 1.

Summarizing and generalizing our synthetic approach, we propose the following scheme :

1) translation of the temporal pattern into a sequence of logical states ;

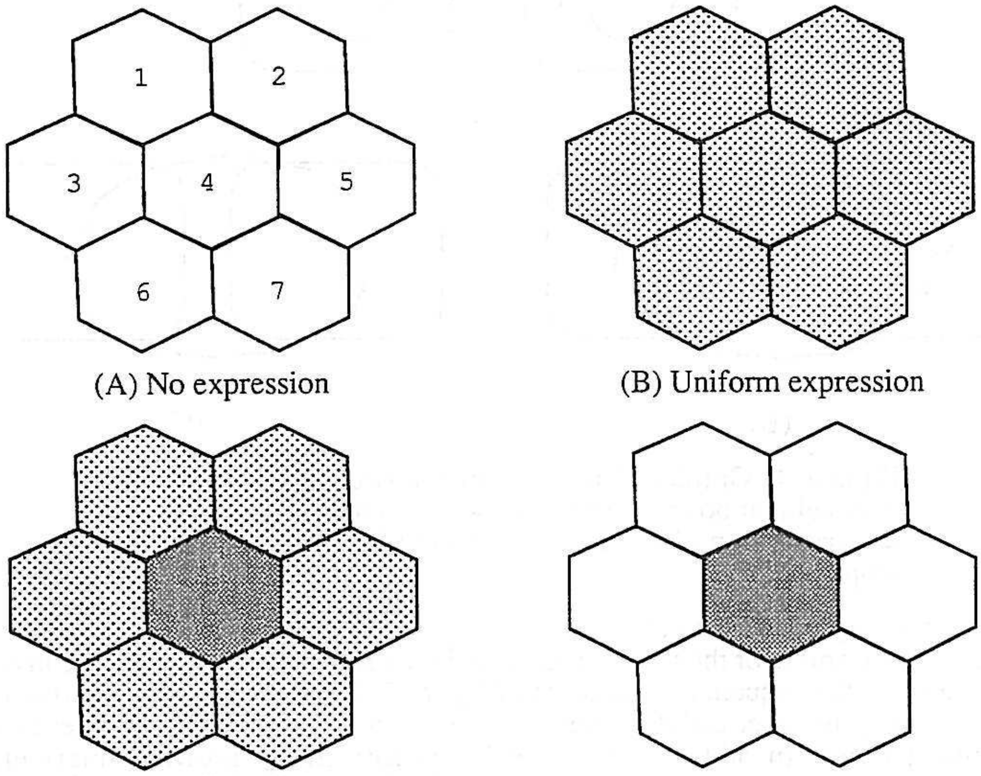
2) if this sequence form a logical cycle, one can localise the corresponding characteristic state ;

3) one can then look for the simplest combination of feedback loops which may produce the required singular state and the corresponding logical cycle (taking into account predefined interactions).

B) Modelising a simple case of pattern formation

Our second example is inspired from a well-studied case of pattern formation in *Drosophila melanogaster* (for reviews, see Ghysen & Dambly-Chaudière, 1989 and 1992). In the larva as well as in the adult fly, the sense organs are arranged in a very precise and reproducible pattern. The determination of this pattern formation seems to involve some kind of preliminary positional information and cellular interactions. The formation of a sense organ implies first the local expression of a set of genes in a group of cells ("proneural cluster"), followed by the commitment of one cell which then laterally inhibits its neighbours and, after a few mitoses, leads to the differentiation of the cells which constitute the sense organ (neuron(s), support cells, etc.). Several genes are involved in these different operations, especially the four genes of the *achaete-scute* complex. The expression of two of them, namely the genes *achaete* and *scute*, is clearly involved in the definition of a proneural cluster as well as in the sorting out of one cell of the cluster. The analysis of mutations and transient expression tests for theses genes indicates that each of them is sufficient to lead to the formation of a sensory organ, but also that the two genes could regulate each other. Nevertheless, the exact interactions and molecular mechanisms involved are still unknown.

In the context of this paper, we apply our synthetic approach to an idealized version of this type of pattern formation. Let us use a convenient two-dimensional symmetrical situation which involves seven identical cells ; by analogy with macrochaete pattern formation in *Drosophila*, let us start with a more or less uniform expression of a gene x in the seven cells ; after some time, the expression of this gene becomes more important in the central cell; finally x expression is inhibited in the peripheral cells and remains only in the central cell (see the figure 3).



(A) No expression (B) Uniform expression

(C) Sorting out of central cell (D) Lateral inhibition

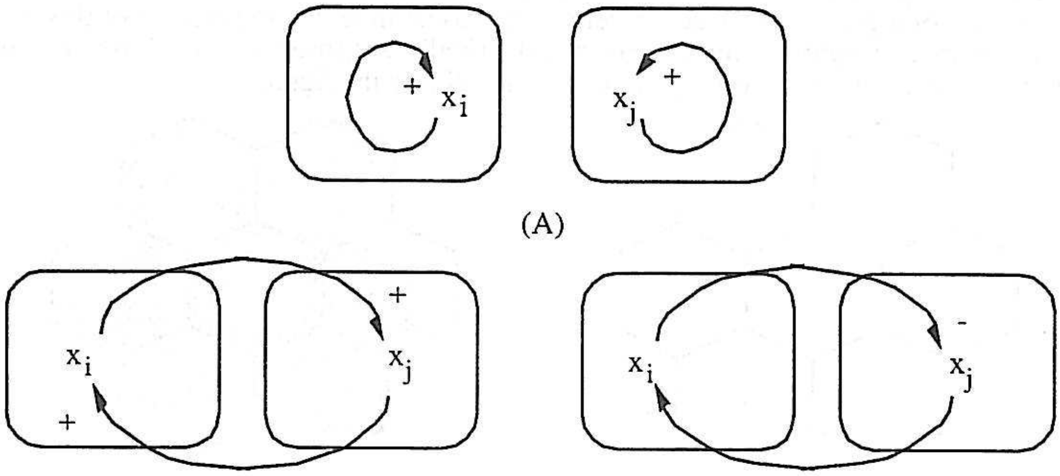
Figure 3. Schematic representation of the main successive steps of a proneural gene expression in a cluster of seven cells (from A to B, to C, to D).

Given this simplified but well defined situation, we now wish to find the basic features which would account for the behaviour of the system. Let us associate a variable x_i to the x gene in each cell. Because all the cells are identical, any interaction proposed in one cell or between two adjacent cells should exist in any of or between any couple of adjacent cells of the system.

The fact that cells which are genetically identical can persist in at least two stable states of gene expression strongly suggest that at least one control gene takes part in a positive feedback loop (direct or indirect autocatalysis). In principle there are two possibilities : either this loop involves only intra-cellular interactions and thus remains functional after cell dissociation ; or this loop involves inter-cellular interactions and thus is broken when cells are dissociated.

Regarding to the last case, there are again two possibilities, according to whether positive or negative interactions are used to construct the inter-cellular positive loops. To these two possibilities correspond clear cut global behaviours :
- On the one hand, positive interactions may lead to cooperative expression of x in the cluster ; starting from a non uniform distribution of x expression, such positive interaction would lead to ubiquitous expression of x in the cluster ; moreover, the central cell being irrigated by a greater number of positive interactions, it could be automatically selected.
- On the other hand, negative inter-cellular interactions may lead to lateral inhibition, i.e. to the inhibition of x expression in the neighbour cells by the cell(s) which express it.

The three types of positive feedback loops ("+", "++", "--") are drawn in the figure 4 (drawn for two adjacent cells).



(A) (B) (C)

Figure 4. Graphs of interaction for two cells i and j. (A) intra-cellular positive feedback loop ; (B) inter-cellular positive feedback loops "++" ; (C) inter-cellular positive feedback loops "--".

Depending on what is given as an entry to our system (to go from A to B in figure 3), one or more of the three types of positive feedback loops have to be involved to account for the sequence of patterns of figure 3. Nevertheless, the positive loops formed by negative inter-cellular interactions seems to be unavoidable in order to reach the final pattern. In addition, positive loops formed by positive inter-cellular interactions could account for a cooperative behaviour of the cells of the cluster, i. e. the reaching of an uniform expression from non uniform entries. Finally, intra-cellular autocatalysis could lead to a certain cell autonomy, so that expressing cells would remain so even after cell dissociation.

Combination of two or three of the positive loops of figure 4 will lead to the combination of their properties for appropriate parameter values. To determine the corresponding range of parameter values, one just has to compute the parametric constraints for each feedback loop to be efficient and, if compatible, fuse them.

Before concluding, let us briefly discuss what are the experimental evidences to involve the three different types of positive loops in respect with the formation of macrochaete pattern in *Drosophila*. (for a review, see Ghysen *et al.*, 1993)

Inhibitory inter-cellular interaction has been involved at several occasions, especially to explain the regular spacing between sensory bristles (see for example Wigglesworth, 1940 ; Richelle and Ghysen, 1979 ; Simpson, 1990). Such inter-cellular inhibition could take place either before ("mutual inhibition") or after ("lateral inhibition") the single out of one cell, but likely could lead to the same final pattern. Recently, two genes, *delta* and *notch*, coding for transmembrane protein, have been proposed to mediate this inter-cellular communication : moreover, *delta* seems to be positively controlled by *achaete-scute*, whereas *notch* could negatively control *achaete-scute*. Other genes seems to be involved in the signal transduction.

Until now, there is no experimental evidence supporting activatory inter-cellular interaction. Thus, cooperative expression doesn't seem to be required in macrochaete pattern formation.

There are a set of experimental evidences in favour of intra-cellular autocatalysis. Effectively, both *achaete* and *scute* seem to be able to activate their own expression ; moreover, they might mutually activate each other. This self activation could constitute the trigger between "competence" and "determination" states. If this is true, once a cell has been triggered to the determined state, it could remains so independently of its neighbours, even after cell dissociation. This is presently under experimental investigation by J. van Helden in the laboratory of A. Ghysen.

To sum up, the present data suggest a model containing both intra-cellular autocatalysis and inter-cellular inhibition. For two cells, we obtain the following graph[3] :

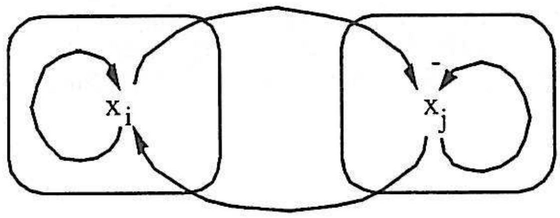


Figure 5. Graph containing the experimental interactions : intra-cellular autocatalysis and inhibitory inter-cellular interactions. For two cells, we thus have 3 positive feedback loops.

ACKNOWLEDGMENTS

We thanks J. van Helden and A. Ghysen for critical readings and sharing ideas about unpublished experimental results. D. T. has been supported by a Grant *Télévie* from the *Fonds National de la Recherche Scientifique*. R. T. acknowledges financial support from *Actions de Recherche Concertée*, Solvay Company and *Fonds National de la Recherche Scientifique*.

REFERENCES

Florine, J. (1964). *La synthèse des machines logiques et son automatisation*, Dunod, Paris.

Friesen, W. O. and Stent, G. S. (1977). Generation of a locomotory rhythm by a neural network with recurrent cyclic inhibition. *Biol. Cybernet.* **28**, p. 40.

Ghysen, A. and Dambly-Chaudière, C. (1989). Genesis of the Drosophila peripheral nervous system. *Trends in Genetics*, **5**, p. 251.

Ghysen, A. and Dambly-Chaudière, C. (1992). Development of the peripheral nervous system in Drosophila. In *Determinants of neuronal identity*, p. 225.

Ghysen, A., Dambly-Chaudière, C., Jan, L. Y. and Jan, Y-N. (1993). Cell interactions and gene interactions in peripheral neurogenesis. *Genes & Development*. In press.

Glass, L., Kauffman, S.A. (1973). The logical analysis of continuous non-linear biochemical control networks. *J. theor. Biol.* **39**, p. 103.

[3]Incidentally, this graph (containing 3 positive feedback loops) can account for up to 11 steady states.

Kauffman, S.A. (1969). Metabolic stability and epigenesis in randomly constructed genetic nets. *J. theor. Biol.* **22**, p. 437.
Kaufman, M. and Thomas R. (1987). Model analysis of the bases of multistationarity in the humoral immune response. *J. Theor. Biol.*, **129**, 141.
Rashevsky, N. (1948). *Mathematical Biophysics*. Univ. of Chicago Press.
Richelle, J. and Ghysen, A. 1979. Determination of sensory bristles and pattern formation in *Drosophila. Dev. Biol.* **70**, p. 418.
Richelle, J. Differentiation and cellular automata. In preparation.
Snoussi, E. H. (1989). Qualitative Dynamics of Piece-Linear Differential Equations: A Discrete Mapping Approach, *Dyn. Stability Syst.* **4**, p. 189.
Snoussi, E. H. and Thomas, R. (1993). Logical identification of all steady states : the concept of feedback loop characteristic states. *Bul. Math. Biol.* In press.
Sugita, M. (1961). Functional analysis of chemical systems in vivo using a logical circuit equivalent. *J. theor. Biol.* **1**, p. 415.
Thomas, R. (ed.) (1979). Kinetic logic: a boolean approach to the ana-lysis of complex regulatory systems. *Lect. Notes Biomath.* **29**, p. 507.
Thomas, R. (1991). Regulatory networks seen as asynchronous automata : a logical description, *J. Theor. Biol.* **153**, 1-23.
Thomas, R. and D'Ari, R. (1990). *Biological Feedback*, CRC-Press.
Thieffry, D. L., M. Colet and Thomas, R. (1993). Logical analysis of regulatory networks, *Math. Modelling. and Scient. Computing*. In press.
Wigglesworth, V. B. (1940). Local and general factors in the development of "patterns" in Rhodnius proxilus (Hemiptera). *J. Exp. Biol.* **17**, p. 180.

APPENDIX

Notations :

K_i is the equivalent of the independant term in differential equations ; it corresponds to what biologists call the **basal** level of expression ;

$K_{i.j}$ corresponds to a positive effect (presence of activator or absence of inhibitor) of the variable j on the variable i;

$K_{i.jk}$ corresponds to joint positive effects of variables j and k ; ...

Oscillatory neural network

The constraints on the logical parameter for which the three negative tree-element loops are efficient and the state ${}^1s^1s^1s^1s$ is steady are simply :.

$K_1 = K_{1.1} = K_{1.2} = 0, K_{1.23} = 1$
$K_2 = K_{2.3} = K_{2.4} = 0, K_{2.34} = 1$
$K_3 = K_{3.1} = K_{3.4} = 0, K_{3.14} = 1$
$K_4 = K_{4.1} = K_{4.2} = 0, K_{4.12} = 1$

For this parameter values, the system follows the cycle presented on the figure 1