

Efficient Attractor Analysis Based on Self-Dependent Subsets of Elements-An Application to Signal Transduction Studies.

Maura Cárdenas-García^{a1} Jaime Lagunez-Otero^{a2} and Nikolai Korneev^{b3}.

^a Instituto de Quimica, UNAM, Ciudad Universitaria, Circuito Exterior CP 04510, Coyoacán, México, D.F. (525) 6224424 Fax (525) 6162203, ^b Instituto Nacional de Astrofísica Óptica y Electrónica, Apartado Postal 51 y 216, Sta. María Tonanzintla Puebla (522) 2472011 Fax (522) 2472940, ¹ maura@servidor.unam.mx, ² lagunez@servidor.unam.mx, ³ korneev@inaoep.mx

Abstract

External signals are transmitted to the cells through receptors activating signal transduction pathways. These pathways form a complicated interconnected network, which is able to answer to different stimuli.

Here we analyze an important pathway for oncogenesis namely *RAS/MAPK signal transduction pathway*. We show that the interaction of the elements of this pathway induces topological structure in the element set and that the knowledge of the topology simplifies the analysis of the set. With a computer algorithm, we isolate from a large and complex group, smaller, independent, more manageable subsets, and build their hierarchy.

Subsets introduction makes easier the search for attractors in discrete dynamical system, it permits the prediction of final states for elements involved in signal transduction pathways.

Introduction

Biological systems often include a large number of interacting elements. As an example, consider proteins in a cell, which are involved in different reactions. Protein A can be produced as a result of the interaction of protein B and C. Further, protein A can be involved in reactions determine the state of protein D and E, etc.

In Figure 1, we present a model of interactions, which are specific to a RAS/MAPK signal transduction pathway. This pathway is important because some elements are proto-oncogenes and others are mutated in a high number of cancers (Burgering and Bos, 1995; Cárdenas-García, 2000; Kawasaki et al., 1996; Khosravi-Far et al., Lowy, 1993; Macara, 1991; Wasylyk, Hagman and Gutierrez-Hartmann, 1998).

The arrow from one number to another indicates that the first element “determines” the state of the second one, whether this is, activation, phosphorylation or complex formation. If the element has two or more arrows reaching it, this means that its state is determined by a cooperative action of all the elements connected to it, or each one determines a state. Not all members of this pathway are proteins, so we will prefer the term *elements* in the text.

Figure 1 looks chaotic, and it is desirable to organize it. One can try to divide interacting elements into subgroups and analyze them separately. In this paper we give a mathematical procedure leading to such organization.

In our paradigm, all pathway elements constitute the *set*. We introduce the concept of *self-determined subset (SDS)*, i.e. subsets of the elements which do not depend

on any element which is not included in the subset. Any *SDS* is autonomous. It is entirely independent on all other elements of the system and can be analyzed separately. On the other hand, any *SDS* acts on the elements, which are not included in it as an external influence. In this paper we discuss how one can systematically find all self-defined sets of a system and show that they form a topology. Knowing the set topology, we show how to introduce hierarchic structure.

Discrete dynamical systems such as cellular automata or random Boolean networks are used to model complicated biological processes with a big number of elements involved (Paton 1993, Somogyii 1996, Kholodenko et al., 1997, Lagunez-Otero, et al., 1996, 1998).

In discrete dynamical system, every element has *states*, which can be numbered. The time development is modeled by steps, and the state of elements at the step $k+1$ can be determined by logical rules from states of elements at the step k .

It is well known that the *final state* of such a system is either a steady one, or a limit cycle. The space of initial states can be divided into basins of attraction, so that initial states corresponding to every basin lead to the same final state or cycle.

If the system has around 20 elements with two states each, the calculation of basins and final states can be done directly by taking each possible initial state and looking for its final one. If the number of elements is larger, the direct approach is not practical because of the prohibitively heavy calculation required.

We consider a system including 41 elements. The direct analysis of this system using Pentium PC is estimated to take some months of calculation, yet we are able to do it in less than 1 hour. The approach we use is to isolate subsets, which are self-determined and analyze them separately.

The number of final states of the subset is usually much smaller than the number of its initial states. These final states can be used as external conditions to determine final states of elements, which do not belong to *SDS*.

RAS/MAPK Signal Transduction Pathway.

Each cell receives a great number of chemical signals generated by other cells or by themselves. These signals regulate their metabolism, determine their differentiation, and indicate when to divide or when enter a death process. In general the external signals are transmitted to the cells through receptors activating signal transduction, which can follow a same way and to generate a final answer or to branch and give rise to the same answer or to others. The RAS/MAPK signaling pathway is a complex mechanism involving the activation of protein kinase cascades. In a cell it is used to transmit signals from

growth-factor receptors protein-tyrosine kinases to transcription factors such TCF/Elk-1 and AP-1 which regulate the genes transcription induction. This pathway is particularly important because some elements are proto-oncogenes and others are mutated in a high number of cancers, (Hesketh 1994; Lacal and McCormick 1993).

In Figure 1 we present the model of element interaction, each number corresponds to a element. In Table 1 we show the names of 41 elements and their dependencies, (Denhardt 1996; Yamamoto, Taya y Kaibuchi 1999).

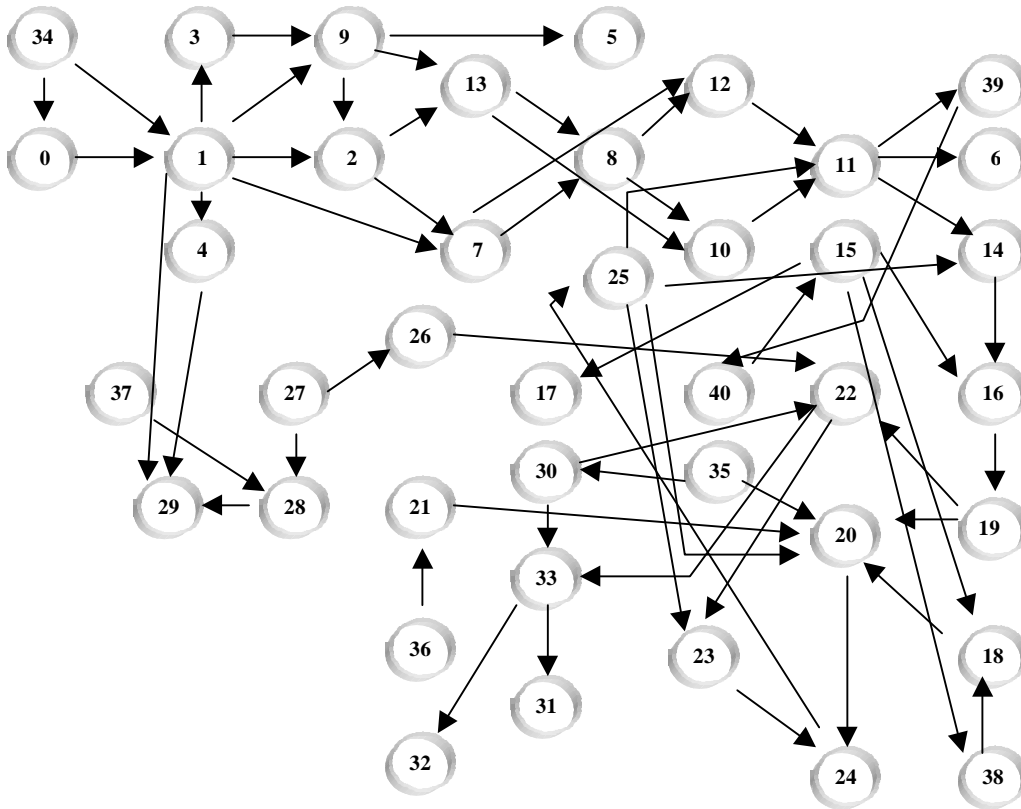


Figure 1. RAS/MAPK signalling pathway interaction model. Numbers are individual elements, arrows represent relation between two elements, two or more arrows mean cooperative or alternative action.

Table 1. RAS/MAPK signal transduction elements and their dependencies. The numbers correspond to Figure 1.

Number	Element name	Dependes on
0	EGFR	EGF
1	EGF/EGFR	EGF EGFR
2	Grb2	EGF/EGFR EGF/EGFR/Shc
3	Shc	EGF/EGFR
4	PLCgamma	EGF/EGFR
5	PI3K	EGF/EGFR/Shc
6	GAP	Ras
7	EGF/EGFR/Grb2	EGF/EGFR Grb2
8	Sos	EGF/EGFR/Grb2
9	EGF/EGFR/Shc	EGF/EGFR Shc
10	EGF/EGFR/Shc/Grb2/Sos	EGF/EGFR/Shc/Grb2 Sos
11	Ras	EGF/EGFR/Shc/Grb2/Sos EGF/EGFR/Grb2/Sos TRE
12	EGF/EGFR/Grb2/Sos	EGF/EGFR/Grb2 Sos
13	EGF/EGFR/Shc/Grb2	EGF/EGFR/Shc Grb2
14	Raf1	Ras TRE
15	MEKK1	PAK
16	MEK	Raf1 MEKK1
17	JNKK	MEKK1
18	JNK	MEK4
19	ERK	MEK
20	cJun	ERK JNK CKII Phtasa TRE
21	Phtasa	TPA
22	Elk1	ERK SRF PKC
23	cFos	Elk1 TRE
24	AP1	cJun cFos
25	TRE	AP1
26	PKC	DG
27	DG	IP3
28	IP3/DG	IP3 DG
29	PLCgamma/EGFR/EGF	PLCgamma EGF/EGFR IP3/DG
30	SRF	CKII
31	EBS	SRF/Elk1
32	SRE	SRF/Elk1
33	SRF/Elk1	SRF Elk1
34	EGF	EGF
35	CKII	CKII
36	TPA	TPA
37	IP3	DG
38	MEK4	MEKK1
39	Rac	Ras
40	PAK	Rac

Topology of interacting elements

In all the following we treat only finite sets. We say that the interaction is introduced on the set if we know for any element A the elements A, B, C, D etc which *determine* this element. We postulate that the element is always determined at least by itself.

The *self-determined subset (SDS)* is a subset, whose elements are determined only by elements, which also belong to the same subset. The empty set is *SDS* by

definition. The most important properties of *SDS*'s are resumed in the following statements:

Statement 1.

The intersection or union of any two *SDS*s is also a *SDS*.

That the elements of union are self-determined is quite evident.

For the intersection the proof is the following: Let us consider the element A, which belongs to two *SDS*'s S_1 and S_2 . If A depends on the element B which belongs to S_1 only, then S_2 cannot be self-determined since there is the element B which determines the element from S_2 but does not belong to S_2 . Thus, the element belonging to intersection can depend only on elements also belonging to the intersection, which proves the statement.

Statement 2.

All *SDS* constitute a topology of the set.

We can identify self-determined subsets with open subsets. By statement 1, any intersection and union of *SDS*s is also *SDS*, thus this property of topology is fulfilled, (O'Neil, 1982). The whole set is obviously self-determined, the empty set is *SDS* as well. Thus, all usual requirements for topology are satisfied

Now we give an explicit way to construct the topology. For this let us define the *minimal SDS* including the given element A, as the intersection of all *SDS* which have A as their element. In topological terms, it is the minimal neighborhood of A which always exists for a finite set.

Practically, the minimal SDS of A can be constructed by making a list which at the first step includes only A, on the second step includes all elements which define A, on the 3rd step includes all elements, which define elements defining A etc. Finally, the list will stop to grow and it will be the minimal *SDS* of A. Extreme cases are the whole set, or only the element A, if it does not depend on anything except itself. It is possible that two different elements A and B can have the same minimal *SDS*.

Statement 3.

Any *SDS* can be represented as a union of minimal *SDS*s.

If the element A is included in S, then the minimal *SDS* of A also belongs to S. Thus, S is the union of minimal *SDS*s of all its elements.

From statement 3 it follows that minimal *SDS* form what is called a *base* of topology (O'Neil, 1982).

Thus, from the practical point of view, to know all *SDS* of the set it is sufficient to determine minimal ones; the unions give all the others.

Statement 4.

It is possible to build a system of dependencies corresponding to any given topology.

The procedure for doing this is presented. For a given finite topology we can determine a minimal open subset including the element A as the intersection of all open subsets including A. Then we can say that A is defined by all elements which enter in this minimal subset. Repeating the procedure for all elements, it is possible to determine all interactions. It is easy to see that generally the interaction corresponding to a given topology is not unique (Fig.2), thus the procedure gives only one of possible interactions.

Statements 2 and 4 mean that for finite systems the topological structure and interaction between system elements are directly related one to another.

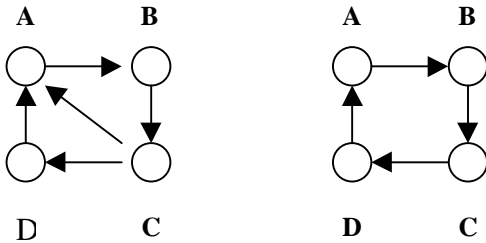


Figure 2. These two interactions have the same topology.

Building the hierarchical structure.

Taking all *SDS*s as we did above, gives the most complete, or the richest topology. The disadvantage of it is that two *SDS* can share some elements. It can be convenient to choose a smaller topology by uniting some *SDS*, so that the resulting division has the following property:

Two *SDS*s have either zero intersection or the intersection is equal to one of them. Such topology base introduces in the system the hierarchy; and we give here a method of building it.

First, let us define the 0^{th} order *SDS* as a minimal one that has no *SDS* inside. The intersection of two different 0^{th} order *SDS* is empty (otherwise the intersection *SDS* will lie inside intersecting subsets, which contradicts the definition).

There is at least one 0^{th} order *SDS* in a set.

Proof: Let us take any *SDS*. If it has no *SDS* within, it is what is needed. If it does, we repeat the search, recursively, for the inner *SDS*. Finally, as the set is finite we will arrive to the 0^{th} order *SDS* we are looking for.

The 1^{st} order minimal *SDS* is defined as having only 0^{th} order *SDS* inside. Intersection of two 1^{st} order minimal subsets is the 0^{th} order or empty. Similarly to Statement 5, if the whole set is not 0^{th} order, there will always be at least one 1^{st} order *SDS*. Generally N^{th} order minimal *SDS* has only *SDS* up to $(N-1)^{th}$ order inside.

Now let us construct the hierarchy. We start from 0^{th} order subsets. At the next step, we add 1^{st} order minimal

subsets and unite all of them which have non-empty intersections. As a result, we get subsets formed of minimal ones with order less than 2 that do not intersect. At the next step we add 2^{nd} order minimal subsets and unite them in the same way. Finally, we will get the base with desired property. The structure can be visualized as a topographic map with height levels. The subset of elements within any height line is self-determined. Elements, which correspond to mountain peaks, have the highest rank, they are determined only by themselves, the higher ones govern elements that lie lower.

For a simple demonstration of the approach we have chosen a model system of 13 elements with arbitrarily taken interactions between them. This system is summarized in Table 1. In Table 2 we give all minimal *SDS* with their orders. The corresponding hierarchical structure is presented in figure 3.

Table 2. Interaction table for the example model.

Element number	Element name	Depends on
0	A	L
1	B	A M
2	C	B
3	D	B C F
4	E	M
5	F	E
6	G	M
7	H	GF
8	I	DH
9	J	I
10	K	M
11	L	K
12	M	M

Table 3. Minimal *SDS* and their orders.

	Elements	Order
1	0, 11, 10, 12	3
2	1, 0, 12, 11, 10	4
3	2, 1, 0, 12, 11, 10	5
4	3, 1, 2, 5, 0, 12, 4, 11, 10	6
5	4, 12	1
6	5, 4, 12	2
7	6, 12	1
8	7, 6, 5, 12, 4	3
9	8, 3, 7, 1, 2, 5, 6, 0, 12, 4, 11, 10	7
10	9, 8, 3, 7, 1, 2, 5, 6, 0, 12, 4, 11, 10	8
11	10, 12	1
12	11, 10, 12	2
13	12	0

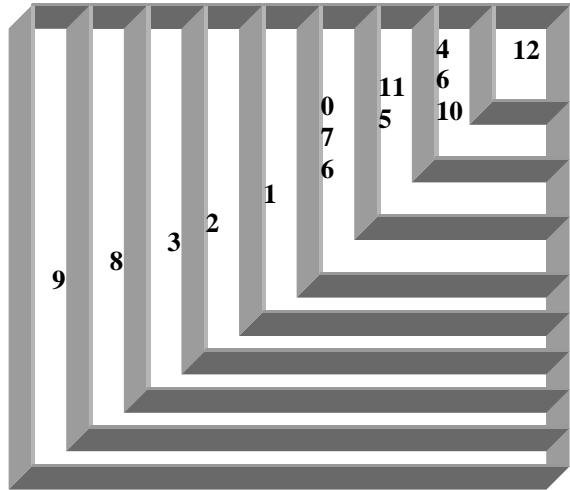


Figure 3. The hierarchical structure.

For the model with 13 elements all necessary operations can be made by hand, but for larger systems this becomes impractical. We have written a computer program doing the job. The algorithm input is the interaction table giving for any element the elements, which determine it. The next step is the calculation of minimal *SDS* for all elements. After this, the order is calculated for all minimal *SDS* and they are arranged into the hierarchical structure.

Topology of RAS/MAPK signal transduction pathway and its application to the attractor finding.

We are interested in describing the element interaction dynamics of the RAS/MAPK pathway with a discrete model, so we assume that any of elements can be found in one of two states (1 or 0), present or absent, active or inactive, interacting or non-interacting, phosphorylated or desphosphorylated, etc. As discussed earlier, we are interested in determining all possible final states of elements involved in a cell signal transduction pathway. To obtain them we take initial states and apply logical rules, which model interactions between elements. Examples of such rules are:

// rule 0, element is activated by another
 // rule 1, element is activated when 2 others are active
 // rule 2, element is activated by any of 2 others, but not activated when both are present
 etc.

The rules are applied to all elements at every step. After some steps the system arrives to equilibrium (states of elements do not change anymore). These final equilibrium states are called attractors of the network. For 41

elements, the number of possible initial states is 2^{41} . This makes the exhaustive search unpractical.

The result of introducing to RAS/MAPK signal transduction pathway the hierarchical structure is presented in Figure 4. Every rectangle in this figure has self-dependent subset inside and can be analysed independently on all the elements which lie outside of this rectangle. The complexity of the task grows with a number of elements N as 2^N . Thus, producing smaller groups, which can be analyzed independently one from another, we can greatly simplify the analysis. We introduce the internal subset which is self-determined, the middle subset which depends on the internal one, and the external which depends on the middle. The size of these subsets is determined by the fact that approximately 20 elements can be analysed at once. Making smaller groups will produce more intermediate steps which seems to be a complication.

The mathematical division in subsets follows from the abstract procedure. Nevertheless, it reflects some functional dependences as well. In particular, the membrane complex which is triggered by the EGF (34) is contained in the inner *SDS*. In the external group, some elements are transcription factors or closely related to them.

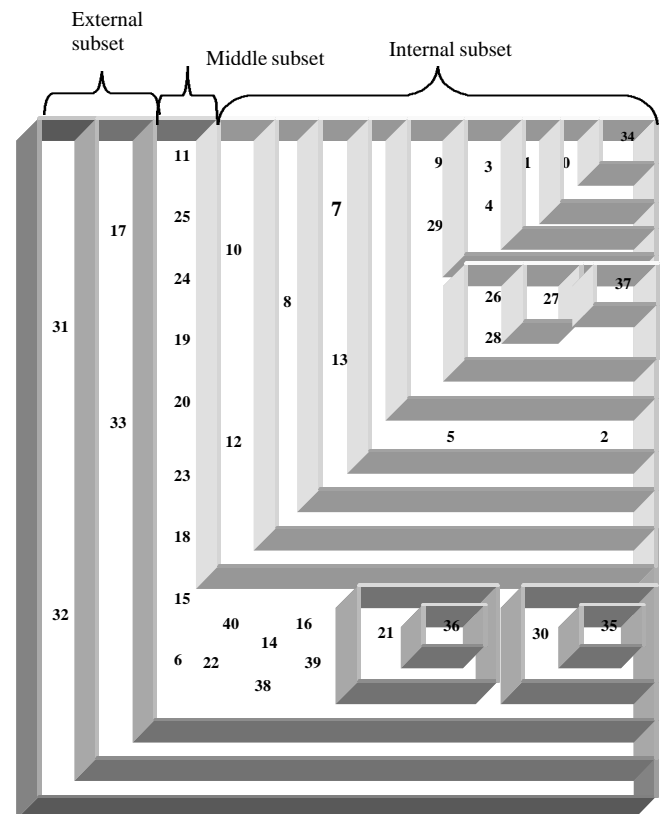


Figure 4. The hierarchical structure of RAS/MAPK signal transduction pathway.

Thus, the general hierarchy corresponds to the path from the stimuli to the response. Similar tendencies are observed for the hierarchy we made for a more complete model of the same pathway which includes 121 elements.

To determine all possible final states of elements we start with the internal *SDS*. For it, this can be done because it is sufficiently small and independent on all the rest.

The number of possibilities depends on the state of the key element EGF (number 34). If EGF is present (in the state 1), we obtain for the internal group only 16 possible final states, listed as columns in Table 5.

Table 4. Possible final states (F) for the internal subset elements with a number EN.

EN	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F
0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
37	0	1	0	1	0	1	0	1	0	0	1	0	1	0	0	1
27	0	0	1	1	0	0	0	0	1	0	0	0	0	1	0	0
26	0	0	1	1	1	1	0	0	1	1	1	0	0	1	1	1
28	0	0	0	1	0	0	1	1	1	1	1	0	0	0	0	0
9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
29	0	0	0	1	0	0	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
34	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Knowing the final states of the internal subset, it is possible to fix them and to determine states of the middle and external subsets by making exhaustive search in them. The analysis is further simplified because, for example, only elements 10, 12 and 26 from the internal group affect the middle group elements. It is seen that elements 10 and 12 always have the final state 1.

Element	Final States															
25	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
24	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
22	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
40	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0
6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
39	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14	1	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1
19	1	1	0	1	1	1	0	1	1	1	0	1	1	1	0	1
30	0	0	0	0	1	1	1	1	1	1	0	0	0	1	1	1
35	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1
36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
38	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
23	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
26	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Table 5. Middle subset possible final states.

We obtained that for the middle group, if the elements 18, 20, 23 and 26, have a state 1, there are 36 final states. They are listed in the table 3.

Looking at these 36 final states, it is interesting to note that phosphatases and kinases (16, 19, 21) vary their activity. Ras (number 11), known as oncogene, is always active. This is also observed in microarrays experiments

Discussion

We succeeded in determining all possible final states of a big system of interacting elements by isolating self-dependent subsets. The final state of the subset depends on the elements of the subset only. We suggest an algorithm for determination of such groups. It is shown that the method can be useful for the analysis of complicated biological systems.

Subsets introduction makes easier attractor search for discrete dynamical systems (Boolean networks), this approach permits the prediction of final states for elements involved in signal transduction pathways and facilitate possible therapeutic target detection.

The model we report is a simplification of reality. Many important elements are not included into analysis, and the validity of introducing discrete states governed by logic rules is by itself questionable. In particular, some of the final states that are predicted cannot occur in the nature, or occur but do not seem probable.

Nevertheless, the example that we consider demonstrates that the topological procedure suggested to introduce self-dependent subsets of elements for their subsequent analysis could be useful for complicated biological systems.

References.

- Burgering, B.M.T. and Bos, J.B. 1995. *Regulation of Ras-mediated signalling: more than one way to skin a cat*. *Trend. Biochem Sci* 20:18-22.
- Cárdenas García, M. 2000. *Search of therapeutic targets using a model of intracellular signals mediated by RAS*. D. Thesis, Facultad de Química e Instituto de Química, UNAM.
- Denhardt, D.T. 1996. *Signal-transducing protein phosphorylation cascades mediated by Ras/Rho proteins in the mammalian cell: the potential for multiplex signalling*. *Biochem. J.* 318:729-747.
- Feig, L. A. 1993. *The Many Roads That Lead to Ras*. *Science* 260:767-768
- Hesketh, R. 1994. *The oncogene handbook*. Academic Press.
- Kawasaki, H., Moriguchi, T., Matsuda, S., Li, H.Z., Nakamura, S., Shimohama, S., Kimura, J., Gotoh, Y., Nishida, E. 1996. *Ras-dependent and Ras-independent activation pathways for the stress-activated-protein-kinase cascade*. *Eur. J. Biochem* 241: 315-321.
- Khosravi-Far, R., Campbell, Sh., Rossman, K. L. And Der, Ch. J. 1998. *Increasing Complexity of Ras Signal Transduction: Involvement of Rho Family Proteins*. *Adv. Cancer Res.* 230:57-107.
- Kholodenko, B.N., Hoek, J. B., Westerhoff, H. V. and Brown, G. C. 1997. *Quantification of information transfer via cellular signal transduction pathways*. *FEBS.* 414: 430-434.
- Lacal, J. C., and McCormick F. eds. 1993. *The ras Superfamily of GTPases*. CRC Press.
- Lagúñez-Otero, J., Strauch, M. and Armas Alemán, M. 1996 *Towards Phenotypic Engineering, Folding and Design* 1, S4-S5.
- Lagúñez-Otero, J. 1998 *The Cell as an Expert System*. Dagstuhl Seminar Report 215, 60. ISSN 0940-1121.
- Lowy, D. R. 1993. *Funtion and Regulation of Ras*. *Ann. Rev. Biochem* 62: 851-91.
- Macara, I. G. 1991 *The Ras superfamily of molecular switches*. *Cell Signalling* 3 (3) 179-187.
- O'Neil, P. V., 1982. *Fundamental concepts of topology*. Gordon and Breach New York, 2nd edition.
- Paton, R.C. 1993. *Some computational models at the cellular level*. *BioSystems*, 29: 63-75.
- Somogyii, R. 1996. *Modeling the complexity of genetic networks: Understanding multigenic and pleiotropic regulation*. *Complexity* 45: 63-68.
- Wasylyk, Hagman and Gutierrez-Hartmann 1998. *Ets transcription factors: nuclear effectors of the Ras-MAP kinase signaling pathway*. *Trends. Biochem. Sci.* 26: 213-216.
- Yamamoto, T., Taya, Sh and Kaibuchi, K. 1999. *Ras induced transformation and signaling pathway*. *J. Biochem.* 126: 799-803.