

A Loop Conjecture for Metabolic Closure

D.A. Contreras^{1,‡}, U. Pereira^{1,2,‡}, V. Hernández^{2,‡}, B. Reynaert^{2,‡}, J.C. Letelier^{2,‡}

¹ Departamento de Física, Facultad de Ciencias, Universidad de Chile, Casilla 653, Santiago, Chile.

² Departamento de Biología, Facultad de Ciencias, Universidad de Chile, Casilla 653, Santiago, Chile.
letelier@uchile.cl

Abstract

Although in the last few decades a variety of theoretical tools have been developed to better understand living organisms, their impact on experimental research has been rather limited. A common element between these theories is the idea of metabolic closure, i.e., the systems that produce all their metabolites and catalysts. In spite of an increasing consensus on the relevance of closure, a formal and operative definition has remained elusive. In this paper we revisit RAF sets and chemical organization theory and show how these two theories overlap and could help bring forth real world results. We also state a theorem ensuring the presence of a cycle of interdependent catalysts for RAF sets and conjecture that these cycles give stability to the network. This conjecture is illustrated and supported by computer simulations. Unavoidably, our viewpoint introduces the notion of fluxes and thus a temporal dimension to the purely algebraic model of RAF sets. The results of this work show that the incorporation of closure, topological and dynamical tools altogether is a promising path for a deeper understanding of living systems.

Introduction

In the last thirty years there have been many efforts directed to develop theories to understand biological systems in terms of metabolic closure or, equivalently, systems that produce and maintain themselves. Two crucial models that definitively put metabolic closure at the very center of biological organization are: Autopoiesis, formulated by Maturana and Varela (Maturana and Varela, 1980), and Rosen's (M,R) Systems (Rosen, 1958). But these two theoretical studies and similar theories (like the Chemoton or Autocatalytic sets), although very clarifying in basic aspects, have not yet produced technical results that illuminate the daily life of bench biologists involved in experimental research.

In the past (Jaramillo et al., 2010) we have emphasized that a little known formalism called RAF sets (Hordijk and Steel, 2004) is a particularly suited technical tool to understand closure in general and autocatalytic sets in particular. Here we study the relation between RAF sets and the chemical organization theory (COT), which is a theory that

adds to the dynamical aspects by introducing the notion of metabolic fluxes to the purely algebraic vision of RAF sets, an idea deeply embedded in basic metabolic engineering. This is accomplished by expressing the kinetic behaviour of the components (molecules) in terms of a stoichiometric matrix, which then leads directly to the concepts of rates and fluxes, introducing the temporal dimension. This approach can be used to expand the original RAF sets theory, which we consider to be highly valuable for biology, but unfortunately too algebraic to be of use, in particular lacking a way to describe the time behaviour of the systems, which is of most importance in the direction of a more realistic biological context.

Here we will show how notions from chemical kinetics can be fused with RAF sets to search for closure in metabolic networks. Although the results presented here seems, initially, as a mere technicalities without theoretical relevance, they open new research paths as we adjoin highly theoretical notions (RAF set and the metabolic closure) with an accepted used tool to understand metabolism in steady state. In particular we show the logical relation between COT and RAF sets.

RAF sets and COT in a Nutshell

We now give a brief introduction to the work of Hordijk and Steel (2004), who came up with a formal framework to study a autocatalytic systems. Their main aim appears to have been to develop algorithms with which autocatalytic systems in Kaufmann's sense (1993) could be described and found computationally. They have produced a powerful approach that can be used to analyze a wide variety of systems. Their formalism is based on the following two important sets: X is the set of molecules involved in metabolism (i.e. metabolites, catalysts or external input material, termed *food set* in the formalism), and \mathcal{R} is the set of reactions that define the metabolic network. Each reaction r is represented as a tuple (A, B) , where $A, B \subset X$, $A \cap B = \emptyset$, A are the reactants and B the products of reaction r .

Further, to formalize the notion of catalysis, a specific set C (called the set of "catalyzations" by Hordijk and Steel)

[‡]All authors contributed equally

is introduced. Each catalyzation c is a tuple (x, r) , where $x \in X$ is the catalyst and $r \in \mathcal{R}$ is the reaction catalyzed by x . Additionally, the subset of molecules that are used but are not produced by metabolism is called food and denoted by F . Thus, a *catalytic reaction system* over a food source F is composed by a triplet $\mathcal{L} = (X, \mathcal{R}, C)$ which defines the universe of molecules (X), the reactions occurring among these molecules (\mathcal{R}) and the identity of the catalyst involved in each reaction (C). Note that this already provides, although at a very simple level, a way to refer to a system, and distinguish the inner and outer components and the transformations that the components undergo.

The following additional functions are defined: $\rho(r) = A$ and $\pi(r) = B$, which return the reactants and the products of any given reaction r , respectively, and the function $\text{supp}(r) = \rho(r) \cup \pi(r)$. With the help of these elementary functions, the same notion can be extended to a *set of reactions* \mathcal{R}' as $\rho(\mathcal{R}') = \bigcup_{r \in \mathcal{R}'} \rho(r)$, where $\mathcal{R}' \subseteq \mathcal{R}$. This definition captures the conglomerate of molecules that participate as reactants for a set of reactions. A similar definition holds for $\pi(\mathcal{R}')$, the products of a subset of reactions. With these ideas we can define the closure of a subset $X' \subseteq X$ relative to $\mathcal{R}' \subseteq \mathcal{R}$ ($\text{cl}_{\mathcal{R}'}(X')$) as the set of reachable molecules that can be synthesized by starting from X' and iteratively applying all the reactions in \mathcal{R}' . Note that this definition is of most importance, as it follows that a set of molecules which is closed (i.e. it is equal to its closure) under a set of reactions will not generate any new molecule and thus, conserves its identity. This operation captures the central idea of metabolic closure, which is fundamental for achieving organizational invariance in autopoietic systems. A *catalytic reaction system* is *reflexively autocatalytic* if for each $r \in \mathcal{R}$ there is an $x \in \text{supp}(\mathcal{R})$ such that $(x, r) \in C$. In other words, every catalyst must be a reactant or product of a reaction in the same system. The system is F -generated if every reactant is either produced by the system or incorporated as a food item (i.e. formally $\rho(\mathcal{R}) \subseteq F \cup \pi(\mathcal{R})$). A system that is *reflexively autocatalytic* and F -generated is called a RAF set (see figure 1).

RAF sets can be understood informally as an interdependent set of biochemical reactions where all of the metabolites, with the exception of the so-called food set, are produced by the collection of reactions \mathcal{R} . This self generation, a defining feature of autopoietic and (M,R) systems, is the core of metabolic closure. Thus, RAF sets, autopoietic and (M,R) systems overlap to a great extent; positioning RAF sets as an operative theory to metabolic closure. The advantage of RAF set formalism is that it is precise enough to be coded in well defined algorithms that exploit its intrinsic recursiveness. To check if a given collection of biochemical reactions is indeed a RAF set, Hordijk and Steel (2004) developed algorithms aimed to analyze the interdependence between a given catalyst and its production pathway.

The chemical organization theory, initially developed by

Dittrich and Di Fenizio (2007), deals with chemical reaction networks. In what is called *static analysis*, the part of this theory that is concerned with the topology of the system, molecules and reactions are defined in a very similar way as in RAF sets. Most notably, both theories share the definition of the closure operator. But while COT makes no explicit mention to catalysts and therefore distances itself from biological systems in which this concept is fundamental, it does incorporate tools to study the dynamical behaviour of chemical reaction networks, thus provides a connection between the structure of a system and the dynamical aspects of it. This is accomplished by first expressing the system in terms of the stoichiometric matrix and associated differential equations.

In COT, it is useful to recognize systems fulfilling certain properties, such as closure. For example, a system is self-sustained if it is able to generate every molecule that is used up. When this topological consideration is transported to the time domain, we can define mass-maintaining systems. A system is said to be mass-maintaining when:

1. All reactions that can be fired by the molecules in the system occur at some positive rate
2. Reactions whose reactants are missing from the system do not occur
3. There is a combination of reaction rates such that all molecules increase or maintain their concentration.

A system which is both closed and mass-maintaining is called an organization. Organizations are interesting as they resemble very closely autopoietic systems. Also, organizations are at the center of many theorems in COT. This theory and RAF sets deal with closure. While one makes no distinction between catalysts and metabolites, the other one lacks the notion of time, which are essential elements of living systems. In the next paragraph we will show an relation between these two theories.

Kinetics in RAF sets

If a theory is to have impact on real biochemical world, it must deal with the notions of that domain, thus, to gain a full understanding of closure we must complement the purely algebraic nature of RAF sets with ideas taken from Metabolic Control Analysis (MCA), a field generated to understand and measure fluxes in biochemical systems which is of common use in the field of metabolic engineering.

Current MCA is a quantitative theory that does not consider closure, as catalysts (i.e. enzymes) are placed in the network, but the reactions generating them are not taken into account. By putting the quantitative aspects of MCA and applying them to RAF sets, side by side, we can gain insight in how to study closure quantitatively. All the theories of metabolic closure (Autopoiesis, (M,R) systems, Autocatalytic sets, etc) are essentially algebraic or conceptual

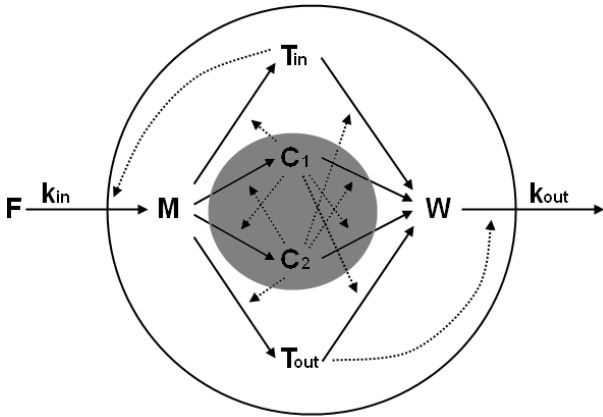


Figure 1: A simple example of a RAF set. Food elements F are incorporated into the system and generate metabolites M , which are transformed into two different catalysts sets: a) T_{in} which regulates the inflow of F and C_2 which catalyzes its generation and b) T_{out} which regulates the outflow of waste metabolites W and C_1 who catalyzes its formation. In addition, C_1 and C_2 catalyze the formation and destruction of the transporter catalysts (T_{out} and T_{in} , respectively), and also they mediate the generation and consumption of each other, forming the Reflexive Autocatalytic core of the system. Finally, growth is regulated by modulating the inflow of F and the outflow of W . We want to highlight the loop defined by metabolites M which turn into C_2 who regulates the formation of C_1 starting from M , a reaction regulated by C_2 .

models centered on connectivity but not in dynamics. To go further in our understanding we must include the time course evolution of the concentrations inside the system.

Fortunately, the formalism of Reder (1988) that uses the stoichiometric matrix and the matrix $D_x v$ to study rates, can be applied almost *verbatim* to analyze if a RAF set will grow or disappear. The great advantage of applying MCA formalisms is that we can quantitatively study how a system with metabolic closure can grow or disappear.

RAF sets are sets of coupled biochemical reactions with the attribute that the catalytic dependences between reactions and their catalysts are explicitly given. As said, RAF sets demand that almost all the molecules that conform a system can eventually be generated, directly or indirectly, from certain food materials and that all catalysts are produced by the system.

The *transformation* part of a RAF set can be represented by the formalism of the stoichiometric matrix, a well known tool extensively used in fields like MCA and Systems Biology in which every reaction is written as a column and every metabolite is referred to as a row. For example, the matrix N of the system described in figure 1 would be as following:

$$N = \begin{pmatrix} 1 & -1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & -1 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

By using the column representation of reactions, it is convenient to define the addition of reactions as a standard addition operation of vectors. This operation expresses the occurrence of both reactions as a single net reaction.

Note that the *catalytic* part lies outside the stoichiometric matrix and cannot be deduced from it. But, in an idea that can be traced back at least to Reder (1988), the *catalytic* part can be represented by a matrix $D_x v$ (also known as the Jacobian of the system) that contains all partial derivatives relating every reaction with every metabolite (or catalyst) in the system. Thus, the catalysts for a given reaction can be discovered by ranking the partial derivatives of the rate of this reaction respective to all metabolites (molecules) in the system. For example, the Jacobian matrix $D_x v$ of the system described in figure 1 would be:

$$D_x v = \begin{pmatrix} \partial_M v_1 & 0 & 0 & 0 & \partial_{T_{in}} v_1 & 0 \\ \partial_M v_2 & 0 & \partial_{C_1} v_2 & \partial_{C_2} v_2 & 0 & 0 \\ \partial_M v_3 & 0 & \partial_{C_1} v_3 & \partial_{C_2} v_3 & 0 & 0 \\ \partial_M v_4 & 0 & \partial_{C_1} v_4 & 0 & \partial_{T_{in}} v_4 & 0 \\ \partial_M v_5 & 0 & 0 & \partial_{C_2} v_5 & 0 & \partial_{T_{out}} v_5 \\ 0 & \partial_W v_6 & \partial_{C_1} v_6 & \partial_{C_2} v_6 & 0 & 0 \\ 0 & \partial_W v_7 & \partial_{C_1} v_7 & \partial_{C_2} v_7 & 0 & 0 \\ 0 & \partial_W v_8 & 0 & \partial_{C_2} v_8 & \partial_{T_{in}} v_8 & 0 \\ 0 & \partial_W v_9 & 0 & \partial_{C_2} v_9 & 0 & 0 \end{pmatrix}$$

Every RAF set can then be described by two matrices; N , that shows the network connectivity, and $D_x v$, that quantifies catalizations.

The necessity of using the $D_x v$ matrix to analyze RAF sets lies in the fact that autocatalysis is a phenomenon that does not depend *only* on connectivity. As it has been shown recently by Plasson et al. (2010) and Piedrafita et al. (2010), the stability of an autocatalytic set depends on the relative rate of some reactions. Thus, two systems with identical connectivities but with different kinetics for some reactions can have vastly different behaviors.

As stated above, another theory concerned with formalizing biological organization is COT, a theoretical framework also centered in the idea of closure differing from RAF as the idea of catalyzation, perhaps the hallmark of RAF sets, is not considered. On the other hand COT brings an idea, the importance of fluxes in a network, that are not considered in RAF sets which is a purely algebraic approach to the description of biological organization. Thus an important question arises: can these two models be related? Can they support each other, in the sense of across fields fertilization? In the next section we clarify some relations between these two models.

An observation needed at the very beginning is that analysis using RAF sets and COT belong to two very different viewpoints as crucial elements in one theory are totally absent in the other. Thus as organizations (in the sense of COT) require that the overall flux across a relevant subset of reactions is maintained (thus avoiding the disappearance of crucial metabolites that, if absent, will produce network collapse). A mirror like situation can be stated with respect catalyzations, a cornerstone idea in RAF sets, and (surprisingly) an idea that is absent from COT. Thus we should expect that if a system is a RAF set it is not immediate that it is also an organization. Only in some especial conditions we should be able to find how these ideas can be concurrently applied.

A hidden relation between RAF sets and COT

A further dissection of RAF sets shows that, although fluxes and reaction rates initially seem to be absent from this model, kinetic ideas do exist just below the surface. In effect we propose two lemmas and a theorem that will bring new light to the problem of comparing both approaches:

Lemma 1 *If a catalytic reaction system $\mathcal{L} = (X, \mathcal{R}, C)$ is F -generated, then for all metabolites x (including catalysts) produced by any reaction $r \in \mathcal{R}$, $x \in \text{supp}(\mathcal{R})$ there is a positive linear combination of reactions $\bar{r}_x = \sum_i \alpha_i r_i$ such that the metabolite x belongs to the products of the reaction \bar{r}_x , $x \in \pi(r_0)$ and the reactants of \bar{r}_x belong to the Food set, i. e., $\rho(r_0) \subseteq F$.*

Proof: Considering the algorithm used to find the closure of \mathcal{L} (Hordijk and Steel, 2004), let $W = F$. Then add

the products of reactions $\mathcal{R}_0 = \{r \in \mathcal{R} | \rho(r) \subseteq W\}$ to W . Adding all reactions in \mathcal{R}_0 gives a global net reaction \bar{r}_0 that consumes metabolites from the Food set F only and produces each metabolite in W . If this process is repeated, considering $W = F \cup \pi(\mathcal{R}_0)$, it is possible to build the set \mathcal{R}_1 that contains all reactions that have their reactants in W , but excluding the reactions from \mathcal{R}_0 . Adding all reactions in \mathcal{R}_1 we obtain a new reaction \bar{r}'_1 that requires metabolites from W only and produces any metabolite in $\pi(\mathcal{R}_1)$. To obtain the fact that this last reaction uses only metabolites from the Food source, let $\bar{r}_1 = \alpha \bar{r}_0 + \bar{r}'_1$, where α is the most negative stoichiometric coefficient of the reaction \bar{r}'_1 . This procedure takes enough metabolites from F to generate $\pi(\mathcal{R}_1)$. If we repeat this algorithm until it is not possible to find new metabolites, we will have generated $cl_{\mathcal{R}}(F) = W$. If the system is F -generated, according to Hordijk and Steel (2004), we have that $cl_{\mathcal{R}}(F) = F \cup \text{supp}(\mathcal{R})$. We have shown that for every metabolite $x \in cl_{\mathcal{R}}(F)$ a composite reaction exists which generates it consuming food items only, in fact it is one of the r_i .

Lemma 2 *If a catalytic reaction system $\mathcal{L} = (X, \mathcal{R}, C)$ is F -generated, there is a strict positive linear combination of reactions $\hat{r} = \sum_i \alpha_i r_i$ with $\alpha_i > 0$ such that all metabolites are products of this reaction, i. e., \hat{r} is a strictly positive vector.*

Proof: From Lemma 1 it follows that for each metabolite m_j there is a positive linear combination of reactions $\bar{r}_{m_j} = \sum_i \alpha_{i,j} r_i$ such that this metabolite is produced exclusively from the Food set. This linear combination \bar{r}_{m_j} is the resultant net reaction associated with the path of reactions j that generate each metabolite m_j . If for all metabolites we add their generating reactions $\bar{r}_{sum} = \sum_i \alpha_i r_i$ with $\alpha_i = \sum_j \alpha_{i,j}$, we have from the Lemma 1 that \bar{r}_{sum} is strictly positive. We must note that not all reactions will be used. We refer to these reactions generally as r_s , having $\alpha_s = 0$ in r_{sum} for those reactions. If we consider the sum of this reactions $r_{not} = \sum_s r_s$, it will consume metabolites. To maintain \hat{r} positive and still fire these non-essential reaction (this will be needed later), we set $\hat{r} = \beta r_{sum} + r_{not}$ with β sufficiently large. Then we can construct a strictly positive linear combination $\hat{r} = \beta r_{sum} + r_{not} = \sum_i \bar{\alpha}_i r_i$ with:

$$\bar{\alpha}_i = \begin{cases} 1 & \text{if } r_i \in \{r_s\} \\ \beta \alpha_i & \text{if } r_i \notin \{r_s\} \end{cases} \quad \text{note } \bar{\alpha}_i > 0$$

and all the metabolites are product of this reaction r_0 .

These lemmas, framed completely in the language of RAF sets, could be interpreted as mere technical results about RAF sets. In essence they state that every metabolite can be generated from the food set and makes explicit the overall reaction producing each, non-food, metabolite. But every time we use a stoichiometric matrix N we are implying a given kinetics because of the necessary equation

relating N to the change of concentrations: $N \cdot v = dX/dt$, where v is the vector of rates. Thus the requirement, in COT, that ($dX/dt \geq 0$) can be phrased as a condition on the components of v . These lemmas show how some (not all) Organizations could be RAF sets, and it is interestingly that they are proved by using notions of linear algebra. Also note that the positive linear combination predicted by the lemmas explicitly shows how to combine individual reactions in any RAF set to attain mass-maintenance.

Once we have established this link we can a little bit further and search for deeper connections. The next theorem continues to exploit matrix N to sketch how some RA sets are F-generated using the stoichiometric matrix N .

Theorem 1 *If a catalytic reaction system is F-generated, then there is a strictly positive rate vector v , such that $N \cdot v = dX/dt$ is also strictly positive, where N is the stoichiometric matrix of the system.*

Proof: We note that the operation $N \cdot v = dX/dt$ is equivalently mathematical to make a linear combination of reactions $r = \sum_j \alpha_j r_j$ if we consider each reaction as a column and α_j as the velocity of reaction r_j . From Lemma 2, if we take the reaction r_0 and choose $v_j = \bar{\alpha}_j$ (normalizing time units), therefore a v exists with components $v_j > 0$ associated to a flux vector dX/dt that satisfies $dX_i/dt > 0$, equivalent to the column representation of r_0 with all of their components also positive.

Corollary 1 *If a catalytic reaction system is F-generated, then it is also an organization.*

Proof: An F-generated system is, by definition, *closed* and as theorem 2 shows, it also satisfies the property of *mass-maintenance*. Thus, it is an organization.

This theorem explains the existing relations between Organizations, F-generated sets, RA sets and RAF (see figure 2). Essentially, we have proved that all F-generated sets are *organizations* and a subset of them are also *reflexive autocatalytic*. This subset is the RAF sets. Theorem 1 is a simple one that has the virtue of illuminating how these two theoretical frameworks are related to each other.

This result is important because some new technical theorems have been obtained by Dittrich's group, for example, on how to detect organizations among real metabolic networks (Centler et al., 2010, 2008). Thus, our theorem shows that these new tools, developed to find organizations, could be also used to search for RAF sets.

In addition, we will make a definition to the sets that are organizations and RA at the same time.

Definition 1 *If a catalytic reaction system is Reflexive Autocatalytic and an Organization, then it is a Reflexive Autocatalytic Organization, RAO.*

These sets are reaction systems that can be sustained, but not necessarily can be generated from a food set F exclusively. We have shown that all F-generated sets are orga-

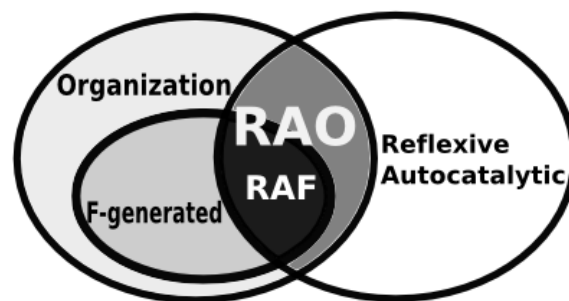


Figure 2: Venn's diagram depicting the logical relations between RA, RAF, and F-generated sets and *organizations* under COT's definition. All RAF sets are organizations, but whether all organizations are F-generated is an ambiguous matter.

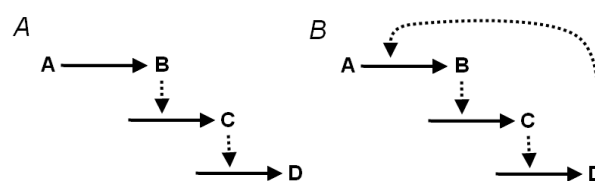


Figure 3: A: A chain of dependent catalyzations. B: A catalyzation loop

nizations, but the converse result (all organizations are F-generated sets) is more difficult to handle. We propose two different approaches: First, if one decides that the Food set corresponds only to the molecules generated from the empty set (in COT's phrasing of reactions), then it is clear that there are organizations which are not F-generated. On the other hand, for any organization it is always possible (due to the closure property) to find a suitable set (generally not unique) F such that the corresponding F-generated set is equal to the given organization. Thus, the extent to which organizations and F-generated sets overlap depends on which approach one takes to express COT systems in terms of RAF sets.

The Loop Theorem

As most theories on biological organization are centered in the notion of closure (Letelier et al., 2011), RAF sets formalism give a succinct and useful description of closure. First, we shall consider a *chain* of catalyzations in which a product from one reaction catalyzes another reaction in the chain (figure 3A). If eventually a product catalyzes an earlier step (figure 3B), we have a catalyzation loop. As we shall see, in a RAO it is always possible to find such a catalyzation loop if the catalysts are not part of the food set. This condition seems natural for systems with metabolic closure.

Considering this definition we propose:

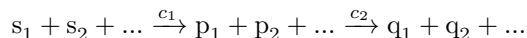
Theorem 2 *If all catalysts in a RAO are generated by the system, then there is at least one catalyzation loop.*

Proof: In such a RAO every catalyst must be generated by a reaction, which in turn must be catalyzed too. In this sense, the production of every catalyst is directly dependent on another and indirectly dependent on a sequence of catalysts. The number of catalysts is finite therefore, at some point a catalyst must depend indirectly on itself, thus, forming at least one catalyzation loop. Note that not every catalyst is part of a loop as it is allowed that some catalysts may catalyze reactions which yield no catalysts as products, yet the system as a whole must have at least one catalyzation loop. Note that in case of direct autocatalysis, the loop is trivial. Also, whether there is more than one catalyzation loop is a question that must be addressed in each particular case.

An unsuspected consequence of the loop theorem is that some of the catalysts inside the catalyzation loop must have a dual catalytic role, that is enzymes that catalyze at least the creation of other two enzymes, if not happen the trivial case of all enzymes catalyze the creation of another one enzyme. This is interesting, as one modern re-interpretation of Rosen's results about how living systems avoid infinite regress is by having enzymes with dual functions (Letelier et al., 2006). Thus, this systemic result (i.e. existence of moonlighting enzymes) can be achieved by two different methods.

This theorem is a basic result that follows directly from the basic definitions of RAOs, but it shows an important property that needs to be underlined: the catalyzation loops (one or more) inside a RAO may be considered as its autocatalytic core and, functionally, there is a difference between the catalysts of the loop and the ones outside it.

We conjecture that the functional segregation hinted has important consequences. In effect, to confer stability to the core the catalysts outside it control the inflow and outflow of matter to and from the core. Thus, the net flow of matter inside it must be controlled, as a large flow would generate an exponential runaway and a small one would extinguish some core components, destroying its organization. Keeping this balance between in and outflow will be seen as homeostatic regulation. In summary, we conjecture that the catalyzation loop confers long term stability to the network. The analytical proof of this result seems difficult, but we did computer simulations in small (toy-like) systems and using mass-action kinetics, expressed for reactions:



By the formula:

$$\frac{d[p_k]}{dt} = k_1[c_1] \prod_i [s_i] - k_2[c_2] \prod_j [p_j]$$

Figure 4 shows one example for the temporal evolution of the concentrations of molecules for the RAF system of

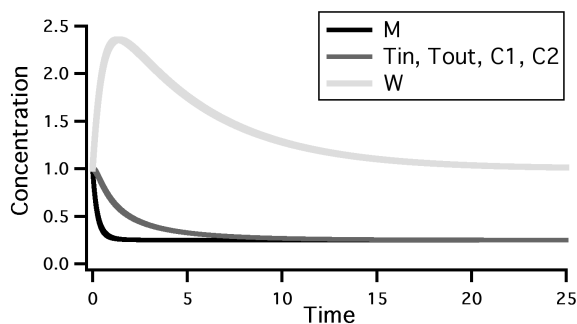


Figure 4: Temporal evolution of RAF toy system (see figure 1). The system reaches a steady state in which all concentrations are different from 0.

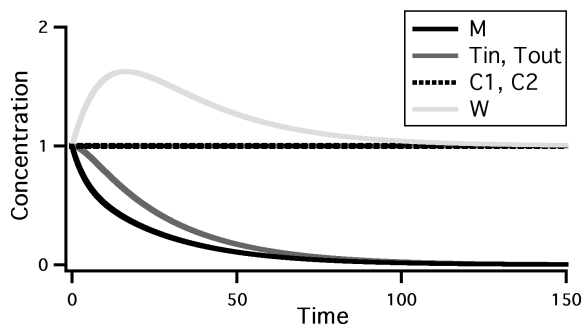


Figure 5: Temporal evolution of a non-RAF system. Although the concentrations of the catalysts C_1 and C_2 were fixed, the system decays until its components disappear.

figure 1. We can see that a steady state is achieved. In figure 5 we simulated a similar system but without the reactions that generate or destroy the catalysts C_1 or C_2 , removing the catalyzation loop and making the system a non-RAF set. In this last case the concentrations of many components decay to zero, stopping the network dynamics.

We also did the bifurcation analysis by varying rate constants k_{in} and k_{out} corresponding to the reactions $r_1 : \phi \rightarrow M$ and $r_9 : W \rightarrow \phi$ respectively. For the RAF set, almost every combination of parameters k_{in} and k_{out} leads to a steady state, except at the border where $k_{in} = 0$ or $k_{out} = 0$ and some regions close to these.

On the other hand, the non-RAF set has no stable points for most values of k_{in} and k_{out} . This puts in evidence the relevance of the autocatalytic core, so that the growth of one part of the system encompasses the rest and grows harmonically and coherently between the inflow and outflow of matter.

Growth and Homeostasis in Autopoietic Systems

Any increase in the concentration of a loop catalyst will translate into an increased concentration of every other cata-

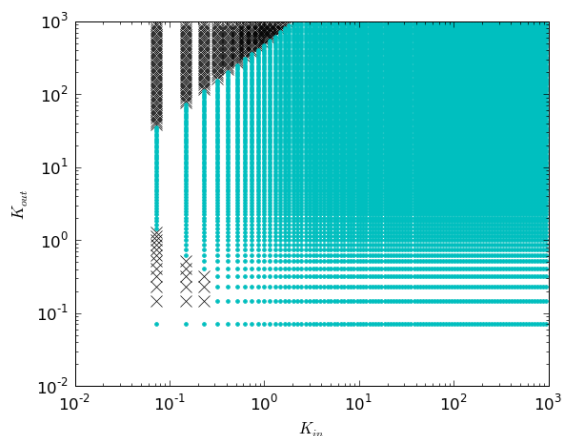


Figure 6: Bifurcation diagram of the RAF set in figure 1, for variables k_{in} and k_{out} . A similar diagram for a non-RAF system has no stable region. (cyan dots=stable, black Xs=unstable).

lyst in the loop, which would consequently lead to a further increment of the first one, exhibiting an apparent autocatalytic behavior. At the same time, this loop must be connected to side branches that lead to the production of catalysts that are not directly related to autocatalysis, but with the obtention and processing of the food sources that sustain it. An interesting type of branches are the ones that lead to the regulation of the enzymes that control the in and outfluxes, because they are supposed to regulate the whole metabolisms growth rate by coordinating these fluxes. This fact shows us the importance of the topology, because every enzyme must be whether part of a loop or a branch of it, thus, a change in an enzyme concentration which is part of the loop will have repercussions in the whole systems growth, as it also affects the enzymes that regulate the fluxes. Thus, the RAF sets may help to understand the dynamics of the homeostatic process. This is not mutually exclusive with the fact that an increase in the concentration of an enzyme outside the loop may have a direct repercussion on global growth.

The loop theorem has an important application for autopoietic systems, that can be defined as self-encapsulated RAF sets. For an autopoietic system (that according to the above theorem must contain at least one autocatalytic loop) to be stable in time, there must be a fine balance between the generation and destruction of molecules. But there must also be a balance respective to control its volume in order to keep the concentrations unaltered. Thus, the organization (*à la* Autopoiesis) of a RAF set must be under a precise homeostatic control, as growth must be promoted, but in the context of compensating for the volume increase without suffering the consequences of autocatalytic growth. Thus, in a first

approach, we must allow for a system to grow in terms of the net amount of molecules, but not in concentration. This implies that volume must be under active control and that allowing the system to grow would not be a contradiction with homeostatic principles.

Discussion and Conclusion

As we have previously stated (Jaramillo et al., 2010) we conclude again that RAF sets formalism is particularly suited to study closure. Of course many aspects of metabolic closure escape this theory (the operator of organizational invariance β of (M,R) systems is a prime example), but this framework provides a solid starting point. The loop theorem proved here, which is a property shared by RAOs, Autopoietic and (M,R) systems is a good example of its power.

Another important point of the present study is to apply the analysis of COT to RAF sets. As it is usual in theoretical biology, the different frameworks generated to explain living organization exist in closed universes without dialogue between competing theories. Here, we partially break this isolation by showing how organizations *à la* COT contain all RAF sets, but not all RA sets. This inclusion, although obvious and expected from a theoretical viewpoint, is not easy to prove. We have developed demonstrations using arguments from linear algebra, instead of the set theory arguments favored in RAF. The most unexpected result is the uncovering of chemical kinetics arguments in RAF sets. In effect, RAF sets appear to be a purely algebraic entity, without considerations for time nor kinetics; but as soon as their stoichiometric matrix is expressed, the kinetic arguments of COT are made obvious. Thus our lemmas and theorems show deep relations between the pure algebraic formulation of RAF sets with the dynamics of organizations in COT. Perhaps this same reasoning could be also be applied to (M,R) systems. Taken together, the results shown here show the value of putting all the different notions of metabolic closure under a common analytical umbrella.

COT has already produced an interesting number of results on the dynamics of reaction networks, in particular regarding to the long-term temporal behaviour and stability of these systems (Dittrich and Di Fenizio, 2007). An interesting result from this theory, which complements the loop theorem presented here, is the decomposition theorem for organizations (Veloz et al., 2011). This theorem states that, under certain conditions, it is possible to split a system into subsystems whose dynamic behaviour are weakly coupled. Thus, an open question is to investigate how our loop theorem, which seems to indicate that systems cannot be segmented, is compatible with such uncoupling of subsystems. In effect a catalyzation loop might constitute the minimum decomposable unit.

We presented the conjecture that systems with at least one catalyzation loop are more stable than similar systems without such loops. This is a powerful result that will unavoid-

ably demand tools from MCA, the most elaborate theory about fluxes in biochemical networks, to be proven or refuted.

In summary, our efforts show that closure is a conceptual key to understand biological organization, as an example we have come close to use closure as an argument to prove one theorem (loop theorem), which we believe is a valuable conceptual step and a fertile direction for theoretical biology.

Acknowledgements

We would like to thank the reviewers for their comments. They were very helpful to improve this manuscript. Funded by FONDECYT grant 1110247, CONICYT scholarship 22101236 and 22110804.

In memoriam

This work reflects the school of thinking generated by Francisco Varela in Chile during the 70's. Francisco was a valued friend, a guide and an inspiration for new generations of biologists.

References

- Centler, F., Kaleta, C., Speroni di Fenizio, P., and Dittrich, P. (2008). Computing chemical organizations in biological networks. *Bioinformatics*, 24(14):1611 – 1618.
- Centler, F., Kaleta, C., Speroni di Fenizio, P., and Dittrich, P. (2010). A parallel algorithm to compute chemical organizations in biological networks. *Bioinformatics*, 26(14):1788 – 11789.
- Dittrich, P. and Di Fenizio, P. S. (2007). Chemical organization theory. *Bulletin of Mathematical Biology*, 69:1199–1231.
- Hordijk, W. and Steel, M. (2004). Detecting autocatalytic, self-sustaining sets in chemical reaction systems. *Journal of Theoretical Biology*, 227:451–461.
- Jaramillo, S., Honorato-Zimmer, R., Pereira, U., Contreras, D., Reynaert, B., Hernández, V., Soto-Andrade, J., Cárdenas, M., Cornish-Bowden, A., and Letelier, J. (2010). (M,R) Systems and RAF sets: Common ideas, tools and projections. *Proceedings of the Twelfth International Conference on the Synthesis and Simulation of Living Systems*, pages 94 – 100.
- Kauffman, S. (1993). *The Origins of Order*. Oxford University Press, New York.
- Letelier, J. C., Cárdenas, M. L., and Cornish-Bowden, A. (2011). From *L'Homme Machine* to metabolic closure: steps towards understanding life. *Journal of Theoretical Biology (in print)*.
- Letelier, J. C., Soto-Andrade, J., Guíñez Abarzúa, F., Cornish-Bowden, A., and Cárdenas, M. L. (2006). Organizational invariance and metabolic closure: analysis in terms of (M,R) systems. *Journal of Theoretical Biology*, 238:949–961.
- Maturana, H. and Varela, F. (1980). *Autopoiesis and Cognition: the Realisation of the Living*. D. Reidel Publishing Company, Dordrecht, The Netherlands.
- Piedrafitá, G., Montero, F., Morán, F., Cárdenas, M. L., and Cornish-Bowden, A. (2010). A simple self-maintaining metabolic system: Robustness, autocatalysis, bistability. *PLoS Comput Biol*, 6(8):e1000872.
- Plasson, R., Brandenburg, A., Jullien, L., and Bersini, H. (2010). Autocatalyses. *Proceedings of the Twelfth International Conference on the Synthesis and Simulation of Living Systems*, pages 4 – 11.
- Reder, C. (1988). Metabolic control theory: A structural approach. *Journal of Theoretical Biology*, 135(2):175 – 201.
- Rosen, R. (1958). The representation of biological systems from the standpoint of the theory of categories. *Bulletin of Mathematical Biology*, 20(4):317–341.
- Veloz, T., Reynaert, B., Rojas, D., and Dittrich, P. (2011). A decomposition theorem in chemical organizations. *Proceedings of the Eleventh European Conference on the Synthesis and Simulation of Living Systems (this volume)*.