A decomposition theorem in chemical organizations

Tomas Veloz[1], Bryan Reynaert[2], David Rojas-Camaggi[2] and Peter Dittrich[3]

[1]Department of Psychology, University of British Columbia
[2]Department of Biology, Universidad de Chile
[3]Friedrich-Schiller-University Jena, Institute of Computer Science, Bio Systems Analysis Group, D-07743 Jena
tomas.veloz@ubc.ca

Abstract

The Chemical Organization Theory (COT) is an abstract reaction network model that has a deep connection to autopoiesis as they share the same central topic: Organization. The main characteristic of autopoietic systems is that they preserve their own organization; this constitutes their identity. In terms of COT, organizations are special reaction networks which are closed and self-maintaining. Organizations compose the majority of stable behaviours of a reaction network (Peter and Dittrich, 2011), in particular every fixed point can be mapped to an organization (Dittrich and Di Fenizio, 2007). Obtaining the set of organizations of a network is a central objective in COT, but it is usually a complex computational task. This work intends to reveal the underlying mathematical structure of organizations. We state a theorem of decomposition for organizations to understand the difficulties of verifying if a set of molecular species is an organization. This suggests a step towards the development of more efficient algorithms and the classification of reaction networks in terms of how complex it is to obtain its set of organizations. We also discuss the consequences of this theorem in relation to autopoietic systems.

Introduction

During a 30-years period, from the 1950's to the 1980's, the field of biological systems and their generalized properties saw the birth of multiple theories (Eigen and Schuster, 1977; Kauffman, 1969; Maturana and Varela, 1973; Wiener, 1948; von Bertalanffy, 1968; Rosen, 1958). A wealth of formalisms were laid out, which focused on different perspectives on the fundamental properties of living systems, but as it was to be expected, there have been deep similarities between most of these theories (Hordijk and Steel, 2004; Jaramillo et al., 2010; Letelier et al., 2003).

Since their conception, most of these theories have been consigned to the theoretical domain having little incidence in applied sciences, with the possible exception of what is currently known as *systems biology*.

This situation may be because the process of translation between the language employed in these theories and the language commonly used in biology is not trivial (Cornish-Bowden et al., 2007). The chemical organization theory, inspired by Fontana and Buss (1994), provides an interesting departure from this tendency as it provides a language which is not only clear and well-defined but also correlates directly to the biomolecular domain. Due to its mathematical foundations theorems can be formally proven and developed (Benkö et al., 2009; Peter and Dittrich, 2011; Peter et al., 2010). Also, COT is a powerful tool to analyze the asymptotic behaviour of reaction networks that other analytic or simulation methods cannot cope with. In particular, the chemical organization theory has been applied to biochemical domains (Centler et al., 2008b; Kaleta et al., 2006; Matsumaru et al., 2006), atmospheric photochemistries (Centler and Dittrich, 2007), and as a tool for the study of P-systems (Peter et al., 2010). It also has been proposed as a theoretical framework to design chemical computers (Matsumaru et al., 2007), and recently, for the study of social systems (Dittrich and Winter, 2008).

Thus, COT is very well suited to study autopoietic systems as both theories focus on the problem of self-maintaining organizations. At first it may seem inappropiate that a theory developed around artificial chemistries may be used to study autopoietic systems, but it should be noted that autopoietic systems are not obliged to a molecular structure or realization, that just happens to be the case of living organisms. Furthermore, the "protobio" (Varela et al., 1974) was both an early attempt to simulate autopoietic systems and an artificial chemistry. Therefore, any advance in COT might be transported directly to the theory of autopoietic systems, independent of the domain in which they are actually realized.

In this paper, we first introduce COT and its relation to autopoietic systems. Then, we present a decomposition theorem from COT and finally analyze its consequences for the long-term time behavior of biological systems.

Autopoiesis and Chemical Organization Theory

Autopoiesis was developed as a theory for living systems by Maturana and Varela (1973). The central idea is that a living organism is a machine, constituted as a unit in space, which maintains its organization through its operation. Moreover, a living organism performs a set of pro-

cesses which generate the components necessary to realize these processes. Thus, the notion of organization as a network of interacting components which stably maintains itself in time is of most importance in this theory. Hence, a theory which concerns itself with such a concept may relate closely to autopoiesis.

The COT, which was introduced in Dittrich and Di Fenizio (2007) in the context of algebraic chemistries, is a mathemathical theory, that by using the structures of sets and matrices, is able to formalize chemical reaction systems at a topological and dynamical level. In this theory, an organization is a reaction network which has the potential of being self-maintaining and thus matches very closely to the definition given by Maturana and Varela. Moreover, as "an autopoietic system is an homeostatic machine which has its organization as the variable it maintains constant", organizations must be stable in time. The COT explores these considerations and has already had important results in this regard. In particular, in this work we present a decomposition theorem for organizations. In order to present our main result, we must first introduce the basics of COT.

Chemical Organization Theory

Basic Definitions

At the most basic level of this theory, we deal with two types of objects: molecular species (from now on species) and reactions. The species are the elements of a species set $\mathcal{M} = \{m_1, \ldots, m_n\}$, and each reaction R is modeled by a pair $R = (A, B) \in \mathcal{P}_M(\mathcal{M}) \times \mathcal{P}_M(\mathcal{M})$, where $\mathcal{P}_M(\mathcal{M})$ denotes the set of all the multisets formed by elements in $\mathcal{M}$. A multiset is defined by a pair (X, η_X), where X is a set and the function $\eta_X : X \to \mathbb{N}_0$ states the number of ocurrences $\eta_X(x)$ (multiplicity) of x in the multiset. In order to be consistent with the usual notation of chemical reactions, we will write the multiset (X, η_X) by $\sum_{x \in X} \eta_X(x)x$. Moreover, we will refer to the reaction $R = (A, B)$ by $R = A \to B$, where $A = (\mathcal{M}, \eta_A)$ and $B = (\mathcal{M}, \eta_B)$.

From now on, let $\mathcal{R} = \{R_1, ..., R_k\}$, where $R_i = A_i \to B_i$, with $A_i = a^{i1}m_1 + \cdots a^{in}m_n$ and $B_i = b^{i1}m_1 + \cdots b^{in}m_n$, for $i = 1, ..., k$ and $j = 1, ..., n$. a^{ij} corresponds to the stoichiometric coefficient of m_j in reaction R_i, that is, the multiplicity $\eta_{A_i}(m_j)$ of molecule m_j in A_i; b^{ij} is defined in a similar way. Now we can define an Algebraic Chemistry, which captures the notion of system, as follows:

Definition 1 *An Algebraic Chemistry is a pair* $\langle \mathcal{M}, \mathcal{R} \rangle$.

A species $m \in \mathcal{M}$ is said to be *present* in a multiset $(X, \eta_X) \in \mathcal{P}_M(\mathcal{M})$ if and only if its multiplicity $\eta_X(m)$, is greater than zero. The *reactants* and *products* of a reaction $R = A \to B$ are the species present in A and in B respectively. The reaction R *can be fired* by a set $X \subseteq \mathcal{M}$ if and only if all species present in A are in X.

From now on let $X \subseteq \mathcal{M}$. Note that there exists a maximal set of reactions $\mathcal{R}_X \subseteq \mathcal{R}$ which can be fired by X. R_X is composed by the reactions $R_i = A_i \to B_i$ such that, if m is present in A_i, then $m \in X$. We call $\mathcal{R}_X$ the *possible reactions* set of X.

In order to deal with the dynamical aspects of any system, it is desirable that the system maintains its identity. This leads to the question of whether the system, left to react for an arbitrary amount of time, will generate species which where originally absent. Note that in a general chemical setting, in which no species will be used up completely, all the reactions that can be fired will fire at some positive rate; therefore, it suffices to check if the set of possible reactions for the system produces any novel species. If it does not, we say that the set of species is closed. The following definition states this formally:

Definition 2 *We say X is closed if and only if for all $R = A \to B \in \mathcal{R}_X$, m is present in B implies $m \in X$. Let $G_{CL}(X)$ be the closure of X, then it is the smallest closed set containing X.*

Remark The closure of a set has been proven to be unique (Dittrich and Di Fenizio, 2007).

Thus, any given set of species will react growing in qualitative novelty until it reaches its closure, but it is unclear whether the set will be stable in time, considering that during reactions species are consumed and their concentration could drop to zero. This motivates the study of dynamical properties of sets of species.

Dynamical Aspects

The *stoichiometric matrix* $\mathbf{S} = (s_{ij})$ associated with $\langle \mathcal{M}, \mathcal{R} \rangle$ is a $n \times k$ matrix, where s_{ij} is the stoichiometric coefficient of species m_i in the reaction R_j (s_{ij} is negative if species m_i is consumed by reaction R_j). Indeed, $s_{ij} = b^{ji} - a^{ji}$. The stoichiometric matrix is at the core of current systems biology (Schuster et al., 1999; Schilling and Palsson, 1998) and its properties have been extensively studied (Kacser and Burns, 1973; Heinrich and Rapoport, 1974). Let the *flux vector* $\mathbf{v} = (v_1, ..., v_k)$ be a non negative vector such that the application of $\mathbf{v}$ on the stoichiometric matrix $\mathbf{S}$ represents a reaction process, i.e. for $i = 1, ..., k$, the rate of the reaction R_i in the system is given by v_i. We define the *production rate vector* by $\mathbf{f} = \mathbf{Sv}$. Thus, for $i = 1, ..., n$, we have that $\mathbf{f}_i$ is the rate of production of the species m_i in the reaction process determined by $\mathbf{v}$.

We can describe the dynamics of the species concentrations $\mathbf{x} = (\mathbf{x}_1, ..., \mathbf{x}_n)$ by the system of ODEs

$$\dot{\mathbf{x}} = \mathbf{Sv}(\mathbf{x}, \mathbf{k}), \tag{1}$$

where according to mass-action kinetics

$$\mathbf{v}_i = \mathbf{k}_i \prod_{j=1}^{n} \mathbf{x}_j^{a_{ij}}$$

for $i = 1, \ldots, k$, is the *flux*, and $\mathbf{k} = (\mathbf{k}_1, ..., \mathbf{k}_k)$ is a strictly positive vector denoting reaction rate constants. We call ODE (1) a *chemical reaction system*.

In order to relate the statical domain with the dynamical domain, we introduce the idea of abstractions and instances:

Definition 3 *The **abstraction** of state* $\mathbf{x}$ *is the set* $\phi(\mathbf{x})$ *with*

$$\begin{aligned} \phi :& \mathbb{R}^n_{\geq 0} \mapsto \mathcal{P}(\mathcal{M}) \\ & \mathbf{x} \mapsto \phi(\mathbf{x}) \qquad \equiv \{m_i \in \mathcal{M} : \ \mathbf{x}_i > \epsilon\}, \end{aligned} \tag{2}$$

where $\mathbb{R}^n_{\geq 0}$ *denotes the set of non-negative real numbers, and* ϵ *is a concentration threshold. Moreover, given a set of species* $X \subseteq \mathcal{M}$*, a state* $\mathbf{x}$ *is an **instance** of* X *if and only if its abstraction equals* X*.*

Chemical Organizations

The following definition is at the core of chemical organization theory:

Definition 4 *A subset of species* $X \subseteq \mathcal{M}$ *is an **organization** if and only if*

1. X *is **closed** and*
2. X *is **self-maintaining**, i.e. there is a strictly positive flux vector* $\mathbf{v}$ *so that*

$$\mathbf{S}_X \mathbf{v} \geq \mathbf{0}$$

where $\mathbf{S}_X$ *is the stoichiometric matrix associated to the Algebraic Chemistry* $\langle X, \mathcal{R}_X \rangle$*.*

Organizations are sets of species which cannot produce new species by their possible reaction set. Also, it is possible that during the operation of an organization, the concentration of none of the species decreases; thus, an organization either maintains itself in time or grows in terms of the concentration of its species. This definition shares fundamental properties with that of autopoietic systems to the extent that all autopoietic systems are organizations. Note that not all organizations are autopoietic systems as an organization that keeps growing is not homeostatic and will eventually rupture its container. This motivates the study of the fixed points and other attractors of the chemical reaction systems.

The following theorem relates fixed points and organizations[1].

Theorem 1 *If* $\mathbf{x}$ *is a fixed-point of the ODE (1), i.e.* $\mathbf{S}\mathbf{v}(\mathbf{x}, \mathbf{k}) = \mathbf{0}$*, then the abstraction* $\phi(\mathbf{x})$ *is an organization.*

Fixed points are related to the dynamic stability of chemical systems. Moreover, since fixed points determine most of the characteristics of the dynamic systems they belong to (Strogatz, 2000), Theorem 1 provides a link between the long-term behavior of a chemical reaction system and its underlying reaction network. This allows the study of the system's dynamics by the chemical organization theory. Furthermore in (Peter and Dittrich, 2011), Theorem 1 is extended to other stable asymptotic behaviours such as periodic orbits and limit cycles. In addition, the necessary conditions for the existence of adequate flux vectors are explored in (Peter et al., 2010). Note that a fixed point in this context does not refer to thermodynamic equilibrium but to the maintenance of the size of the system in terms of the number of its components. The question about stability refers to the conservation of the structure or organization of the processes in a given timescale as the system is also subject to an evolutionary dynamic which can lead to change or desintegration. Now that we have introduced the idea of organization and shown some relevant aspects, we will focus on our main result; the decomposition theorem.

[1] Proof can be found in (Dittrich and Di Fenizio, 2007).

Species Role in a Network

The idea behind the role of a species is that it can be classified in relation to a set of species by how it behaves in the set of possible reactions.

Reactivity and Catalysts

Definition 5 *Let* $m \in X$*, then*

- m *is non-reactive w.r.t* X *if and only if for all reactions* $R = A \to B \in \mathcal{R}_X$*,* m *is not present in* A *nor in* B*.*
- m *is a catalyst w.r.t* X *if and only if for some reaction* $R' = A' \to B' \in \mathcal{R}_X$*,* m *is present in* A *and for all reactions* $R = A \to B \in \mathcal{R}_X$*,* $\mathcal{A}(A, m) = \mathcal{A}(B, m)$*.*
- m *is reactive w.r.t* X *if and only if for some reaction* $R' = A' \to B' \in \mathcal{R}_X$*,* m *is present either in* A' *or in* B' *and for some reaction* $R = A \to B \in \mathcal{R}_X$*,* $\mathcal{A}(A, m) \neq \mathcal{A}(B, m)$*.*

We say that $Y \subseteq X$ *is a non-reactive, catalytic or reactive set of* X*, if for all* $m \in Y$*,* m *is non-reactive, a catalyst or reactive w.r.t* X *respectively.*

The following lemma is straightforward

Lemma 1 *There is a unique maximal non-reactive, catalytic and reactive set of* X*.*

Definition 6 *The maximal non-reactive, catalytic and reactive sets of* X *are called the non-reactive, catalytic and reactive sets of* X *respectively.*

Overproduction

Definition 7 *Consider the Algebraic Chemistry* $\langle \mathcal{M}, \mathcal{R} \rangle$ *and a non-negative flux vector* $\mathbf{v}$ *such that* $(\mathbf{S}\mathbf{v})_i = \mathbf{f}_i \geq 0$ *for* $i = 1, ..., n$*. If* $\mathbf{f}_j > 0$ *for some* $j = 1, ..., n$*, we say that* m_j *is an overproduced species in* $\langle \mathcal{M}, \mathcal{R} \rangle$*.*

Overproduced species have a positive production rate for certain flux vectors which do not lead to the consumption of any other species. We remark that on the one hand, the definition of overproduced species does not demand that the system is self-maintaining because the flux vector is only required to be non-negative, but on the other hand, overproduced species definition not only requires the non-negative production of all the species, but also the positive production of at least one species. Thus, overproduced species are the species that can be *indefinitely* produced by some *reaction pathway*. Note that although this seems to violate the law of mass conservation, real systems require a constant input of mass or energy, and thus, it is usual when simulating or analyzing chemical networks to include an outer source of mass which does not decrease when consumed by a reaction. The relevance of these species is that they can actually be overproduced without consuming any of the inner species of the system; hence, they embody the notion of input. The following lemma is straightforward.

Lemma 2 *Let an overproduced species $m \in X$ in $\langle X, \mathcal{R}_X \rangle$. If $X \subset Y$, then m is overproduced in $\langle Y, \mathcal{R}_Y \rangle$.*

Corollary 1 *If X is a set of overproduced species in $\langle \mathcal{M}, \mathcal{R} \rangle$, then its closure $G_{CL}(X)$ is also overproduced.*

Lemma 3 *There exists a unique maximal set F of overproduced species in $\langle X, \mathcal{R}_X \rangle$.*

Proof If there are no overproduced species in X then the maximal overproduced set is the empty set. Otherwise the set containing all the overproduced species in $\langle X, \mathcal{R}_X \rangle$ leads to a maximal overproduced set. Now we are going to prove that the maximal overproduced set is unique. Suppose that there are two maximal overproduced sets $F_1, F_2 \in X$ and $F_1 \neq F_2$, let $\mathbf{v_1}, \mathbf{v_2}$ the flux vectors required to verify the overproduced property of F_1 and F_2 w.r.t $\langle X, \mathcal{R}_X \rangle$ respectively. Trivially, $\mathbf{v}_1 + \mathbf{v}_2$ verifies the overproduced property of $F_1 \cup F_2$ w.r.t $\langle X, \mathcal{R}_X \rangle$, and $F_i \subset F_1 \cup F_2$ for $i = 1, 2$. As the inclusion is strict we have a contradiction.

Definition 8 *The maximal set of overproduced species F with respect to X is called the overproduced set of X.*

Remark Consider the situation of adding a species m to an organization O. The fact that m is overproduced in $O' = O \cup \{m\}$ does not guarantee that O' is an organization. For example, consider the set of species $O' = \{a, b, c\}$ and the set of reactions

$$\begin{aligned} &R_1 = a \to b, \quad R_2 = b \to a, \\ &R_3 = a + c \to a + 2c, \quad R_4 = b + c \to \emptyset. \end{aligned} \tag{3}$$

We have that $O = \{a, b\}$ is an organization, c is overproduced in O', but O' is not an organization.

Then, at first sight the overproduced species could be seen as a useless definition concerning the self-maintainance of a reaction network because the overproduced species (c in the example) can catalyze the consumption of species that cannot be recovered in the network (b through reaction R_4 in the example). However, the identification of the overproduced molecules of a set X simplifies the verification of the self-maintaining condition of any set that contains X. Indeed in the example above, we have that c is overproduced in $\{a, c\}$, thus we can avoid the calculation of the production of the species c when verifying the self-maintainance of O'.

Roles and Organizations

From now on let N, E, F the non-reactive, catalyst and overproduced set of X respectively.

Definition 9 *$X - (F \cup E \cup N)$ is the potential active cycle (PAC) of X w.r.t F.*

Remark For any given flux vector which verifies the self-maintainance of X, the PAC has a production rate equal to zero. But PAC should not be confused with the set of species with a production rate equal to zero. Indeed, the non-reactive and catalytic species have a production rate zero, but they do not belong to the PAC. The following lemma states that no species can be only produced or only consumed in the PAC of an organization:

Lemma 4 *Let C be the PAC of X. If X is an organization, then for every $m \in C$ we have that m is consumed by some reaction $R \in \mathcal{R}_X$ and produced by other reaction $R' \in \mathcal{R}_X$.*

Proof Let $m \in C$, then m cannot be non-reactive either catalyst. As m has production zero, then m is a reactive species w.r.t X. Then, there must exist a reaction $R = A \to B \in \mathcal{R}_X$ s.t $\mathcal{A}(A, m) \neq \mathcal{A}(B, m)$. If $\mathcal{A}(A, m) > \mathcal{A}(B, m)$, as X is an organization, there has to exist some reaction $R' = A' \to B'$ s.t $\mathcal{A}(A', m) < \mathcal{A}(B', m)$. On the other hand, if $\mathcal{A}(A, m) < \mathcal{A}(B, m)$, as m is not overproduced (because $m \in C$), there has to exist some reaction $R' = A' \to B'$ s.t $\mathcal{A}(A', m) > \mathcal{A}(B', m)$.

PAC and Dependent Connectivity

We are going to define a special notion of connectivity which will allow us to separate the PAC of a set X in a number of partially non-overlapping sub-PACs, such that the self-maintainance of X can be studied from the self-maintainance of the sub-PACs. From now on, let X be a closed set[2].
The following definition of connectivity appears in (Centler et al., 2008a):

Definition 10 *Two species m_i and m_j in X are directly connected in $\langle X, \mathcal{R}_X \rangle$ if and only if there exist a reaction $R = A \to B \in \mathcal{R}_X$ such that $\{m_i, m_j\} \subseteq A \cup B$.*

[2]Verifying the closed property, and obtain the closure of a set of species is trivial compared with verifying its self-maintainance (Centler et al., 2008a)

Definition 11 *Two species m_i and m_j in X are connected in $\langle X, \mathcal{R}_X \rangle$ if and only if there exist a sequence of species $m_0, ..., m_p \in X$ such that $m_i = m_0$, m_k and m_{k+1} are directly connected in $\langle X, \mathcal{R}_X \rangle$ for all $k = 0, ..., p-1$ and $m_p = m_j$.*

We present a more restricted notion of connectivity than definition 11. This restriction only connects species that are non-independent when verifying self-maintainance.

Definition 12 *Two species m_i and m_j in X are dependently connected in $\langle X, \mathcal{R}_X \rangle$ if and only if there exists a sequence of species $m_0, ..., m_p \in X - (E \cup F)$ such that $m_i = m_0$, m_k and m_{k+1} are directly connected in $\langle X, \mathcal{R}_X \rangle$ for all $k = 0, ..., p-1$ and $m_p = m_j$.*

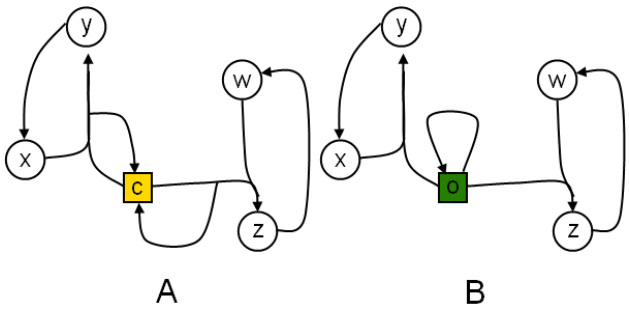

Figure 1: Following definition 11, both networks **A** and **B** are fully connected. Note that in **A**, the self-maintainance of $C_1 = \{x, y, c\}$ and $C_2 = \{z, w, c\}$ are independent. The same situation occurs in **B** with sets $C_1' = \{x, y, o\}$ and $C_2' = \{z, w, o\}$. *Dependent connection* allows to connect all the species in C_1 without connecting them to species in C_2 and viceversa because C_1 and C_2 are connected through a catalyst. Analogously C_1' and C_2' are not dependently connected because they are connected through an overproduced species.

Lemma 5 *Let $m, \bar{m} \in X$. m is dependently connected in $\langle X, \mathcal{R}_X \rangle$ to $\bar{m}$ if and only if $\bar{m}$ is dependently connected in $\langle X, \mathcal{R}_X \rangle$ to m.*

To continue, we need to mention that a computer science formalism called Petri Nets (Murata, 1989), has been considered as an interesting source of insights for the biochemical pathways research (Reddy et al., 1993, 1996). Petri Nets arose from the necessity to formalize concurrent processes. We will incorporate some fundamental topological parameters of Petri Nets to our analysis: the set of input transitions of a place and the set of input places of a transition.

Definition 13 *Let $m \in \mathcal{M}$. We define*

$$Act(m, \mathcal{R}) = \{R = A \to B \in \mathcal{R} \mid m \text{ is present in } A\}.$$

We say $Act(m, \mathcal{R})$ is the activable set of reactions of m in $\mathcal{R}$.

Definition 14 *Let $R = A \to B \in \mathcal{R}$. We define*

$$Req(R) = \{m \mid m \text{ is present in } A\}.$$

We say $Req(R)$ is the required set of species of R. Furthermore, for a set of reactions $S \subseteq \mathcal{R}$ we define

$$Req(S) = \bigcup_{R \in S} Req(R).$$

The set of input places in Petri Nets corresponds to the set Req$(\cdot)$, and the set of input transitions corresponds to Act$(\cdot, \mathcal{R})$.

Definition 15 *We define $Causal^*(m, \mathcal{R}_X)$ as the set of dependently connected species in $\langle X, \mathcal{R}_X \rangle$ to m, and $Causal(m, \mathcal{R}_X) = Req(Act(Causal^*(m, \mathcal{R}_X)))$.*

The following lemmas are derived straightforward from lemma 5 and definition 15:

Lemma 6 *Let $m, \bar{m} \in \mathcal{M}$. $\bar{m} \in Causal^*(m, \mathcal{R}_X)$ if and only if*

$$Causal^*(m, \mathcal{R}_X) = Causal^*(\bar{m}, \mathcal{R}_X).$$

Lemma 7 *Let $R \in \mathcal{R}_X$, $m, \bar{m} \in X - (E \cup F)$ s.t $m \notin Causal(\bar{m}, \mathcal{R}_X)$. If $R \in Act(Causal(m, \mathcal{R}_X) \cap Causal(\bar{m}, \mathcal{R}_X), \mathcal{R}_X)$ then $R \in \mathcal{R}_{E \cup F}$.*

$Causal^*(\cdot, \cdot)$ provides a way to split a set X of species in dependent connected subsets. It is necessary to identify a the catalytic set E and the overproduced set F to generate such separation. The more elements are in $E \cup F$, the more chance of recognize the independent causal connected sets we have.

Lemma 8 *Let D be the PAC of X. Then*

$$D = \bigcup_{m \in D} Causal^*(m, \mathcal{R}_X)$$

Proof Note that $D \subseteq \bigcup_{m \in D} Causal^*(m, \mathcal{R}_X)$. Let $m \in \bigcup_{m' \in D} Causal^*(m', \mathcal{R}_X)$, then for some species $m' \in D$ we have m is dependently connected to m', then m' is also dependently connected to m. This means m is a reactive, non overproduced, and non-catalytic species. Then $m \in D$.

From now on we let $D \subseteq X$ be the potential active cycle of X.

Definition 16 *Any set $D' \subseteq D$ s.t $D = \bigcup_{m \in D'} Causal^*(m, \mathcal{R}_X)$ is called a base of D. Any minimal cardinality base of D is called a minimal base of D.*

Lemma 9 *Let D', D'' be two minimal bases of D. Then every species in D' is dependently connected to one and only one species of D''.*

Proof Let $m \in D'$ and suppose that there is no species in D'' dependently connected to m. By corollary 6 we have $Causal^*(m, \mathcal{R}_X)$ is not contained in $\bigcup_{m' \in D''} Causal^*(m, \mathcal{R}_X)$. Then there has to be at least one species dependently connected to m in D''. Now suppose there is more than one species dependently connected to m in D''. Let $m_1, m_2 \in D''$ such species. As m_1 and m_2 are dependently connected to m, then m_1 and m_2 are dependently connected. By corollary 6 we have $Causal^*(m_1, \mathcal{R}_X) = Causal^*(m_2, \mathcal{R}_X)$. Then D'' is not a minimal base of D.

A minimal base of D is a set which generates all the non dependent sub-PACs of the D. We are going to prove that the self-maintainance of a potential active cycle can be obtained from the self-maintainance of its non dependent sub-PACs.

Lemma 10 *Let D' a minimal base of D. Then*

$$Act(D, \mathcal{R}_X) = \bigcup_{m \in D'} Act(Causal^*(m, \mathcal{R}_X), \mathcal{R}_X).$$

Proof Note that

$$\bigcup_{m \in D'} Act(Causal^*(m, \mathcal{R}_X), \mathcal{R}_X) \subseteq Act(D, \mathcal{R}_X).$$

Let $R \in Act(D, \mathcal{R}_X)$ then for some $m \in D$ we have $R \in (m, \mathcal{R}_X)$. From definition 16 we have that there is $m' \in D'$ s.t $m \in Causal^*(m', \mathcal{R}_X)$. Then by corollary 6 we have

$$R \in Act(Causal^*(m', \mathcal{R}_X), \mathcal{R}_X)$$
$$\subseteq \bigcup_{m \in D'} Act(Causal^*(m, \mathcal{R}_X), \mathcal{R}_X).$$

Decomposition Theorem for Organizations

Theorem 2 *Let $D' = \{\bar{m}_1, ..., \bar{m}_d\}$ a minimal base of D. For $i = 1, ..., d$ let*

$$D_i = Causal(\bar{m}_i, \mathcal{R}_X),$$
$$F_i = Causal(\bar{m}_i, \mathcal{R}_X) \cap F.$$

Let $\emptyset \to Y = \{\emptyset \to y \ / \ y \in Y\}$. X is self-maintaining if and only if for all $i = 1, ..., d$ we have that D_i is self-maintaining in the subnetwork $\langle D_i, \mathcal{R}_{D_i} \cup \emptyset \to F_i \rangle$.

Proof $\Rightarrow$: Let $\bar{F} = \bigcup_{i=1}^{d} F_i$. Note that X is self-maintaining in $\langle X, \mathcal{R}_X \rangle$ if and only if X is self-maintaining in $\langle X, \mathcal{R}_X \cup \emptyset \to \bar{F} \rangle$. Let $\mathbf{v}$ be a vector which verifies the self-maintainance of X in $\langle X, \mathcal{R}_X \rangle$. Let $Act(D_i, \mathcal{R}_X) = \{R_{\alpha_1}, ..., R_{\alpha_l^i}\}$, then $\bar{\mathbf{v}}$ lead to a non-negative production on all the species of $Causal^*(\bar{m}_i, \mathcal{R}_X)$ where

$$\bar{\mathbf{v}}_i = \begin{cases} \mathbf{v}_i & \text{if } i = \alpha_j \text{ for some } j, \\ 0 & \text{else} \end{cases}$$

As the rest of species belong to $\bar{F}$, to reach their non-negative production we use the reactions in $\emptyset \to \bar{F}$.

$\Leftarrow$: Let $\mathbf{v_1}, ..., \mathbf{v_d}$ the flux vectors which verifies the self-maintainance of $\langle D_i, \mathcal{R}_{D_i} \cup \emptyset \to F_i \rangle$, $i = 1, ..., d$ and $\mathbf{v}^F$ the flux vector which verifies the potential overproduction of F w.r.t X. Then there exist a non-negative number β s.t $\beta \mathbf{v}^F + \sum_{i=1}^{d} \bar{\mathbf{v}}_\mathbf{i}$ verifies the self-maintainance of X, where $\bar{\mathbf{v}}_\mathbf{i}$ is the flux $\mathbf{v_i}$ represented as a flux vector for $\mathcal{R}_X$, i.e. completed with zeros in the coordinates representing reactions that are not in $\mathcal{R}_{D_i}$.

Corollary 2 *Let D' a minimal base of D. Then X can be non-overlapping decomposed (partitioned) as*

$$X = N \cup E \cup F \cup D_1^* \cup \cdots \cup D_d^*. \quad (4)$$

With $D_i^* = Causal^*(m_i, R_X)$, and m_i the $i-$th element of D'. Moreover X is self-maintaining if and only if $E \cup F \cup D_i^*$ is self-maintaining for $i = 1, ..., d$.

Stability, COT and Autopoiesis

Living organisms are systems far from thermodynamic equilibrium, therefore the question about stability refers to the conservation of the organization of the organism's processes. This issue requires some attention, as the processes an autopoietic system are homeostatic in essence, but they are always potentially subject to dramatic changes. A clear example of this is the cell cycle, which is driven by cyclic, i.e. non stationary, processes like the cyclins proteins family expression patterns. In relation to this, the autopoiesis theory describes living organisms as processes that produce the components that give rise to those processes, where some attributes are preserved and others may change (Varela et al., 1974) allowing a structural drift (Maturana and Mpodozis, 2000). Therefore, in a way, autopoietic systems are not obliged to exhibit stability in the long run.

The basis of structural drift in biochemical networks is metabolic regulation. This is managed by the modulation of enzymes expression and also by means of a co-catalysis phenomenon in which coenzymes and regulators interact with the enzyme structure-function relation. This metabolic regulation determines the cell's developmental *direction* between a wide range of possible organizations. In order to illustrate this idea, imagine a system *following* an attractor when suddenly the concentration of a given regulator reaches a level that triggers a deep change in the structure of the network. Now, the attractor mentioned above is absent and the system takes a different pathway in which another phase space shift can occur. If this process becomes cyclic, the entire loop can be described as a limit cycle, but a decomposition of phase space in a set of *contextual* mini phase spaces that emerge in different regulatory scenarios could contribute both to the understanding of biological operations as to the algorithms research in biology inspired AC simulations. Therefore, in the COT, the changes in structure can be represented as changes in the phase space, leading to a dependence of phase space with the state of enzymes and regulators.

Some future goals in biologically inspired AC are the study of cyclic behaviour and autopoiesis theory's structural coupling. Preliminar studies concerning artificial autopoietic systems have been done in (Peter et al., 2010), by making use of the P-systems formalism. A P-system is formed by a set reaction networks, each one enclosed by a membrane. The reaction networks can interact diffusing particles through membranes. In particular, it is shown how a bistable cyclic process can be reached among two different unstable reaction networks, by exchanging in low rates their subproducts, i.e. forming together an organization.

Conclusions

In this paper we have shown how the chemical organization theory connects deeply to notions of autopoietic systems. Moreover, every autopoietic system is an organization, and thus, theorems derived for organizations are valid for autopoietic systems.

We introduce the notion of the role played by a species in a subnetwork of an reaction network. The different possible roles that a species can play in a subnetwork (non-reactive, catalyst, overproduced, active cyclic) give information about the structure of the subnetwork (lemmas 1, 3 and 4). We also introduce the notion of dependent connectivity in a reaction network (definition 12), which is useful to split a reaction network into the minimal parts required to verify the self-maintainance (Theorem 2). This theorem helps to simplify the organization verification not only by decomposing the set, but also when we keep track of the decomposition of the set to verify the self-maintainance of sets that contain it.

The fact that an organization can be subdivided into self-maintaining subnetworks which are mostly independent one from another (and not necessarily closed), is both striking and noteworthy. The decomposition theorem stated in this work shows that the long-term behaviour of an organization can depend on sets of species whose states are weakly coupled. This result opens new paths of analysis for a broad set of fields, from metabolic dynamics to ecological networks. This result also relates directly to a debated subject which is the composition of autopoietic systems by other autopoietic systems.

At this point a comment on the domain of applicability of this theorem is convenient. On some cases, like reactive flow systems where there is a spontaneuous decay of every species, the decomposition is trivial, i.e., the system cannot be further subdivided. This is because each species of an organization would be overproducible (to counteract the decay). So, this paper addresses those systems where some species take part in reactions but do not decay spontaneously. In the case of living systems, every molecule decays spontaneously or, equivalently, dilutes when the system grows. This seems to make an argument towards the unapplicability of the decomposition theorem to living systems. However, the fact that every molecule decays in a living system is only revelant at the right timescale. For a smaller timescale some molecules do not decay. Therefore, it is important to examine living systems at different timescales. Choosing a very long time scale, basically no system would continue to exist so that autopoiesis would not become visible. While with a smaller timescale more and more elements (molecules) would become stable and would not decay spontaneously. Here is where the decomposition can potentially be applied.

This decomposition theorem is suggested as a starting point for the complexity analysis of the organization verification problem as well as for the classification of reaction networks in terms of how complicated it is to compute their organizational structure. Although this is a relevant result for current systems biology as it simplifies analysis and simulations, and for artificial life, at it may guide the construction of artificial self-sustaining systems, it also demands revisiting questions on biological systems concerning their modularity, autonomy, and even the notion of their identity.

References

Benkö, G., Centler, F., Dittrich, P., Stadler, B. M. R., and Stadler, P. F. (2009). A topological approach to chemical organizations. *Artif. Life*, 15:71–88.

Centler, F. and Dittrich, P. (2007). Chemical organizations in atmospheric photochemistries: A new method to analyze chemical reaction networks. *Planet. Space Sci.*, 55:413–428.

Centler, F., Kaleta, C., Di Fenizio, P. S., and Dittrich, P. (2008a). Chemical organizations in a toy model of the political system. *Advances in Complex Systems*, 11(4):609–627.

Centler, F., Kaleta, C., Di Fenizio, P. S., and Dittrich, P. (2008b). Computing chemical organizations in biological networks. *Bioinformatics*, 24:1611–1618.

Cornish-Bowden, A., Cárdenas, M. L., Letelier, J. C., and Soto-Andrade, J. (2007). Beyond reductionism: Metabolic circularity as a guiding vision for a real biology of systems. *PROTEOMICS*, 7:839–845.

Dittrich, P. and Di Fenizio, P. S. (2007). Chemical organization theory. *Bulletin of Mathematical Biology*, 69:1199–1231.

Dittrich, P. and Winter, L. (2008). Computing chemical organizations in biological networks. *Bioinformatics*, 24(14):1611–1618.

Eigen, M. and Schuster, P. (1977). A principle of natural self-organization. *Naturwissenschaften*, 64:541–565.

Fontana, W. and Buss, L. (1994). The arrival of the fittest: Toward a theory of biological organization. *Bulletin of Mathematic Biology*, 56:1–64.

Heinrich, R. and Rapoport, T. A. (1974). A linear steady-state treatment of enzymatic chains. *European Journal of Biochemistry*, 42:89–95.

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., Hernndez, V., Soto-Andrade, J., Crdenas, M., Cornish-Bowden, A., and Letelier, J. C. (2010). (m,r) systems and raf sets: Common ideas, tools and projections. In Fellermann, H., Drr, M., Hanczyc, M. M., Laursen, L. L., Maurer, S., Merkle, D., Monnard, P. A., Stoy, K., and Rasmussen, S., editors, *Artificial Life XII*, pages 94–100. MIT Press, Cambridge, MA.

Kacser, H. and Burns, J. A. (1973). The control of flux. *Symposia of the Society for Experimental Biology*, 27:65–104.

Kaleta, C., Centler, F., and Dittrich, P. (2006). Analyzing molecular reaction networks from pathways to chemical organizations. *Molecular Biotechnology*, 34:117–123.

Kauffman, S. (1969). Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology*, 22:437–467.

Letelier, J. C., Marin, G., and Mpodozis, J. (2003). Autopoietic and (M,R) systems. *Journal of Theoretical Biology*, 222:261 – 272.

Matsumaru, N., Centler, F., di Fenizio, P., and Dittrich, P. (2006). Chemical organization theory applied to virus dynamics. *Information Technology*, 48:154–160.

Matsumaru, N., Centler, F., di Fenizio, P., and Dittrich, P. (2007). Chemical organization theory as a theoretical base for chemical computing. *International Journal on Unconventional Computing*, 3:285–309.

Maturana, H. and Mpodozis, J. (2000). The origin of species by means of natural drift. *Revista chilena de historia natural*, 73(2).

Maturana, H. and Varela, F. (1973). *De Máquinas y Seres Vivos: Una teoría sobre la organización biológica*. Editorial Universitaria.

Murata, T. (1989). Petri nets and its applications. *Proceedings of the IEEE*, 77:541–580.

Peter, S. and Dittrich, P. (2011). On the relation between organizations and limit sets in chemical reaction systems. *Advances in Complex Systems*, 14:77–96.

Peter, S., Veloz, T., and Dittrich, P. (2010). Feasibility of organizations - a refinement of chemical organization theory. *In: M. Gheorghe, T. Hinze, G. Paun (Eds.), Proc. of the Eleventh International Conference on Membrane Computing*, pages 369–382.

Reddy, V. N., Liebmant, M. N., and Mavrovouniotis, M. L. (1996). Qualitative analysis of biochemical reaction systems. *Comput Biol. Med.*, 26:9–24.

Reddy, V. N., Mavrovouniotis, M. L., and Liebmant, M. N. (1993). Petri net representations in metabolic pathways. *ISMB-93 Proceedings*, pages 328–336.

Rosen, R. (1958). A relational theory of biological systems. *Bulletin of Mathematical Biophysics*, 20:317–341.

Schilling, C. H. and Palsson, B. O. (1998). The underlying pathway structure of biochemical reaction networks. *Proceedings of the National Academy of Sciences*, 95:4193–4198.

Schuster, S., Dandekar, T., and Fell, D. A. (1999). Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. *Trends Biotechnology*, 17:53–60.

Strogatz, S. (2000). *Nonlinear Dynamics and Chaos*. Westview Press, Cambridge.

Varela, F., Maturana, H., and Uribe.R (1974). Autopoiesis: The organization of living systems, its characterization and a model. *Biosystems*, 5:187–196.

von Bertalanffy, L. (1968). *General System Theory: Foundations, Development, Applications*. George Braziller, New York.

Wiener, N. (1948). *Cybernetics: Or Control and Communication in the Animal and the Machine*. Hermann and Cie Editeurs, Paris.