Discrete Simulations of Biochemical Dynamics

Vincenzo Manca

Department of Computer Science, University of Verona, Italy
vincenzo.manca@univr.it

Abstract. Metabolic P systems, shortly MP systems, are a special class of P systems, introduced for expressing biological metabolism. Their dynamics are computed by *metabolic algorithms* which transform populations of objects according to a *mass partition* principle, based on suitable generalizations of chemical laws. The definition of MP system is given and a new kind of regulation mechanism is outlined, for the construction of computational models from experimental data of given metabolic processes.

Keywords: P Systems, Metabolism, Discrete Biological Models.

1 Introduction

In [7] a discrete perspective was introduced in the analysis of metabolic processes, which was then developed in papers [2,9,4,8,10,11], and which is focused on the notion of *Metabolic* P systems, shortly MP systems. Here, we outline the possibility of deducing an MP model, for a given metabolic process, from a suitable macroscopic observation of its behavior along a certain number of steps.

MP systems are a special type of P systems [12] which were proven to effectively model the dynamics of several biochemical processes: the Belousov-Zhabotinsky reaction (Brusselator) the Lotka-Volterra dynamics, the SIR (Susceptible-Infected-Recovered epidemic) [1], the leukocyte selective recruitment in the immunological response [5,1], the Protein Kinase C activation [2], circadian rhythms, mitotic cycles [8], [6][1].

The perspective introduced by MP systems can be synthesized by a principle which replaces the mass action principle. We call it the *mass partition principle* because, according to it, the system is observed along a discrete sequence of steps, and at each step, all the matter of any kind of substance, consumed in the time interval between two consecutive steps, is partitioned among all the reactions which need it for producing their products. If we are able to determine the amount of reactants that any reaction takes in that step, according to the stoichiometry of the reactions (which we assume to know), we can perfectly establish the amount of substances consumed and produced between two steps, therefore all the dynamics can be discovered. As a consequence of mass partition principle, two important aspects follow. In MP system rules act on

---

[1] The package Psim, developed in Java within the research group on *Natural Computing* led by the author, at the Department of Computer Science of the University of Verona (Italy), provides representations and dynamics generations of MP systems (Psim is available from the site of the *Center for BioMedical Computing*, at the University of Verona: www.cbmc.it).

object populations, rather than on single objects. Moreover, dynamics is deterministic at population level, but nothing can be said about the dynamical evolution of single objects.

2   Metabolic P Systems

MP systems are deterministic P systems where the transition to the next state (after some specified interval of time) is calculated according to a *mass partition strategy*, that is, the available matter of each substance is partitioned among all reactions which need to consume it. A special class of MP systems was proved to be stongly related to differential models [4]. The notion of MP system we consider here is based on those given in [8,10,11].

Let us consider a set X of substances and a set of R of reactions over them, as pairs of strings, represented in the arrow notation according to which any rule $r \in R$ is identified by $\alpha_r \to \beta_r$ with α_r, β_r strings over X (α_r represents the reactants of r, while β_r represents the products of r, for example, $aab \to cd$ is a reaction where two *molecules* of a with a *molecule* of b react by producing a *molecule* of c and a molecule of d).

For a string γ and a symbol x we denote by $|\gamma|_x$ the number of occurrences of the symbol x in γ, while $|\gamma|$ is the length of γ. Then, the *stoichiometric matrix* A_R correspondent to a set R of reactions over a set X of substances is defined by setting $A_R = (A_R(x, r) \mid x \in X, r \in R)$ and, for every $x \in X$ and $r \in R$, $A_R(x, r) = |\beta_r|_x - |\alpha_r|_x$. Moreover, we define $R_\alpha(x) = \{r \in R \mid |\alpha_r|_x > 0\}$. Two reactions r_1, r_2 are *competing* if $r_1, r_2 \in R_\alpha(x)$ for some substance $x \in X$. We call *regulator*, of a reaction r, any reactant of r or any reactant of a reaction which is competing with r.

Definition 1 (MP System). *An MP system is a construct*

$$M = (X, R, Q, U, \nu, \sigma, \tau, q_0, \Phi)$$

where:

- $X = \{x_1, \dots, x_n\}$ *is a finite set of substances (the types of molecules);*
- $R = \{r_1, \dots, r_m\}$ *is a finite set of reactions over X;*
- Q *is the set of states, that is, the functions $q : X \to \mathbb{R}$ from substances to real numbers. The state q of the instant i can be identified as a vector $(x_1[i], x_2[i], \dots, x_n[i])$ of real numbers, constituted by the values which are assigned, by q, to the elements of X.*
- $U = \{u_1, \dots, u_m\}$ *is the set of reaction units, where, for each rule r, u_r is a function from states to real numbers (the amount of molar quantity consumed/produced by the rule r in correspondence to any occurrence of reactant/product occurring in it);*
- ν *is a natural number which specifies the number of molecules of a (conventional) mole of M, as population unit of M;*
- σ *is a function which assigns to each $x \in X$, the mass $\sigma(x)$ of a mole of x (with respect to some measure unit);*

– τ is the temporal interval between two consecutive states;
– $q_0 \in Q$ is the initial state, also denoted by $X[0] = (x_1[0], x_2[0], \dots, x_n[0])$;
– Φ is a set of regulation maps.

The temporal evolution of an MP system M is calculated by means of the following system of autonomous first-order difference equations (1) (2), called metabolic algorithm, where $X[i]$ and $U[i]$ are the vectors of substance quantities and reaction units at step i, A_R is the stoichiometric matrix of dimension $n \times m$ corresponding to the reactions of R (n is the number of different substances and m the number of reactions), Φ is the vector of functions (as many as the reactions), and $\times, +$ are the usual matrix product and vector sum:

$$X[i+1] \;=\; (A_R \times U[i]\,) + X[i] \tag{1}$$
$$U[i] \;=\; \Phi(X[i]) \tag{2}$$

The parameters τ, ν, μ have no role in the mathematical definition of dynamics. Nevertheless, they are essential for giving a determinate physical meaning to the numerical values, according to a specific time/mass measure scale.

3   Metabolic Algorithms and Log-Gain Regulation

Given a real metabolic system that we can observe for a certain number of steps, is it possible to determine an MP system which could predict, within an acceptable approximation, the future behaviour of the given system? We will show how this task could be achieved. In fact, in some cases, we can determine, in a systematic way, an MP system which is an adequate model of some observed metabolic dynamics.

In order to discover the reaction units at each step, we introduce the notion of *log-gain regulation*. In fact, it seems to be perfectly natural that a proportion should exist among the relative variation of substances and the relative variation of the reaction unit of r. The relative variation of a substance x is defined as the ratio $\Delta(x)/x$. In differential notation (with respect to the time variable), this ratio is related to $\frac{dx}{dt}/x$, and from elementary calculus we know that it is the same as $\frac{d(\lg x)}{dt}$. This equation explains the term "log-gain" for expressing relative variations. In this way, we can derive the values of the reaction units at any observation time, therefore, these parameters determine the dynamics of MP systems. More precisely, we set the following principle.

Principle 2 (Log Gain Regulation). *For $i \geq 0$ let $Lg(u_r[i]) = (u_r[i+1] - u_r[i])/u_r[i]$ be the log-gain of the reaction unit u_r at the step i, and let $Lg(x[i]) = (x[i+1] - x[i])/x[i]$ be the log-gain of the substance x at the same step, then $Lg(u_r[i])$ is a linear combination of the log-gains of the regulators of r:*

$$Lg(u_r[i]) = \sum_{x \in X} p_{r,x} Lg(x[i]) + p_r \tag{3}$$

$p_{r,x}$ *with $x \in X$ are the log-gain parameters. If the parameter $p_{r,x} \neq 0$, then x is a regulator of r. The parameter p_r is called the log-gain offset of the rule r.*

Given the dynamics of a system that we *observe* for a sufficient number of steps, we want to know, with a sufficient precision, the (molar) quantities of all different kinds of molecules, for a sequence of steps. Let us denote these quantities with the sequence, for $i = 0, \ldots k$, of vectors:

$$X[i] = (x_1[i], x_2[i], \ldots, x_n[i])$$

Moreover, we assume to know the structure of the system, that is, kinds of substances, reactions, time unit, molar unit, and initial state. We want to predict the vectors $X[i]$ for steps $i > k$, which follows the observation steps. We solve the problem when we discover the regulation maps Φ.

Let us consider the system of equations $LG[i] + \Delta S[i + 1]$, obtained by putting together system (3) at step i with system (1) of Definition 1 at step $i + 1$. We call it **observation module**. This system of equations has $n + m$ equations. The variables of this system are the reaction units and the log-gain parameters (and offsets). In general, the number of these variables: $u_1[i + 1], u_2[i + 1], \ldots, u_m[i + 1], \ldots$ is greater than the number of equations. Moreover, in order to discover the dynamics underlying the passage of the MP system, from one step to its next step, it is enough to know. at any step. the value of reaction units. Despite the difference between the number of equations and the number of variables, the following theorem holds, as a consequence of log-gain principle (we omit the proof).

Theorem 3. *The system $LG[i] + \Delta S[i + 1]$ has one and only one solution.*

Let us assume to know $U[0]$ (in fact, there are some methods for determining it [11]). The value $X[0]$ is known because it corresponds to the initial state of the system. Therefore, if we solve this system for $i = 0$, that is, $LG[0] + \Delta S[1]$, we get the value of $U[1]$. So, if vectors $X[i]$ for $i = 1, \ldots, k$, are given by observation, we can apply the same procedure, again for $i = 1, 2, \ldots, k$, and get $U[2], U[3], \ldots, U[k + 1]$.

Now assume that these vectors depend on the substance quantities with some polynomial dependence of a given degree, say a third degree, then we can use some standard interpolation tools for finding the functional dependence of vector U with respect to the substance quantities. The resulting polynomials are some approximations of the regulation functions Φ we are searching for, and our task was completed. In fact, now we can use the metabolic system (1) (2) of Definition 1 for computing the evolution of the given MP systems in all the steps i for $i > k$.

We applied this method to many metabolic systems (e.g. Lotka-Volterra, Brusselator, and Mitotic Cycles) and we were able to reconstruct, almost exactly their dynamics. But this procedure assumes the knowledge of $U[0]$. Actually, there are several possibilities under investigation. However, we discovered experimentally a very interesting fact, which deserves a more subtle theoretical investigation. If we consider the system $\Delta S[0]$ and choose as $U[0]$ one of its infinite solutions (imposing some additional very natural constraints), then in many cases, we found that, independently from the chosen value of $U[0]$, after a small number of steps, say $k = 3$ steps, our procedure will generate, with a great approximation, the same vectors $U[i + k]$, for all $i > 0$. This means that the data collected in the observation steps are sufficient to determine the functions which, on the basis of substance quantities, regulate the dynamics of the system.

Numerical elaborations of our simulations were performed by MATLAB® standard operators (*backslash* operator for square matrix left division or in the least squares sense solution) and interpolation was performed by polynomials of third degree. Specific observation strategies were adopted, by using about one hundred steps. In almost all cases, the observed dynamics were correctly reconstructed. This means that the regulation functions, deduced according to the outlined method, provided MP systems with the same dynamics of the observed systems. In conclusion, in the case of natural systems, from suitable observations, we could discover, with good approximation, the underlying dynamical regulation maps, and consequently, reliable computational models of their dynamic. However, applications of our method to more complex dynamics and deeper theoretical analyses of the simulation results will be topics for further research.

References

1. Bianco, L., Fontana, F., Franco, G., Manca, V.: P systems for biological dynamics, pp. 81–126. in [3]
2. Bianco, L., Fontana, F., Manca, V.: P systems with reaction maps. International Journal of Foundations of Computer Science 17(1), 27–48 (2006)
3. Ciobanu, G., Păun, G., Pérez-Jiménez, M.J. (eds.): Applications of Membrane Computing. Springer, Heidelberg (2006)
4. Fontana, F., Manca, V.: Discrete solutions of differential equations by metabolic P systems. Theoretical Computer Science 372, 165–182 (2007)
5. Franco, G., Manca, V.: A membrane system for the leukocyte selective recruitment. In: Martín-Vide, C., Mauri, G., Păun, G., Rozenberg, G., Salomaa, A. (eds.) Membrane Computing. LNCS, vol. 2933, pp. 180–189. Springer, Heidelberg (2004)
6. Franco, G., Guzzi, P.H., Mazza, T., Manca, V.: Mitotic Oscillators as MP Graphs. In: Hoogeboom, H.J., Păun, G., Rozenberg, G., Salomaa, A. (eds.) WMC 2006. LNCS, vol. 4361, pp. 382–394. Springer, Heidelberg (2006)
7. Manca, V., Bianco, L., Fontana, F.: Evolutions and oscillations of P systems: Applications to biological phenomena. In: Mauri, G., Păun, G., Pérez-Jiménez, M.J., Rozenberg, G., Salomaa, A. (eds.) WMC 2004. LNCS, vol. 3365, pp. 63–84. Springer, Heidelberg (2005)
8. Manca, V., Bianco, L.: Biological networks in metabolic P systems. BioSystems (to appear) doi:10.1016/j.biosystems.2006.11.009
9. Manca, V.: MP systems approaches to biochemical dynamics: Biological rhythms and oscillations. In: Hoogeboom, H.J., Păun, G., Rozenberg, G., Salomaa, A. (eds.) WMC 2006. LNCS, vol. 4361, pp. 86–99. Springer, Heidelberg (2006)
10. Manca, V.: Metabolic P Systems for Biochemical Dynamics. Progress in Natural Sciences (Invited Paper) 17(4), 384–391 (2007)
11. Manca, V.: The Metabolic Algorithm for P Systems: Principles and Applications. Theoretical Computer Science (to appear, 2007)
12. Păun, G.: Membrane Computing. An Introduction. Springer, Heidelberg (2002)