Goldbeter's mitotic oscillator entirely modeled by MP systems

Vincenzo Manca Luca Marchetti

Center for Biomedical Computation (CBMC)
University of Verona, Department of Computer Science

web-site: http://www.cbmc.it

E-mail: luca.marchetti@univr.it

Eleventh International Conference on Membrane Computing (CMC11)
24-27 August 2010, Jena, Germany

Center for
Biomedical Computation

Outline

1. Introduction:
   - introduction to Metabolic P systems
     (i.e. the mathematical framework used for this work. . .)

     [Vincenzo Manca (2010) Metabolic P systems. Scholarpedia, 5(3):9273]
   - presentation of the Log-Gain Stoichiometric Step-wise regression (LGSS)
     (i.e. the regression algorithm used to create models descripted here. . .)

     [Vincenzo Manca, Luca Marchetti (2010) Log-Gain Stoichiometric Stepwise regression for MP systems. International Journal of Foundations of Computer Science, to appear]
2. Presentation of research results:
   - introduction to mitotic oscillations
     (i.e. the biological phenomenon under examination. . .)

     [A. Goldbeter (1991) A minimal cascade model for the mitotic oscillator involving cyclin and cdc2 kinase. PNAS 88(20): 9107911]
   - classification of mitotic MP models
     (i.e. the topic of our paper!!!)

Center for
Biomedical Computation

Outline

1. Introduction:
   - introduction to Metabolic P systems
     (i.e. the mathematical framework used for this work. . .)

     [Vincenzo Manca (2010) Metabolic P systems. Scholarpedia, 5(3):9273]
   - presentation of the Log-Gain Stoichiometric Step-wise regression (LGSS)
     (i.e. the regression algorithm used to create models descripted here. . .)

     [Vincenzo Manca, Luca Marchetti (2010) Log-Gain Stoichiometric Stepwise regression for MP systems. International Journal of Foundations of Computer Science, to appear]
2. Presentation of research results:
   - introduction to mitotic oscillations
     (i.e. the biological phenomenon under examination. . .)

     [A. Goldbeter (1991) A minimal cascade model for the mitotic oscillator involving cyclin and cdc2 kinase. PNAS 88(20), 91079111]
   - classification of mitotic MP models
     (i.e. the topic of our paper!!!)

Center for
Biomedical Computation

An introduction to MP systems

P systems have been proposed by Gh. Păun in '98 as a discrete computational model inspired by the central role of membranes in the structure and functioning of living cells.

[G. Păun. Computing with membranes. *J. Comput. System Sci.*, 61(1): 108–143, 2000.]

Metabolic P systems are a variant of P systems, apt to express biological processes.

[Vincenzo Manca (2010) Metabolic P systems. Scholarpedia, 5(3):9273]

Main features:

- A fixed membrane structure
  (many time only the skin membrane is used).
- A "biological" interpretation of objects as biological substances and of evolution rules as biological reactions.
- An evolution strategy based on a discrete, deterministic algorithm called Equational Metabolic Algorithm (EMA).

Main components of MP systems

An MP system can be represented by means of MP grammars and MP graphs.

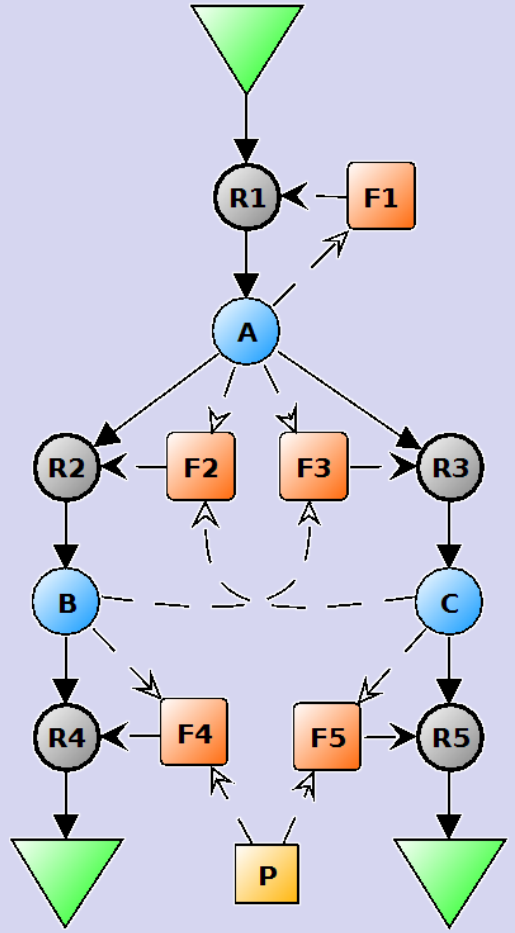
MP grammar

MP reactions	MP fluxes
$r_1 : \emptyset \to A$	$\varphi_1 = 0.1 + 3A$
$r_2 : A \to B$	$\varphi_2 = 0.2C$
$r_3 : A \to C$	$\varphi_3 = 0.1B$
$r_4 : B \to \emptyset$	$\varphi_4 = 0.6B + P$
$r_5 : C \to \emptyset$	$\varphi_5 = 0.4C + P$

$A[0], B[0], C[0] = 1 mol.$

$P[0] = 0.2, P[i+1] = P[i] + 0.2.$

MP graph



Main components of MP systems

- SUBSTANCES -

The types of molecules taking part to reactions...

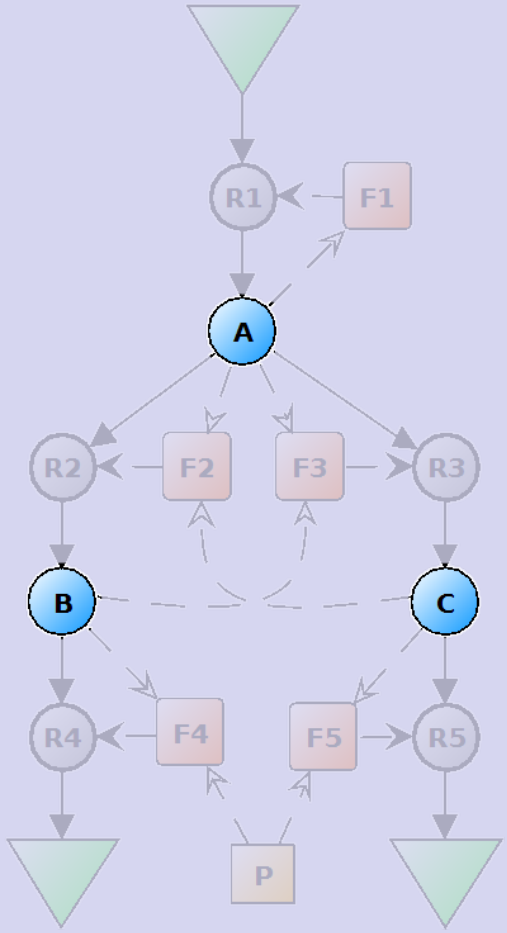
MP grammar

MP reactions	MP fluxes
$r_1 : \emptyset \to A$	$\varphi_1 = 0.1 + 3A$
$r_2 : A \to B$	$\varphi_2 = 0.2C$
$r_3 : A \to C$	$\varphi_3 = 0.1B$
$r_4 : B \to \emptyset$	$\varphi_4 = 0.6B + P$
$r_5 : C \to \emptyset$	$\varphi_5 = 0.4C + P$

$A[0], B[0], C[0] = 1 mol.$

$P[0] = 0.2, P[i+1] = P[i] + 0.2.$

MP graph



Main components of MP systems

- REACTIONS -

Evolution rules for matter transformation...

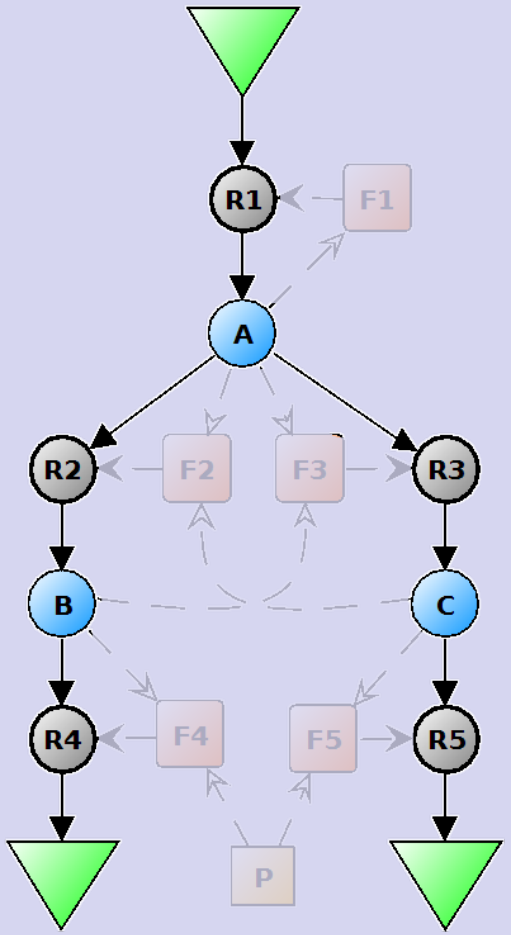
MP grammar

MP reactions	MP fluxes
$r_1 : \emptyset \to A$	$\varphi_1 = 0.1 + 3A$
$r_2 : A \to B$	$\varphi_2 = 0.2C$
$r_3 : A \to C$	$\varphi_3 = 0.1B$
$r_4 : B \to \emptyset$	$\varphi_4 = 0.6B + P$
$r_5 : C \to \emptyset$	$\varphi_5 = 0.4C + P$

$A[0], B[0], C[0] = 1 mol.$

$P[0] = 0.2, P[i+1] = P[i] + 0.2.$

MP graph



Main components of MP systems

- FLUXES -

Functions which give the evolution of the system...

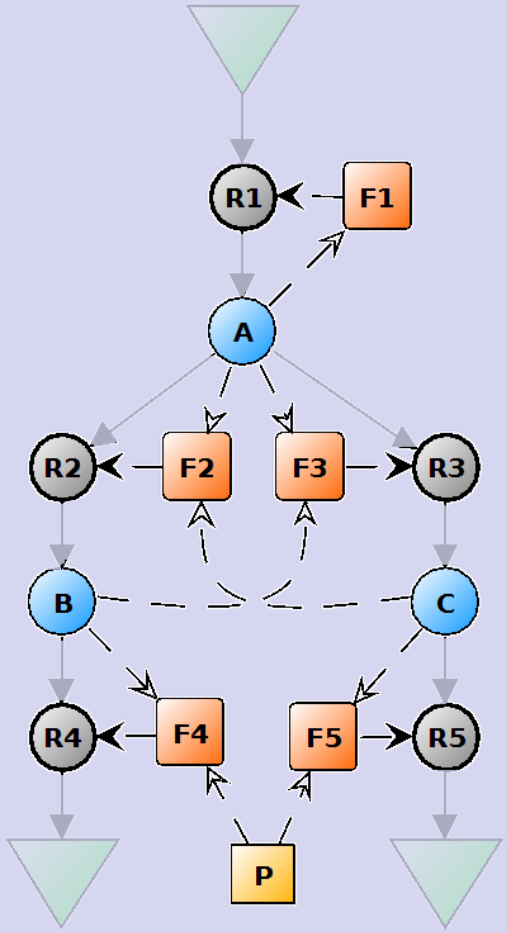
MP grammar

MP reactions	MP fluxes
$r_1 : \emptyset \to A$	$\varphi_1 = 0.1 + 3A$
$r_2 : A \to B$	$\varphi_2 = 0.2C$
$r_3 : A \to C$	$\varphi_3 = 0.1B$
$r_4 : B \to \emptyset$	$\varphi_4 = 0.6B + P$
$r_5 : C \to \emptyset$	$\varphi_5 = 0.4C + P$

$A[0], B[0], C[0] = 1 mol.$

$P[0] = 0.2, P[i+1] = P[i] + 0.2.$

MP graph



Main components of MP systems

Equational Metabolic Algorithm

MP grammar

MP reactions	MP fluxes
$r_1 : \emptyset \to A$	$\varphi_1 = 0.1 + 3A$
$r_2 : A \to B$	$\varphi_2 = 0.2C$
$r_3 : A \to C$	$\varphi_3 = 0.1B$
$r_4 : B \to \emptyset$	$\varphi_4 = 0.6B + P$
$r_5 : C \to \emptyset$	$\varphi_5 = 0.4C + P$

$A[0], B[0], C[0] = 1 mol.$

$P[0] = 0.2, P[i+1] = P[i] + 0.2.$

EMA

For each step i of computation:

1) we compute reaction units:

$$u_{1,2,\ldots,5}[i] = \varphi_{1,2,\ldots,5}[i]$$

$$\begin{aligned} u_1[i] &= 0.1 + 3A[i] \\ u_2[i] &= 0.2C[i] \\ u_3[i] &= 0.1B[i] \\ u_4[i] &= 0.6B[i] + P[i] \\ u_5[i] &= 0.4C[i] + P[i] \end{aligned}$$

Ex: $u_1[i]$ gives the amount of substance which is produced and consumed by r_1 at step i.

Main components of MP systems

Equational Metabolic Algorithm

MP grammar

MP reactions	MP fluxes
$r_1 : \emptyset \to A$	$\varphi_1 = 0.1 + 3A$
$r_2 : A \to B$	$\varphi_2 = 0.2C$
$r_3 : A \to C$	$\varphi_3 = 0.1B$
$r_4 : B \to \emptyset$	$\varphi_4 = 0.6B + P$
$r_5 : C \to \emptyset$	$\varphi_5 = 0.4C + P$

$A[0], B[0], C[0] = 1 mol.$

$P[0] = 0.2, P[i+1] = P[i] + 0.2.$

EMA

For each step i of computation:

1) we compute reaction units:

$$u_{1,2,\ldots,5}[i] = \varphi_{1,2,\ldots,5}[i]$$

2) we compute the variation of each substance $\Delta_{A,B,C}[i]$:

$$\begin{aligned} \Delta_A[i] &= u_1[i] - u_2[i] - u_3[i] \\ \Delta_B[i] &= u_2[i] - u_4[i] \\ \Delta_C[i] &= u_3[i] - u_5[i] \end{aligned}$$

Ex: $\Delta_A[i]$ is increased of $u_1[i]$ because r_1 produces A and decreased of $u_2[i] + u_3[i]$ because r_2, r_3 consume A.

Main components of MP systems

Equational Metabolic Algorithm

MP grammar

MP reactions	MP fluxes
$r_1 : \emptyset \to A$	$\varphi_1 = 0.1 + 3A$
$r_2 : A \to B$	$\varphi_2 = 0.2C$
$r_3 : A \to C$	$\varphi_3 = 0.1B$
$r_4 : B \to \emptyset$	$\varphi_4 = 0.6B + P$
$r_5 : C \to \emptyset$	$\varphi_5 = 0.4C + P$

$A[0], B[0], C[0] = 1 mol.$

$P[0] = 0.2, P[i+1] = P[i] + 0.2.$

EMA

For each step i of computation:

1) we compute reaction units:

$$u_{1,2,\ldots,5}[i] = \varphi_{1,2,\ldots,5}[i]$$

2) we compute the variation of each substance $\Delta_{A,B,C}[i]$:

$$\begin{aligned} \Delta_A[i] &= u_1[i] - u_2[i] - u_3[i] \\ \Delta_B[i] &= u_2[i] - u_4[i] \\ \Delta_C[i] &= u_3[i] - u_5[i] \end{aligned}$$

3) we compute the next state:

$$\begin{aligned} A[i+1] &= A[i] + \Delta_A[i] \\ B[i+1] &= B[i] + \Delta_B[i] \\ C[i+1] &= C[i] + \Delta_C[i] \end{aligned}$$

Main components of MP systems

Equational Metabolic Algorithm

MP grammar

MP reactions	MP fluxes
$r_1 : \emptyset \to A$	$\varphi_1 = 0.1 + 3A$
$r_2 : A \to B$	$\varphi_2 = 0.2C$
$r_3 : A \to C$	$\varphi_3 = 0.1B$
$r_4 : B \to \emptyset$	$\varphi_4 = 0.6B + P$
$r_5 : C \to \emptyset$	$\varphi_5 = 0.4C + P$

$A[0], B[0], C[0] = 1 mol.$

$P[0] = 0.2, P[i+1] = P[i] + 0.2.$

EMA

More "algebrically", the vector of substance variation

$$\Delta[i] = (\Delta_A[i]; \Delta_B[i]; \Delta_C[i])$$

is given by the following matrix product:

$$\underbrace{\underbrace{\begin{pmatrix} 1 & -1 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{pmatrix}}_{\mathbb{A}} \times \underbrace{\begin{pmatrix} u_1[i] \\ u_2[i] \\ u_3[i] \\ u_4[i] \\ u_5[i] \end{pmatrix}}_{U[i]}}_{\Delta[i] = \mathbb{A} \times U[i]}$$

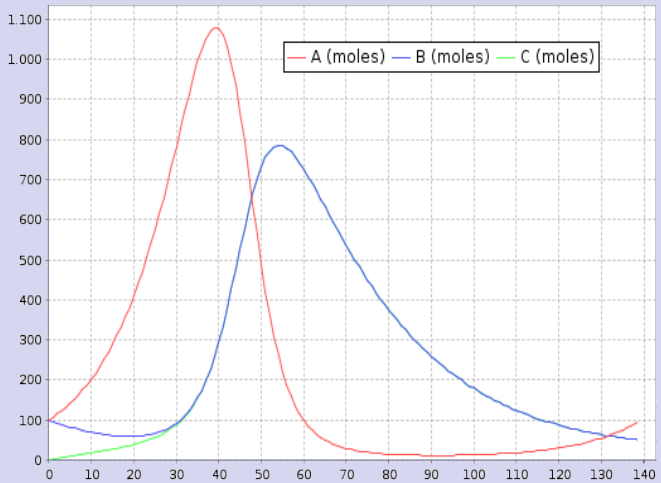
$$Z[i+1] = Z[i] + \Delta[i]$$

where $Z[i] = (A[i]; B[i]; C[i])$.

Main components of MP systems

Equational Metabolic Algorithm

MP simulation



EMA

More "algebrically", the vector of substance variation

$$\Delta[i] = (\Delta_A[i]; \Delta_B[i]; \Delta_C[i])$$

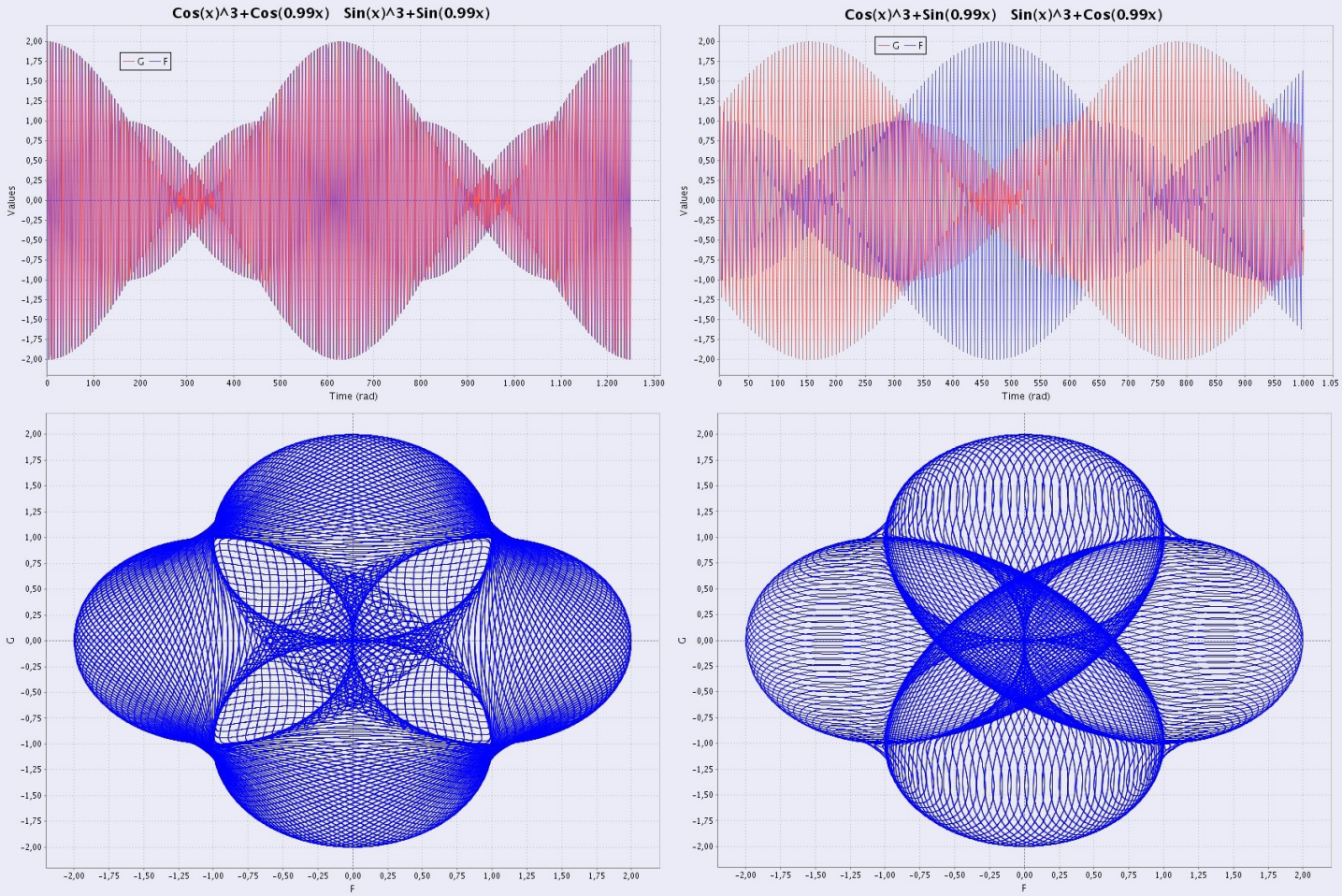
is given by the following matrix product:

$$\underbrace{\underbrace{\begin{pmatrix} 1 & -1 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{pmatrix}}_{\mathbb{A}} \times \underbrace{\begin{pmatrix} u_1[i] \\ u_2[i] \\ u_3[i] \\ u_4[i] \\ u_5[i] \end{pmatrix}}_{U[i]}}_{\Delta[i] = \mathbb{A} \times U[i]}$$

$$Z[i+1] = Z[i] + \Delta[i]$$

where $Z[i] = (A[i]; B[i]; C[i])$.

Some beautiful oscillation patterns which can be achieved with simple MP grammars...



[Vincenzo Manca, Luca Marchetti (2010) Metabolic approximation of real periodical functions. The Journal of Logic and Algebraic Programming 79 (2010), pag.363-373]

Some hints for rearranging ideas about MP systems...

Some good properties of MP systems...

1. Since P systems are inspired by the structure and functioning of living cells, MP systems are natively convenient as a modelling framework of biological systems.
2. MP systems are based on a discrete and deterministic evolution strategy which permit the calculation of the dynamics in a very simple way.

What is missing???

Of course MP systems are useful only if they are equipped with a procedure which permit the definition of new models starting from real observations of a phenomenon!

Some hints for rearranging ideas about MP systems...

Some good properties of MP systems...

1. Since P systems are inspired by the structure and functioning of living cells, MP systems are natively convenient as a modelling framework of biological systems.
2. MP systems are based on a discrete and deterministic evolution strategy which permit the calculation of the dynamics in a very simple way.

What is missing???

Of course MP systems are useful only if they are equipped with a procedure which permit the definition of new models starting from real observations of a phenomenon!

Center for Biomedical Computation

Some hints for rearranging ideas about MP systems...

Some good properties of MP systems...

1. Since P systems are inspired by the structure and functioning of living cells, MP systems are natively convenient as a modelling framework of biological systems.
2. MP systems are based on a discrete and deterministic evolution strategy which permit the calculation of the dynamics in a very simple way.

What is missing???

We need to find a nice way to solve the

INVERSE DYNAMICAL PROBLEM

The inverse dynamical problem

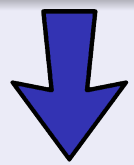
The dynamical problem

What it is given:

1. an MP grammar:
   - stoichiometry;
   - flux maps;
2. an initial state.

What we want:

THE DYNAMICS CALCULATION



EMA
Equational Metabolic Algorithm

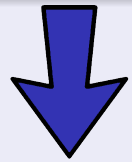
The inverse dynamical problem

What it is given:

1. a time-series of observations (i.e. a sampled dynamics);
2. an idea of stoichiometry.

What we want:

THE MP SYSTEM WHICH REPRODUCES THE OBSERVED DYNAMICS



The inverse dynamical problem

The dynamical problem

What it is given:

1. an MP grammar:
   - stoichiometry;
   - flux maps;
2. an initial state.

What we want:

THE DYNAMICS CALCULATION

EMA
Equational Metabolic Algorithm

The inverse dynamical problem

What it is given:

1. a time-series of observations (i.e. a sampled dynamics);
2. an idea of stoichiometry.

What we want:

THE MP SYSTEM WHICH REPRODUCES THE OBSERVED DYNAMICS

LGSS
Log-Gain Stoichiometric Step-wise regression

How LGSS works...

Hypothesis

Let's suppose to want to study a new phenomenon:. . .

- The set of substances involved are given by the particular phenomenon we want to describe.
- The stoichiometry can be deduced by a basic knowledge of the phenomenon.
- The dynamics of the phenomenon can be obtained by means of some analysis in laboratory which can provide some temporal series of global states.

$$(Z[i]|i = 1, 2, \ldots, t)$$

What we have

Substances involved + basic stoichiometry + the dynamics

How LGSS works...

Hypothesis

Let's suppose to want to study a new phenomenon:. . .

- The set of substances involved are given by the particular phenomenon we want to describe.
- The stoichiometry can be deduced by a basic knowledge of the phenomenon.
- The dynamics of the phenomenon can be obtained by means of some analysis in laboratory which can provide some temporal series of global states.

$$(Z[i]|i = 1, 2, \ldots, t)$$

What we have

Substances involved + basic stoichiometry + the dynamics

Center for
Biomedical Computation

How LGSS works...

Hypothesis

Let's suppose to want to study a new phenomenon:. . .

- The set of substances involved are given by the particular phenomenon we want to describe.
- The stoichiometry can be deduced by a basic knowledge of the phenomenon.
- The dynamics of the phenomenon can be obtained by means of some analysis in laboratory which can provide some temporal series of global states.

$$(Z[i]|i = 1, 2, \ldots, t)$$

What we have

Substances involved + basic stoichiometry + the dynamics

How LGSS works...

Hypothesis

Let's suppose to want to study a new phenomenon:. . .

- The set of substances involved are given by the particular phenomenon we want to describe.
- The stoichiometry can be deduced by a basic knowledge of the phenomenon.
- The dynamics of the phenomenon can be obtained by means of some analysis in laboratory which can provide some temporal series of global states.

$$(Z[i]|i = 1, 2, \ldots, t)$$

What we have

Substances involved + basic stoichiometry + the dynamics

How LGSS works...

Hypothesis

Let's suppose to want to study a new phenomenon:. . .

- The set of substances involved are given by the particular phenomenon we want to describe.
- The stoichiometry can be deduced by a basic knowledge of the phenomenon.
- The dynamics of the phenomenon can be obtained by means of some analysis in laboratory which can provide some temporal series of global states.

$$(Z[i]|i = 1, 2, \ldots, t)$$

What is missing?

We need to calculate the right flux maps

How LGSS works...

Hypothesis

Let's suppose to want to study a new phenomenon:. . .

- The set of substances involved are given by the particular phenomenon we want to describe.
- The stoichiometry can be deduced by a basic knowledge of the phenomenon.
- The dynamics of the phenomenon can be obtained by means of some analysis in laboratory which can provide some temporal series of global states.

$$(Z[i]|i = 1, 2, \ldots, t)$$

What is missing?

LGSS will do this for us!

How LGSS works...

1) Starting from temporal series of global states

$$(Z[i]|i = 1, 2, \ldots, t),$$

2) we can write the substance variation vector

$$Z[i+1] - Z[i] = \Delta[i].$$

3) Assuming n substances and m reactions, we can invert the EMA equation by writing the following system called *ADA* (Avogadro and Dalton Action):

$$\mathbb{A} \times U[i] = \Delta[i]$$

of n equations and m unknowns (the m components of the flux vector $U[i]$) which can be written in the following form:

$$\mathbb{A} \times \begin{pmatrix} \varphi_1(Z[i]) \\ \varphi_2(Z[i]) \\ \ldots \\ \varphi_m(Z[i]) \end{pmatrix} = \Delta[i].$$

How LGSS works...

4)The previous system can be expanded by calculating the matrix product:

$$\begin{aligned}
\Delta_1[i] &= \mathbb{A}(1,1)\varphi_1(Z[i]) + \mathbb{A}(1,2)\varphi_2(Z[i]) + \ldots + \mathbb{A}(1,m)\varphi_m(Z[i]) \\
\Delta_2[i] &= \mathbb{A}(2,1)\varphi_1(Z[i]) + \mathbb{A}(2,2)\varphi_2(Z[i]) + \ldots + \mathbb{A}(2,m)\varphi_m(Z[i]) \\
\ldots\ldots &= \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
\Delta_n[i] &= \mathbb{A}(n,1)\varphi_1(Z[i]) + \mathbb{A}(n,2)\varphi_2(Z[i]) + \ldots + \mathbb{A}(n,m)\varphi_m(Z[i]).
\end{aligned}$$

5) and again, by considering the t elements of our temporal series:

$$\begin{aligned}
\Delta_1[1] &= \mathbb{A}(1,1)\varphi_1(Z[1]) + \mathbb{A}(1,2)\varphi_2(Z[1]) + \ldots + \mathbb{A}(1,m)\varphi_m(Z[1]) \\
\ldots\ldots &= \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
\Delta_1[t] &= \mathbb{A}(1,1)\varphi_1(Z[t]) + \mathbb{A}(1,2)\varphi_2(Z[t]) + \ldots + \mathbb{A}(1,m)\varphi_m(Z[t]) \\
\\
\ldots\ldots &= \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
\\
\Delta_n[1] &= \mathbb{A}(n,1)\varphi_1(Z[i]) + \mathbb{A}(n,2)\varphi_2(Z[i]) + \ldots + \mathbb{A}(n,m)\varphi_m(Z[i]) \\
\ldots\ldots &= \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
\Delta_n[t] &= \mathbb{A}(n,1)\varphi_1(Z[t]) + \mathbb{A}(n,2)\varphi_2(Z[t]) + \ldots + \mathbb{A}(n,m)\varphi_m(Z[t]).
\end{aligned}$$

How LGSS works...

6) Now we can assume that fluxes can be obtained as linear combination of d basic functions

$$g_1, g_2, \ldots, g_d$$

which we call regressors, constructed over substance quantities and parameters.

$$\begin{aligned}
\varphi_1(Z) &= c_{1,1}g_1(Z) + c_{1,2}g_2(Z) + \ldots + c_{1,d}g_d(Z) \\
\varphi_2(Z) &= c_{2,1}g_1(Z) + c_{2,2}g_2(Z) + \ldots + c_{2,d}g_d(Z) \\
\varphi_3(Z) &= c_{3,1}g_1(Z) + c_{3,2}g_2(Z) + \ldots + c_{3,d}g_d(Z) \\
\ldots\ldots &= \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
\varphi_m(Z) &= c_{3,1}g_1(Z) + c_{3,2}g_2(Z) + \ldots + c_{m,d}g_d(Z).
\end{aligned}$$

Then, we can substitute these representations of regulators in the expanded *ADA* system.

How LGSS works...

7) This last substitution gives the following system:

$$\left.\begin{array}{l} \begin{pmatrix} \Delta_1[1] \\ \ldots\ldots \\ \Delta_1[t] \end{pmatrix} = \overbrace{\sum_{l=1}^{m}\sum_{j=1}^{d} c_{l,j}\mathbb{A}(1,l)}^{m\cdot d \text{ unknowns}} \begin{pmatrix} g_j(Z[1]) \\ \ldots\ldots\ldots \\ g_j(Z[t]) \end{pmatrix} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \begin{pmatrix} \Delta_n[1] \\ \ldots\ldots \\ \Delta_n[t] \end{pmatrix} = \underbrace{\sum_{l=1}^{m}\sum_{j=1}^{d} c_{l,j}\mathbb{A}(n,l)}_{m\cdot d \text{ unknowns}} \begin{pmatrix} g_j(Z[1]) \\ \ldots\ldots\ldots \\ g_j(Z[t]) \end{pmatrix} \end{array}\right\} n\cdot t \text{ equations}$$

If our temporal series are long enough,

$$m \cdot d < n \cdot t$$

and the system can be solved by means of a Least Square Estimation according to a stepwise procedure for reaching better approximations.

How LGSS works...

The stepwise procedure of LGSS will solve the previous system automatically by means of a suitable stepwise regression technique which:

1. considers all the information which can be achieved about the phenomenon under examination such as:
   - a sorting among the regressors based on a new formulation of the well-tested log-gain principle apt to preserve the allometry of the system

     [Vincenzo Manca (2010) Metabolic P systems. Scholarpedia, 5(3):9273]
     [L.von Bertalanffy (1967) General Systems Theory: Foundations, Developments, Applications. George Braziller Inc., New York]

     The log-gain principle is a discrete formulation of a general principle which has to be ensured in a biological dynamics: *"the relative variation of a systemic variable has to be proportional to the relative variation of any variable which it depends on"*.
2. tests different combinations of regressors for each flux;
3. selects the best approximations by using a suitable statistical test based on the Fischer test F;
4. gives the flux maps: $\varphi_l = \sum_{j=1}^{d} c_{l,j} g_j,\ l = 1, 2, \ldots, m$.

How LGSS works...

The stepwise procedure of LGSS will solve the previous system automatically by means of a suitable stepwise regression technique which:

1. considers all the information which can be achieved about the phenomenon under examination such as:
   - a sorting among the regressors based on a new formulation of the well-tested log-gain principle apt to preserve the allometry of the system

     [Vincenzo Manca (2010) Metabolic P systems. Scholarpedia, 5(3):9273]
     [L.von Bertalanffy (1967) General Systems Theory: Foundations, Developments, Applications. George Braziller Inc., New York]

     The log-gain principle is a discrete formulation of a general principle which has to be ensured in a biological dynamics: *"the relative variation of a systemic variable has to be proportional to the relative variation of any variable which it depends on".*
2. tests different combinations of regressors for each flux;
3. selects the best approximations by using a suitable statistical test based on the Fischer test F;
4. gives the flux maps: $\varphi_l = \sum_{j=1}^{d} c_{l,j} g_j,\ l = 1, 2, \ldots, m.$

How LGSS works...

The stepwise procedure of LGSS will solve the previous system automatically by means of a suitable stepwise regression technique which:

1. considers all the information which can be achieved about the phenomenon under examination such as:
   - a sorting among the regressors based on a new formulation of the well-tested log-gain principle apt to preserve the allometry of the system

     [Vincenzo Manca (2010) Metabolic P systems. Scholarpedia, 5(3):9273]
     [L.von Bertalanffy (1967) General Systems Theory: Foundations, Developments, Applications. George Braziller Inc., New York]
   - information about the biological meanings of the regressors which we are using (for example we can force the utilization of a regressor if we know that it has an important meaning...)
2. tests different combinations of regressors for each flux;
3. selects the best approximations by using a suitable statistical test based on the Fischer test F;
4. gives the flux maps: $\varphi_l = \sum_{j=1}^{d} c_{l,j} g_j$, $l = 1, 2, \ldots, m$.

How LGSS works...

The stepwise procedure of LGSS will solve the previous system automatically by means of a suitable stepwise regression technique which:

1. considers all the information which can be achieved about the phenomenon under examination such as:
   - a sorting among the regressors based on a new formulation of the well-tested log-gain principle apt to preserve the allometry of the system

     [Vincenzo Manca (2010) Metabolic P systems. Scholarpedia, 5(3):9273]
     [L.von Bertalanffy (1967) General Systems Theory: Foundations, Developments, Applications. George Braziller Inc., New York]
   - information about the biological meanings of the regressors which we are using (for example we can force the utilization of a regressor if we know that it has an important meaning...)
2. tests different combinations of regressors for each flux;
3. selects the best approximations by using a suitable statistical test based on the Fischer test F;
4. gives the flux maps: $\varphi_l = \sum_{j=1}^{d} c_{l,j} g_j$, $l = 1, 2, \ldots, m$.

How LGSS works...

The stepwise procedure of LGSS will solve the previous system automatically by means of a suitable stepwise regression technique which:

1. considers all the information which can be achieved about the phenomenon under examination such as:
   - a sorting among the regressors based on a new formulation of the well-tested log-gain principle apt to preserve the allometry of the system

     [Vincenzo Manca (2010) Metabolic P systems. Scholarpedia, 5(3):9273]
     [L.von Bertalanffy (1967) General Systems Theory: Foundations, Developments, Applications. George Braziller Inc., New York]
   - information about the biological meanings of the regressors which we are using (for example we can force the utilization of a regressor if we know that it has an important meaning...)
2. tests different combinations of regressors for each flux;
3. selects the best approximations by using a suitable statistical test based on the Fischer test F;
4. gives the flux maps: $\varphi_l = \sum_{j=1}^{d} c_{l,j} g_j$, $l = 1, 2, \ldots, m$.

How LGSS works...

The stepwise procedure of LGSS will solve the previous system automatically by means of a suitable stepwise regression technique which:

1. considers all the information which can be achieved about the phenomenon under examination such as:
   - a sorting among the regressors based on a new formulation of the well-tested log-gain principle apt to preserve the allometry of the system

     [Vincenzo Manca (2010) Metabolic P systems. Scholarpedia, 5(3):9273]
     [L.von Bertalanffy (1967) General Systems Theory: Foundations, Developments, Applications. George Braziller Inc., New York]
   - information about the biological meanings of the regressors which we are using (for example we can force the utilization of a regressor if we know that it has an important meaning...)
2. tests different combinations of regressors for each flux;
3. selects the best approximations by using a suitable statistical test based on the Fischer test F;
4. gives the flux maps: $\varphi_l = \sum_{j=1}^{d} c_{l,j} g_j$, $l = 1, 2, \ldots, m$.

Rearranging ideas about LGSS...

Log-Gain Stoichiometric Stepwise regression (LGSS)

[Vincenzo Manca, Luca Marchetti (2010) Log-Gain Stoichiometric Stepwise regression for MP systems. International Journal of Foundations of Computer Science, to appear]

1. It is a completely automatic regression technique which permits to discover the flux maps for an MP system by considering its stoichiometry and the time series of its dynamics.
2. It is based on a stepwise regression algorithm.
3. It automatically uses the stoichiometry of the biological phenomenon during the regression.
4. It considers a new formulation of the Log-gain principle which permits to preserve the principle of allometry.
5. It has been implemented as a set of MATLAB functions.
6. IT RUNS VERY FAST!!!

Rearranging ideas about LGSS...

Log-Gain Stoichiometric Stepwise regression (LGSS)

[Vincenzo Manca, Luca Marchetti (2010) Log-Gain Stoichiometric Stepwise regression for MP systems. International Journal of Foundations of Computer Science, to appear]

1. It is a completely automatic regression technique which permits to discover the flux maps for an MP system by considering its stoichiometry and the time series of its dynamics.
2. It is based on a stepwise regression algorithm.
3. It automatically uses the stoichiometry of the biological phenomenon during the regression.
4. It considers a new formulation of the Log-gain principle which permits to preserve the principle of allometry.
5. It has been implemented as a set of MATLAB functions.
6. IT RUNS VERY FAST!!!

Rearranging ideas about LGSS...

Log-Gain Stoichiometric Stepwise regression (LGSS)

[Vincenzo Manca, Luca Marchetti (2010) Log-Gain Stoichiometric Stepwise regression for MP systems. International Journal of Foundations of Computer Science, to appear]

1. It is a completely automatic regression technique which permits to discover the flux maps for an MP system by considering its stoichiometry and the time series of its dynamics.
2. It is based on a stepwise regression algorithm.
3. It automatically uses the stoichiometry of the biological phenomenon during the regression.
4. It considers a new formulation of the Log-gain principle which permits to preserve the principle of allometry.
5. It has been implemented as a set of MATLAB functions.
6. IT RUNS VERY FAST!!!

Center for Biomedical Computation

Rearranging ideas about LGSS...

Log-Gain Stoichiometric Stepwise regression (LGSS)

[Vincenzo Manca, Luca Marchetti (2010) Log-Gain Stoichiometric Stepwise regression for MP systems. International Journal of Foundations of Computer Science, to appear]

1. It is a completely automatic regression technique which permits to discover the flux maps for an MP system by considering its stoichiometry and the time series of its dynamics.
2. It is based on a stepwise regression algorithm.
3. It automatically uses the stoichiometry of the biological phenomenon during the regression.
4. It considers a new formulation of the Log-gain principle which permits to preserve the principle of allometry.
5. It has been implemented as a set of MATLAB functions.
6. IT RUNS VERY FAST!!!

Rearranging ideas about LGSS...

Log-Gain Stoichiometric Stepwise regression (LGSS)

[Vincenzo Manca, Luca Marchetti (2010) Log-Gain Stoichiometric Stepwise regression for MP systems. International Journal of Foundations of Computer Science, to appear]

1. It is a completely automatic regression technique which permits to discover the flux maps for an MP system by considering its stoichiometry and the time series of its dynamics.
2. It is based on a stepwise regression algorithm.
3. It automatically uses the stoichiometry of the biological phenomenon during the regression.
4. It considers a new formulation of the Log-gain principle which permits to preserve the principle of allometry.
5. It has been implemented as a set of MATLAB functions.
6. IT RUNS VERY FAST!!!

Center for Biomedical Computation

Rearranging ideas about LGSS...

Log-Gain Stoichiometric Stepwise regression (LGSS)

[Vincenzo Manca, Luca Marchetti (2010) Log-Gain Stoichiometric Stepwise regression for MP systems. International Journal of Foundations of Computer Science, to appear]

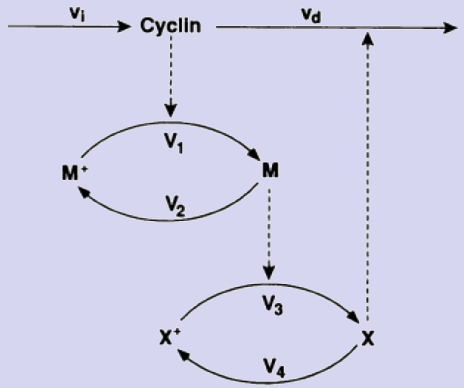
1. It is a completely automatic regression technique which permits to discover the flux maps for an MP system by considering its stoichiometry and the time series of its dynamics.
2. It is based on a stepwise regression algorithm.
3. It automatically uses the stoichiometry of the biological phenomenon during the regression.
4. It considers a new formulation of the Log-gain principle which permits to preserve the principle of allometry.
5. It has been implemented as a set of MATLAB functions.
6. IT RUNS VERY FAST!!!

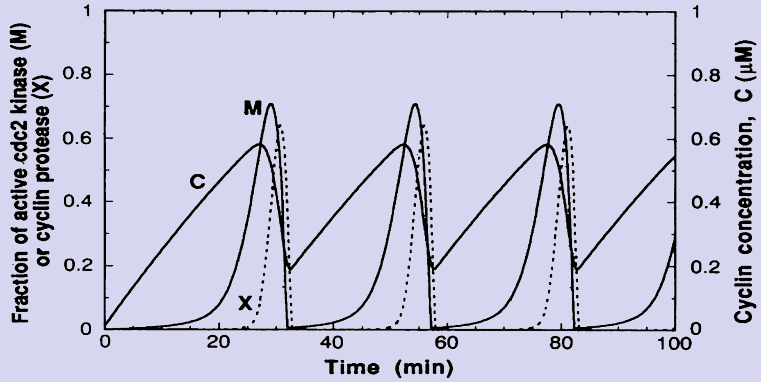
Coming back to the topic of the submitted paper...

The idea

Application of LGSS to get MP models which reproduce the dynamics of the Goldbeter's mitotic oscillator at different time sampling

The Goldbeter's ODE system about mitotic oscillations in early amphibian embryos





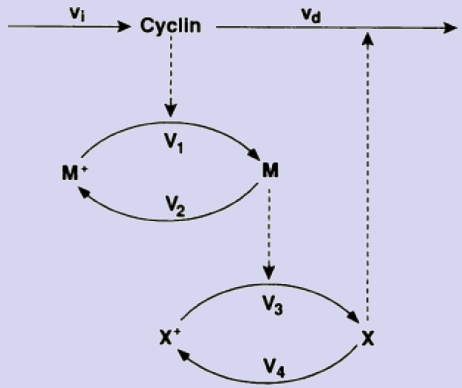
[A. Goldbeter (1991) A minimal cascade model for the mitotic oscillator involving cyclin and cdc2 kinase. PNAS 88(20), 91079111]

Coming back to the topic of the submitted paper...

The idea

Application of LGSS to get MP models which reproduce the dynamics of the Goldbeter's mitotic oscillator at different time sampling

The Goldbeter's ODE system about mitotic oscillations in early amphibian embryos



$$\frac{dC}{dt} = v_i - v_d X \frac{C}{K_d + C} - k_d C$$
$$\frac{dM}{dt} = V_{M1} \frac{C}{K_c + C} \frac{(1-M)}{K_1 + (1-M)} - V_2 \frac{M}{K_2 + M}$$
$$\frac{dX}{dt} = M V_{M3} \frac{(1-X)}{K_3 + (1-X)} - V_4 \frac{X}{K_4 + X}$$

[A. Goldbeter (1991) A minimal cascade model for the mitotic oscillator involving cyclin and cdc2 kinase. PNAS 88(20), 91079111]

The goal of our analysis

The Goldbeter's ODE system is a very famous model, but the differential equations are very complex and it is hard to understand the regulative role of each substance involved in the phenomenon.

Is it possible to do better with MP systems???

The answer is YES!

By applying the LGSS algorithm we have generated automatically 700 models which reproduces the dynamics of the oscillator at different time grains from 0.06 sec. to 42 sec. with increments of 0.06 seconds.

Before the introduction of LGSS this analysis was really impossible... we needed one month to define by hands only one model!

The goal of our analysis

The Goldbeter's ODE system is a very famous model, but the differential equations are very complex and it is hard to understand the regulative role of each substance involved in the phenomenon.

Is it possible to do better with MP systems???

The answer is YES!

By applying the LGSS algorithm we have generated automatically 700 models which reproduces the dynamics of the oscillator at different time grains from 0.06 sec. to 42 sec. with increments of 0.06 seconds.

Before the introduction of LGSS this analysis was really impossible... we needed one month to define by hands only one model!

The goal of our analysis

The Goldbeter's ODE system is a very famous model, but the differential equations are very complex and it is hard to understand the regulative role of each substance involved in the phenomenon.

Is it possible to do better with MP systems???

The answer is YES!

By applying the LGSS algorithm we have generated automatically 700 models which reproduces the dynamics of the oscillator at different time grains from 0.06 sec. to 42 sec. with increments of 0.06 seconds.

Before the introduction of LGSS this analysis was really impossible... we needed one month to define by hands only one model!

Center for Biomedical Computation

The goal of our analysis

The Goldbeter's ODE system is a very famous model, but the differential equations are very complex and it is hard to understand the regulative role of each substance involved in the phenomenon.

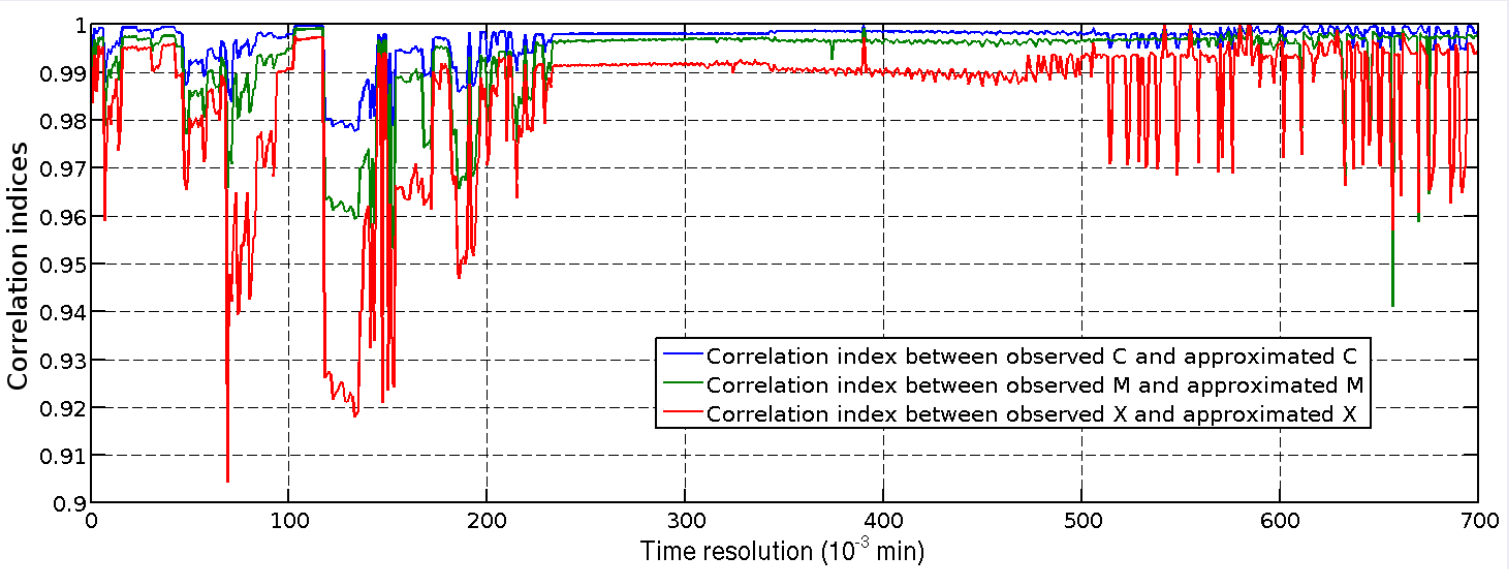
Is it possible to do better with MP systems???

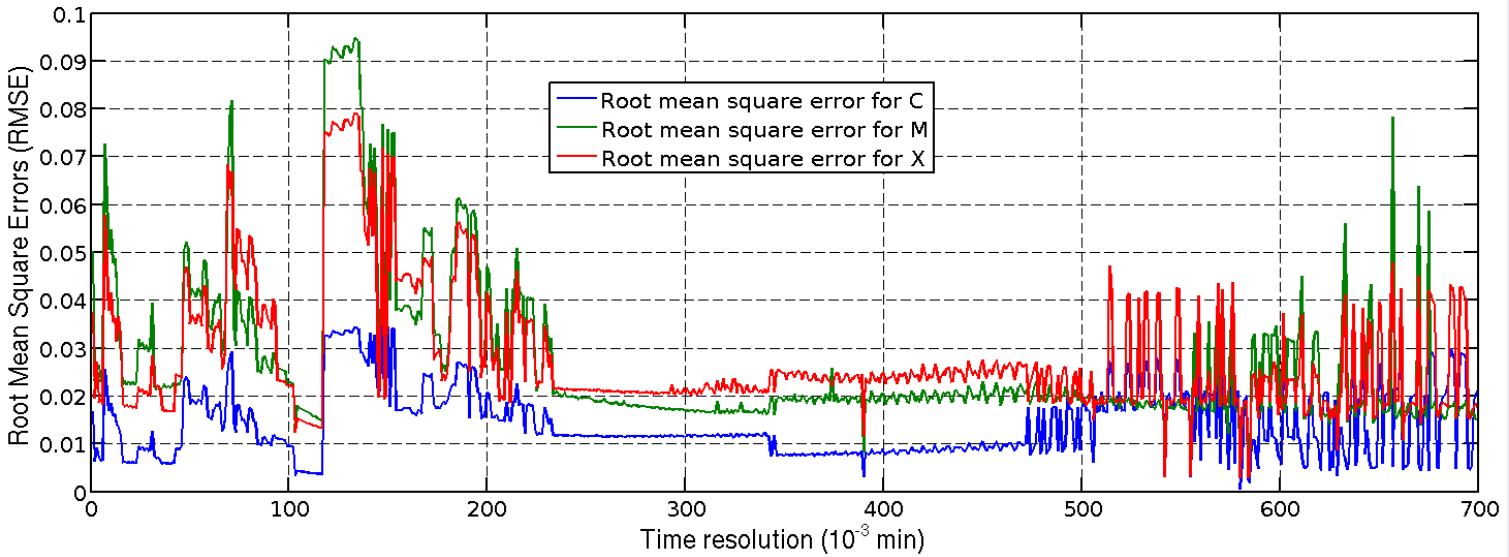
The answer is YES!

By applying the LGSS algorithm we have generated automatically 700 models which reproduces the dynamics of the oscillator at different time grains from 0.06 sec. to 42 sec. with increments of 0.06 seconds.

Before the introduction of LGSS this analysis was really impossible... we needed one month to define by hands only one model!

Correlation indices and RMSE of our MP models





Classifications of MP mitotic models

The flux maps calculated by the LGSS are obtained starting from a dictionary of 20 possible regressors, that is monomials of C, M and X with degree less than or equal to 3.

Ex: C, M, X, C^2, M^2, X^2, CM, CX, MX, C^3, M^3, X^3, C^2M, CM^2, $\ldots$, CMX.

After a deep analysis of the 700 generated MP models we found that, by considering different values of the resolution time, different analytical forms of flux maps were appropriate. Each of them:

- permits the calculation of a dynamics which is highly correlated with the observed one;
- is based on a formula much more simple than the one proposed by Goldbeter;
- gives a clear idea of the regulative role of each substance involved in the phenomenon at the different time-grains.

Classifications of MP mitotic models

The flux maps calculated by the LGSS are obtained starting from a dictionary of 20 possible regressors, that is monomials of C, M and X with degree less than or equal to 3.

After a deep analysis of the 700 generated MP models we found that, by considering different values of the resolution time, different analytical forms of flux maps were appropriate. Each of them:

- permits the calculation of a dynamics which is highly correlated with the observed one;
- is based on a formula much more simple than the one proposed by Goldbeter;
- gives a clear idea of the regulative role of each substance involved in the phenomenon at the different time grains.

Classifications of MP mitotic models

The flux maps calculated by the LGSS are obtained starting from a dictionary of 20 possible regressors, that is monomials of C, M and X with degree less than or equal to 3.

After a deep analysis of the 700 generated MP models we found that, by considering different values of the resolution time, different analytical forms of flux maps were appropriate. Each of them:

- permits the calculation of a dynamics which is highly correlated with the observed one;
- is based on a formula much more simple than the one proposed by Goldbeter;
- gives a clear idea of the regulative role of each substance involved in the phenomenon at the different time grains.

Classifications of MP mitotic models

The flux maps calculated by the LGSS are obtained starting from a dictionary of 20 possible regressors, that is monomials of C, M and X with degree less than or equal to 3.

After a deep analysis of the 700 generated MP models we found that, by considering different values of the resolution time, different analytical forms of flux maps were appropriate. Each of them:

- permits the calculation of a dynamics which is highly correlated with the observed one;
- is based on a formula much more simple than the one proposed by Goldbeter;
- gives a clear idea of the regulative role of each substance involved in the phenomenon at the different time grains.

Classifications of MP mitotic models

The flux maps calculated by the LGSS are obtained starting from a dictionary of 20 possible regressors, that is monomials of C, M and X with degree less than or equal to 3.

After a deep analysis of the 700 generated MP models we found that, by considering different values of the resolution time, different analytical forms of flux maps were appropriate. Each of them:

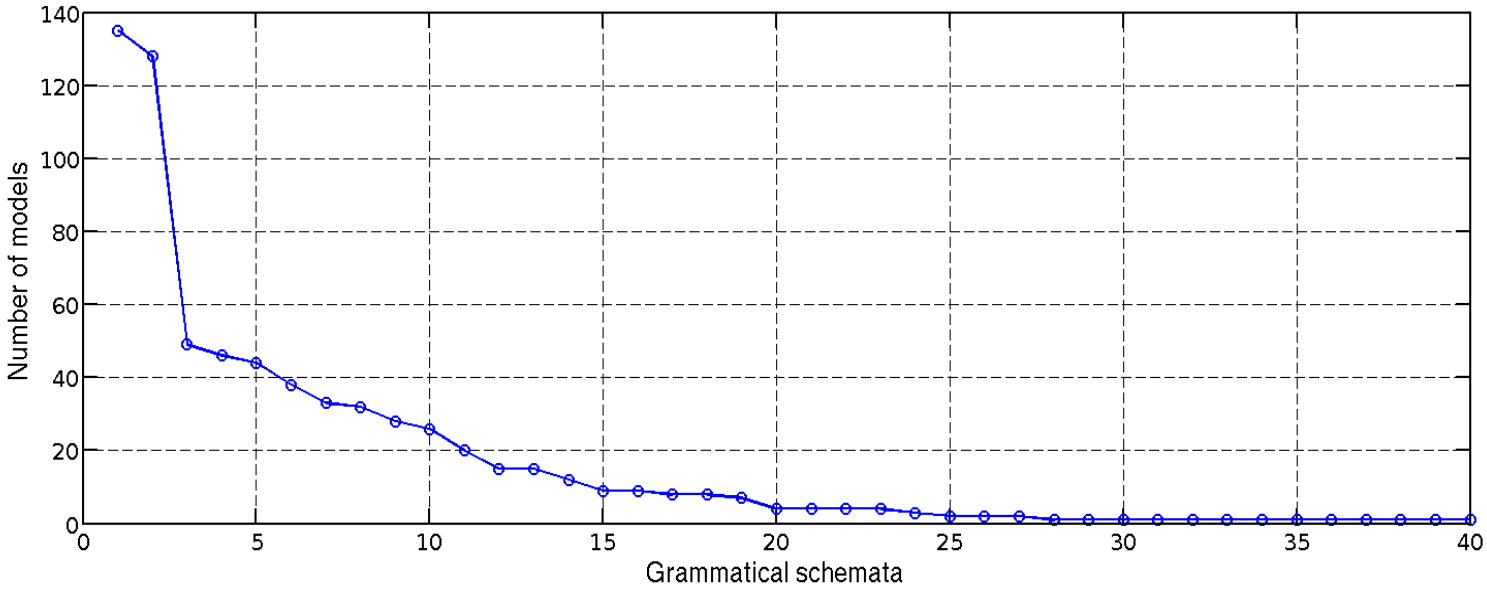
- permits the calculation of a dynamics which is highly correlated with the observed one;
- is based on a formula much more simple than the one proposed by Goldbeter;
- gives a clear idea of the regulative role of each substance involved in the phenomenon at the different time grains.

Classifications of MP mitotic models

The analytical forms of flux maps can be classified in only 40 grammatical schemata which provide the best approximation error for different interval of time sampling...

Grammatical schemata	number of models	number of regressors	total n. of monomials	τ interval (10^{-3} *min*)	best τ (10^{-3} *min*)	best RMSE
1	135	6	16	151 – 345	315	$1.61 \cdot 10^{-2}$
2	128	6	17	343 – 477	401	$1.62 \cdot 10^{-2}$
3	49	6	17	43 – 93	43	$1.84 \cdot 10^{-2}$
4	46	6	16	138 – 232	219	$1.95 \cdot 10^{-2}$
5	44	8	24	1 – 71	40	$1.48 \cdot 10^{-2}$
6	38	6	16	525 – 699	683	$1.78 \cdot 10^{-2}$
7	33	6	16	473 – 563	556	$1.79 \cdot 10^{-2}$
8	32	5	15	514 – 694	602	$2.78 \cdot 10^{-2}$
9	28	7	16	570 – 696	671	$1.09 \cdot 10^{-2}$
10	26	6	16	493 – 684	684	$1.8 \cdot 10^{-2}$

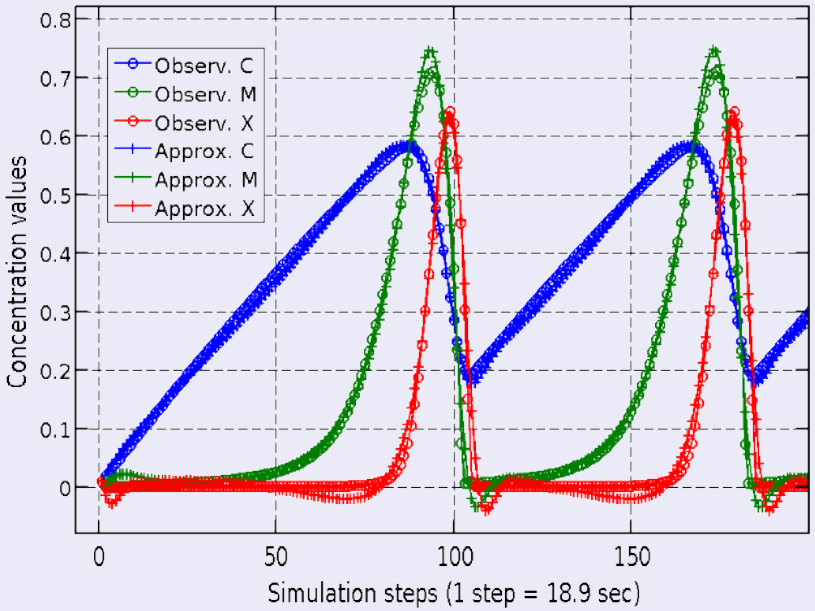
Classifications of MP mitotic models

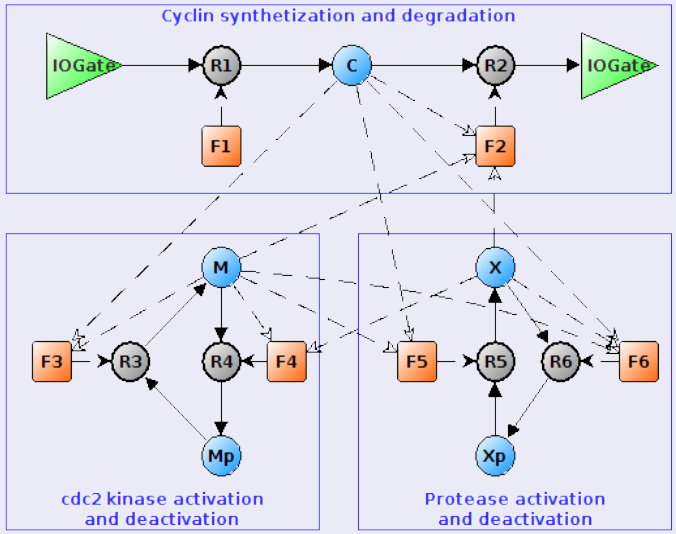


Grammatical schemata	number of models	number of regressors	total n. of monomials	τ interval (10^{-3} *min*)	best τ (10^{-3} *min*)	best RMSE
1	135	6	16	151 – 345	315	$1.61 \cdot 10^{-2}$
2	128	6	17	343 – 477	401	$1.62 \cdot 10^{-2}$
3	49	6	17	43 – 93	43	$1.84 \cdot 10^{-2}$
4	46	6	16	138 – 232	219	$1.95 \cdot 10^{-2}$
5	44	8	24	1 – 71	40	$1.48 \cdot 10^{-2}$
6	38	6	16	525 – 699	683	$1.78 \cdot 10^{-2}$
7	33	6	16	473 – 563	556	$1.79 \cdot 10^{-2}$
8	32	5	15	514 – 694	602	$2.78 \cdot 10^{-2}$
9	28	7	16	570 – 696	671	$1.09 \cdot 10^{-2}$
10	26	6	16	493 – 684	684	$1.8 \cdot 10^{-2}$

An example of MP mitotic model

$$
\begin{aligned}
r_1 &: \emptyset \to C & \varphi_1 &= k_1 + k_2\ C^2 + k_3\ CM \\
r_2 &: C \to \emptyset & \varphi_2 &= k_4\ C + k_5\ M + k_6\ X \\
r_3 &: M^+ \to M & \varphi_3 &= k_7 + k_8\ CM \\
r_4 &: M \to M^+ & \varphi_4 &= k_9\ M + k_{10}\ X \\
r_5 &: X^+ \to X & \varphi_5 &= k_{11}\ C + k_{12}\ M \\
r_6 &: X \to X^+ & \varphi_6 &= k_{13} + k_{14}\ X + k_{15}\ C^2 + k_{16}\ CM
\end{aligned}
$$





Constants and initial values: $k_1 = 0.01$, $k_2 = 0.03327$, $k_3 = 0.0485192$, $k_4 = 0.0168923$, $k_5 = 0.0428226$, $k_6 = 0.054506$, $k_7 = 0.00245843$, $k_8 = 0.540636$, $k_9 = 0.219284$, $k_{10} = 0.14129$, $k_{11} = 0.308615$, $k_{12} = 1.01307$, $k_{13} = 0.0338141$, $k_{14} = 0.468994$, $k_{15} = 0.756053$, $k_{16} = 1.15991$, $C[0] = M[0] = X[0] = 0.01$, $M^+[0] = X^+[0] = 0.99$.

Conclusions and future work...

In this presentation:

- we have introduced the MP systems framework and its evolution algorithm;
- we have explained in its main features the LGSS regression algorithm;
- we have discussed an application of LGSS to the modelling of a real biological dynamics pointing out its points of strength...

Future work:

- theoretical analysis of LGSS in order to evaluate more its regression power;
- application of LGSS to the modelling of real complex biological phenomena.

Conclusions and future work...

In this presentation:

- we have introduced the MP systems framework and its evolution algorithm;
- we have explained in its main features the LGSS regression algorithm;
- we have discussed an application of LGSS to the modelling of a real biological dynamics pointing out its points of strength...

Future work:

- theoretical analysis of LGSS in order to enhance even more its regression power;
- application of LGSS to the modelling of very complex real biological phenomena.

Conclusions and future work...

In this presentation:

- we have introduced the MP systems framework and its evolution algorithm;
- we have explained in its main features the LGSS regression algorithm;
- we have discussed an application of LGSS to the modelling of a real biological dynamics pointing out its points of strength...

Future work:

- theoretical analysis of LGSS in order to enhance even more its regression power;
- application of LGSS to the modelling of very complex real biological phenomena.

Conclusions and future work...

In this presentation:

- we have introduced the MP systems framework and its evolution algorithm;
- we have explained in its main features the LGSS regression algorithm;
- we have discussed an application of LGSS to the modelling of a real biological dynamics pointing out its points of strength...

Future work:

- theoretical analysis of LGSS in order to enhance even more its regression power;
- application of LGSS to the modelling of very complex real biological phenomena.

Conclusions and future work...

In this presentation:

- we have introduced the MP systems framework and its evolution algorithm;
- we have explained in its main features the LGSS regression algorithm;
- we have discussed an application of LGSS to the modelling of a real biological dynamics pointing out its points of strength...

Future work:

- theoretical analysis of LGSS in order to enhance even more its regression power;
- application of LGSS to the modelling of very complex real biological phenomena.

Conclusions and future work...

Thank you!!!