Special session S2

Research questions in state transition models of biomolecular dynamics

Giuditta Franco
University of Verona, Department of Computer Science
Strada Le Grazie, 15, I-37134 Verona
franco@sci.univr.it

Vincenzo Manca
University of Verona, Department of Computer Science
Strada Le Grazie, 15, I-37134 Verona
vincenzo.manca@univr.it

Giuseppe Scollo
University of Catania, Department of Mathematics and Computer Science
Viale A. Doria, 6, I-95125 Catania
scollo@dmi.unict.it

1 Introduction

Recent developments of discrete models to analyse biological processes motivate the revisitation of typical concepts of classical dynamics in a completely discrete context, such as that provided by cellular automata [16], or, more generally, by state transition systems [11,15].

Dynamics of discrete systems, besides being instrumental to the analysis of metabolic processes [1,2,8,9], in the context of P systems [12,13,14,10,4], in general prove most natural in the representation of biomolecular dynamics, where the symbolic entities which come into play are easily amenable to strings or related structures, ranging from multisets to dynamical networks [3,17].

This work aims at identifying a few relevant problems outstanding in biomolecular computing [5], and at outlining research directions and strategies for their solution. The next two sections, which give a contrived summary of basic concepts and relevant research results, form the background for the subsequent outlook into the possible future of this research.

2 State transition dynamics

The classical approach to study dynamical systems is focused on differential equations, that impose local (infinitesimal) relations on quantity variations from which, under suitable hypotheses, one can analytically reconstruct the global dynamical behaviour of the system. In [11], we addressed the problem of considering, in general terms, dynamical systems that are completely discrete in that, not only are the time instants natural or integer numbers, but also the space is a discrete entity. We characterized transitions and space states of a discrete dynamical system by means of *state transition dynamics*, defined as a pair $\langle S, q\rangle$, where S is a set of states and q is a binary relation on S, the *transition relation* on states.

A state transition dynamics is *eternal* if the transition relation q is total. We assume this to be always the case in the systems of our interest, that is, the set of *final* states $S\backslash\mathsf{dom}(q)$ is empty. Every dynamics may be easily extended to an eternal one by turning each final state into a fixed point of the transition relation, that is by adding a transition from that state to itself in the extended dynamics.

As in [11], we call *quasistate* any subset of S. With the relation algebra notation from [15], the *successor* quasistate of state x under q is written $\mathsf{img}(x\,;\,q)$. This is generalized to denote the successor of quasistate x under q, defined as the union of the successor quasistates of states in x. The *orbit* of origin x, or x-orbit, results from iteration of this concept, viz. it is defined as the infinite sequence of quasistates $(x_i \mid i \in \mathbb{N})$ such that $x_0 = x$ (or $x_0 = \{x\}$, if x is an individual state) and $x_{i+1} = \mathsf{img}(x_i\,;\,q)$. The orbit is *periodic* if $\exists n > 0 : x_n = x_0$, while it is *eventually periodic* if, for some $k \geq 0$, it evolves into a periodic one after a k-step *transient*, that is, if $\exists n > 0 : x_{k+n} = x_k$. A *basin* B is a nonempty quasistate that is closed under q, that is, $\mathsf{img}(B\,;\,q) \subseteq B$. Every orbit gives rise to a basin, by taking the union of the quasistates it consists of.

For any state x in S, a *trajectory* of origin x, or x-trajectory, is a function $\xi : \mathbb{N} \to S$ such that, with subscript argument, $\xi_0 = x$ and $\xi_{n+1} \in \mathsf{img}(\xi_n\,;\,q)$. $\xi_{\mathbb{N}}$ denotes the image of this function. An x-*flight* is an injective x-trajectory.

Any x-trajectory thus runs "inside" the x-orbit. The difference between these two concepts is apparent, but it practically disappears for *deterministic* dynamics, viz. those where the transition relation is actually a function—hence orbits are singleton sequences, and one may say *the* x-trajectory, for any state x. In [11,15] the general case of nondeterministic dynamics is considered, whereby classical dynamics concepts such as *attracting set*, *attractor*, and *recurrence* come in two distinct modal flavours. While we refer the interested reader to the cited work for details about the formulations of these concepts, and related results, in the general, nondeterministic case, in the rest of this paper we shall confine ourselves to the deterministic case, where modal differences disappear and the aforementioned concepts take the following shapes.

A state x is *recurrent* if it occurs at least twice (hence infinitely often) in the x-trajectory. An *attracting set* A of a basin B is a nonempty subset of B such that, for every state x in B, the x-trajectory eventually gets into it; more precisely, $\forall x \in B\ \exists k \in \mathbb{N} : \forall n \geq k\ \mathsf{img}(x\,;\, q^n) \subseteq A$, where q^n denotes the n-fold iterated composition of q. An attracting set of B that is minimal under set inclusion is called the *attractor* of B. We may say *the* attractor since this is unique, for a given basin B, whenever it exists[11]. The image of an x-flight is a simple example of a basin that has no attractor. The deterministic-case corollary of a result in [15] states that the attractor of a basin B exists iff B has no flights, and that in this case the attractor of B consists of the set of recurrent states in B.

3 Metabolic P systems

P systems are a computational model based on the *compartmentalization of the workspace* and on *multiset rewriting*. These basic concepts have clear biological counterparts in the role that membranes play in biological organisms, and in the biochemical basis of any biological entity, respectively. We refer to [12,13,14] for the definitions of P systems structure in its most important variants, and for the usual strategies which determine their evolution in time.

However, from a biological viewpoint, the evolution strategies considered in standard P systems do not seem to be fully adequate. Indeed, in these systems, state transitions are usually determined by a maximally parallel application of multiset rewriting rules, but maximal parallelism or other strategies of this kind cannot express the dynamics of a population of chemicals governed by biochemistry laws, *e.g.* in its *regulatory* aspects, which involve a dynamical *change of strategy*, depending on the system state. *Metabolic P systems* [1,2,8,9] address this question by equipping each multiset rewriting rule, also called *metabolic rule*, with a *reaction map*, which is a real-valued function of the system state that yields the "competitive strength" of that rule at the given state.

Metabolic rules thus compete for the allocation of biochemical resources, whose types occur in the rule pattern. Allocation of resources, which are available in finite amounts, is governed by a *mass partition principle*, where reaction maps determine the relative allocation factors. Each metabolic rule acts on a system state by consuming/producing amounts of biochemical resources that are integer multiples of a *reaction unit* for that rule in the given state, thus generalizing the notion of molar unit (Avogradro's principle).

In the notationally simpler case of a system with no workspace partitioning, *i.e.* with only one (outer) membrane, states are functions of type $T \to \mathbb{N}$, where T is the finite alphabet of biochemical resource types, and a deterministic dynamics on such states is computed, for a given set R of metabolic rules on T equipped

with reaction maps $F = (f_r \mid r \in R)$, by the *P metabolic algorithm* (PMA) [8], which formalizes the resource allocation strategy outlined above. A generalization of the PMA to P systems with inner membranes is feasible, as well as to such systems with possible resource flow through membranes and dynamical change of membrane structure, but at the cost of increasing notational complexity.

4 Research outlook

While the investigation into state transition dynamics carried out in [11,15] is motivated by the wish to consider dynamical systems where not only time but also the state space is discrete, it is to be noted that actually no assumption about state space structure is made in the definitions and results obtained in that work. As a matter of fact, one may easily specialize state transition dynamics by making additional assumptions about state space structure, whether discrete or continuous as it proves convenient. Thus, for example, even the simple case of metabolic P systems without inner membranes may get a richer state space structure just by taking states to be real-valued functions on the type alphabet, if continuous values prove more convenient to express mass quantities. Furthermore, regardless of the discrete or continuous nature of the functions' codomain, the state space of metabolic P systems is easily endowed with a metric, viz. it is a finite-dimensional vector space on that codomain (since the type alphabet is finite).

The situation outlined above raises interesting questions relating to further analysis of state transition dynamics for state spaces endowed with topological structure, such as the following ones: 1) Definability of weaker notions of recurrence, where one would replace *exact* recurrent occurrence of the given state in its trajectory with *approximate* occurrence, viz. recurrence of seeing the trajectory get across a sufficiently (or arbitrarily) small neighbourhood of the origin. 2) Definability of weaker notions of attraction, where, similarly to the previous question, exact inclusion of trajectories or orbits in the attracting set is replaced by approximate inclusion. 3) Which (if any) of the possible answers to the previous questions deliver a straightforward generalization of the characterization results, linking recurrence and attractors, already obtained in [15] for structureless state transition dynamics?

A wholly different class of questions arises from a valuable feature underlying the very design of metabolic P systems. Here, two distinct levels of description of biochemical dynamics are readily recognized: the *stoichiometry* level, represented by the metabolic rules, and the *regulation* level, represented by the reaction maps assigned to them. Bidirectional interaction between the two levels is inherent, since state transitions are determined according to the given stoichiometry, but with relative strength determined by the regulation, while, conversely, regulation

depends on the state. A first, basic question is: For a given dynamical *behaviour* of a biochemical system, how can one find the metabolic rules and reaction maps that best reproduce that behaviour? This may be called the general *simulation problem* for metabolic P systems. A somewhat "easier" (but: how much *really* easier?) subproblem of the general one results from assuming that a solution for the stoichiometry level is available somehow, but nothing is known about the reaction maps. This is thus the general *regulation problem* for metabolic P systems. Finally (for the time being), one may happen to know everything about the stoichiometry of the given behaviour, and *something* about plausible reaction maps (*e.g.* their polynomial form), but without knowing the values of specific parameters (*e.g.* polynomial coefficients) which determine those maps uniquely. This is the *tuning problem*, a subproblem of the regulation problem as it were.

Different strategies for the solution of the simulation problem are exemplified in [9], where the mitotic oscillator model proposed by Goldbeter [6,7] is taken as a case study. This model proves adequate to account for the simplest form of mitotic oscillations, as found in early amphibian embryos. Goldbeter's model consists of an autonomous system of ordinary differential equations. Three different metabolic P system models are developed in [9] for the case under study. Two of them are directly derived from Goldbeter's model under different "readings" of the differential equations in terms of metabolic rules and reactions maps; this clearly tells that, in general, the "translation" of a differential model into a metabolic P model is not unique. Furthermore, the third P model is directly synthesized from the biochemical description of the phenomenon under study; it exhibits the desired oscillatory behaviour as well as the two other P models do, and it is even a simpler model.

REFERENCES

1. L. Bianco, F. Fontana, G. Franco and V. Manca, P Systems for Biological Dynamics, in: G. Ciobanu, Gh. Păun, M.J. Perez-Jimenez (Eds.), *Applications of Membrane Computing*, Springer, Natural Computing series (2006) 81–126.

2. L. Bianco, F. Fontana and V. Manca, P Systems with reaction maps, *Int'l J. of Foundations of Computer Science*, vol. 17:1, 2006, 27–48.

3. C. Bonanno and V. Manca, Discrete dynamics in biological models, in: Gh. Păun, C. Calude (Eds.), *Romanian Journal of Information Science and Technology*, vol. 1-2:5, 2002, 45–67.

4. F. Fontana, L. Bianco and V. Manca, P systems and the modeling of biochemical oscillations, in: R. Freund, Gh. Păun, G. Rozenberg, M. Yung, A. Salomaa (Eds.), *WMC 2005*, Springer, LNCS 3850 (2006) 200–209.

5. G. Franco, *Biomolecular Computing – Combinatorial Algorithms and Laboratory Experiments*, Doctoral dissertation, University of Verona (2006).

6. A. Goldbeter, A Minimal Cascade Model for the Mitotic Oscillator Involving Cyclin and cdc2 Kinase, *PNAS*, vol. 88:20, 1991, 9107–9111.

7. A. Goldbeter, *Biochemical Oscillations and Cellular Rythms. The molecular bases of periodic and chaotic behaviour*. Cambridge University Press, New York, USA, 2004.

8. V. Manca, Topics and Problems in Metabolic P Systems, Fourth Brainstorming on Membrane Computing, Sevilla, 2006.

9. V. Manca and L. Bianco, Biological Networks in Metabolic P Systems, University of Verona, Dep't of Computer Science, submitted (2006).

10. V. Manca, L. Bianco and F. Fontana, Evolutions and oscillations of P systems: Theoretical considerations and applications to biochemical phenomena, in: G. Mauri, Gh. Păun, M.J. Pérez-Jiménez, G. Rozenberg, A. Salomaa (Eds.), *Membrane Computing*, Springer, LNCS 3365, (2005) 63–84.

11. V. Manca, G. Franco, G. Scollo, State transition dynamics: basic concepts and molecular computing perspectives, in: M. Gheorghe, M. Holcombe (Eds.), *Molecular Computational Models: Unconventional Approaches*, Idea Group, Hershey, PA, USA (2005) 32–55.

12. Gh. Păun, Computing with membranes, *J. Comput. System Sci.*, vol. 61:1, 2000, 108–143.

13. Gh. Păun, *Membrane Computing. An Introduction*, Springer, Berlin, Germany, 2002.

14. P Systems Web Page, `http://psystems.disco.unimib.it` .

15. G. Scollo, G. Franco, V. Manca, A Relational View of Recurrence and Attractors in State Transition Dynamics, March 2006, submitted.

16. S. Wolfram, *Theory and Application of Cellular Automata*, (1986) Addison-Wesley.

17. A. Wuensche, Basins of Attraction in Network Dynamics: A Conceptual Framework for Biomolecular Networks, in: G. Schlosser and G.P. Wagner (Eds.), *Modularity in Development and Evolution*, (2002), Chicago University Press.