Cell Communication in Tissue P Systems and Cell Division in Population P Systems

Francesco BERNARDINI, Marian GHEORGHE

Department of Computer Science
The University of Sheffield
Regent Court, Portobello Street, Sheffield, S1 4DP, UK
E-mail: {F.Bernardini,M.Gheorghe}@dcs.shef.ac.uk

Abstract. Two classes of tissue P systems based on evolution communication rules are introduced, some results are proved, but many more are listed as further research problems. A framework to develop population P systems is defined and a number of variants formulated with a strong biological motivation.

1 Introduction

P systems represent a class of distributed and parallel computing devices of a biological type that was introduced in [15]. Several variants of this model have been investigated and the literature on the subject is now rapidly growing. The main results in this area show that P systems are a very powerful and efficient computational model [16], [18], [13].

The main ingredient of a every P system is a membrane structure defined as a hierarchical arrangement of different membranes embedded in a unique main membrane, which identify several distinct regions inside the system. Each region gets assigned a finite multiset of objects and a finite set of rules for modifying the objects or moving them from a place to another one. The structure of a P system is usually represented as a tree. Tissue P systems has been then proposed as a variant of membrane systems where the structure of the system is defined as an arbitrary graph. Nodes in the graph represent membranes that are able to communicate objects alongside the edges of the graph [16].

From a biological point of view, tissue P systems can be interpreted as a straightforward model of cell behaviour in multicellular organisms. In such organisms, cells are specialized members of a multicellular community. They collaborate with each other to form a multitude of different tissues, arranged into organs performing varied functions. These important features of cell behaviour in tissues are shared by biological system at various levels. In general, they may be regarded as composed from many individually components that cooperate in a coherent way by interacting each other. This population of individuals is usually far from being stable: mechanisms enabling new individuals to be introduced, and mechanisms causing the removal of some individuals play a fundamental role in the evolution of a biological system as a community of interacting/cooperating components.

The present paper deals with both aspects of cell communication in tissue P systems and aspects of cell proliferation in population P systems. In section 2 we consider a variant of tissue P systems that is inspired by the general mechanism of cell communication that is based on signal molecules and receptor proteins. Section 3 presents a general framework for population P systems, a class of P systems relying on cell division with a dynamic structure of the underlying graph. Many variants are then derived from this framework and research topic formulated.

All over the paper, the reader is supposed to be familiar with the basic knowledge of membrane computing [16]. We just adopt a slight different terminology: the term membrane is replaced by the term cell to denote the basic functional unit of a tissue P system or a population P system. Furthermore, we will use the notion of matrix grammars with or without appearance checking and the well-known hierarchy $CF \subset MAT \subset MAT_{ac} = RE$ [8], where CF, MAT, MAT_{ac}, and RE denote the family of context-free, matrix, matrix with appearance checking, and recursively enumerable languages, respectively.

2 Cell Communication

Mechanisms enabling one cell to influence the behaviour of another one play a fundamental role in multicellular organisms where cells have to be able to coordinate their own behaviour for the benefit of the organism as a whole. These communication mechanisms depend heavily on extracellular *signal molecules*, which are produced by a cell to signal their neighbours or cells that are further away. Most of these signal molecules are secreted from the signaling cell into the extracellular space by diffusion through the plasma membrane. Some signals remain instead tightly bound to signaling cell surface, and they are able to influence cells that are in direct contact with the signaling cell. Cells can respond to external signaling by means of some proteins called *receptors*; receptors are able to recognise external signals by binding signal molecules to cell surface. Each receptor specifically binds a signal molecule (or a class of signal molecules) and initiates a specific response inside the target cell. In multicellular organisms, cells of different types are designed to respond to different classes of signal molecules and have got assigned different classes of receptor proteins. When a receptor binds a signal molecule, they form a single active unit that generates a cascade of intra-cellular signals, that eventually alter a target protein in a way that changes the behaviour of the cell [2].

From a P system point of view, cell communication based on signals and receptors represents an interesting combination of communication (a signal object is produced inside a cell and reaches another cell identified as target cell) and evolution (the target cell responds by altering its internal status). This leads to two reasonable interpretations for the cell communication mechanism in the context of P systems: *signals* and *receptors* as a mechanism to move objects from a cell to another one, or signals and receptors as a mechanism to trigger particular transformations inside a cell once a signal has been recognised by means of some receptor. In a sense, this latter interpretation can be related to two common notions in membrane computing: the notion of catalyst (the transformation of an object is mediated by a specific catalyst object) and the notion of promoter (the application of some rules inside a region depends on the presence of particular objects called promoters) (see [16]).

In the rest of this section, we want to explore both these possibilities by firstly considering an evolution-communication model where communication is mediated by special

receptor objects, and then a variant of tissue P systems where signals are used for activating different sets of rules inside a cell.

2.1 Evolution-Communication Model

Evolution-communication P systems represent a class of P systems where objects are transformed and communicated by means of two separate mechanisms, which are usually expressed in terms of a finite set of transformation rules and a finite set of communication rules. The term evolution-communication is mainly used to distinguish these systems from purely communicative P systems (the objects can only be moved from one place to another one without any chance of being modified) and from the basic model of membrane systems (where communication is a consequence of transformation rules) [16]. Two main variants of evolution-communication P systems have been considered so far: P systems with boundary rules [5], and evolution-communication P systems with symport/antiport as communication rules [6].

We introduce here a variant of tissue P systems where transformation rules are usual rewriting rules on multisets of objects whereas, communication among cells is mediated by special receptor objects. A receptor is represented as an object $\bar{x}$ where x is a finite multiset of objects that stands for the multiset recognised by that receptor . This notation is used to specify that a receptor $\bar{x}$ acts as a single unit with respect to the objects in x, which are therefore received in lumps. Communication among cells is then achieved by allowing a multiset x to move from a cell to another one if and only if, the selected target cell contains a receptor $\bar{x}$. This target cell is chosen in a non-deterministic way among those that are in direct contact with the cell that contains the multiset x. Furthermore, we assume that a receptor is immediately destroyed after having been used for some communication. New receptors can be created by using transformation rules that, besides standard objects, are able to produce receptors as well. An example is pictorially described in Figure 1 that shows how communication among cells is achieved by means of receptors.

From a biological point of view, these features of receptors reflect the fact that there is a continuous turnover of molecules at cell surface level, and in many cases, the binding of a signal molecule to a receptor modifies the receptor conformation, which need then to be replaced in order to initiate a new response.

Remark 2.1. The communication model based on receptors we are considering here has evident similarities with some of the notions already considered in the area of membrane computing: symport/antiport rules with promoters [16], boundary rules [5], and the so-called one-way communication model mainly considered in the context of P automata [7], [9]. In those variants, one can in fact express communication rules of the form $x\,[_i\,y \rightarrow [_i\,xy$ where x, y are arbitrary multisets of arbitrary sizes. Rules of this form specify that a multiset x can enter a membrane i if and only if, membrane i contains a multiset y; that is, the multiset y acts as a receptor for the multiset x. In this respect, one can easily reconsider boundary rules in the context of tissue P systems by adopting for instance communication rules of the form $x\,]_j\,[_i\,y \rightarrow]_j\,[_i\,xy$. This is to say that a multiset x can be moved from a cell j to a cell i if and only if, cell i contains a multiset y. Notice that in this case, communication is no longer one-way as the mechanism turns to be symmetric (i.e., objects can enter a cell and can exit a cell as well). However, the mechanism based on receptors is more restrictive in what concerns the following aspects:

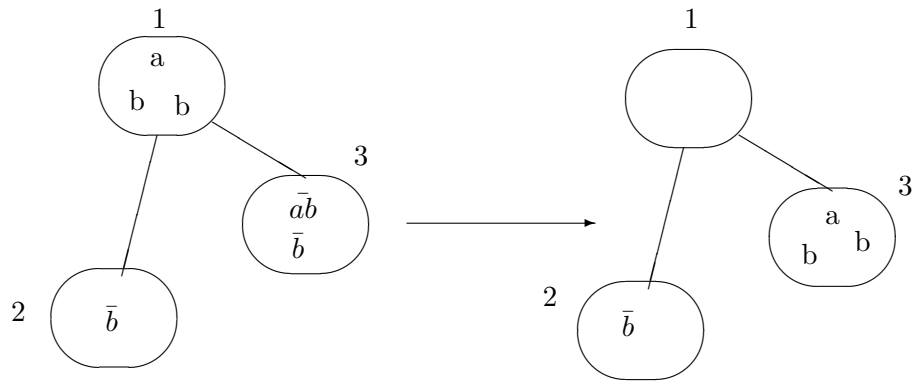

Figure 1: *The multiset ab is moved into cell* 3 *from cell* 1 *by means of the receptor* $\bar{ab}$*; the object b is moved into cell* 3 *from cell* 1 *by means of the receptor* $\bar{b}$*. The receptors in cell* 3 *are destroyed whereas, a receptor* $\bar{b}$ *remains in cell* 2 *as it has not been used in any communication.*

- a receptor is meant to be very specific for a particular multiset of objects: it can be used only in communications that involve that particular multiset;
- a receptor is immediately destroyed after having been used for some communication;
- the target cell is chosen in a non-deterministic way.

In other words, we are considering communication rules of the form

$$x\,]_i\,[\,\bar{x}\,\rightarrow]_i\,[\,x, \tag{1}$$

where: the receptor $\bar{x}$ is consumed by the rule; the receptor $\bar{x}$ can only appear in communication rules like (1) specifically coupled with the multiset x and vice versa; a second label is not specified as the target cell is non-deterministically chosen.

We are now ready to give a formal definition for tissue P systems with receptors.

Definition 2.1. A tissue P system with receptors is a construct

$$\mathcal{T} = (V, \gamma, C_1, C_2, \ldots, C_n, c_O)$$

where:

1. V is a finite alphabet of symbols called objects;
2. $\gamma = (\{1, 2, \ldots n\}, A)$, is a finite undirected graph that defines the structure of the system where $A \subseteq \{\{i, j\} \,|\, 1 \leq i \neq j \leq n\}$;
3. $C_i = (w_i, \rho_i, R_i)$, for each $1 \leq i \leq n$, with:
   - $w_i \in V^*$ a finite multiset of objects,
   - $\rho_i \in \{\,\bar{u} \,|\, u \in V^*\,\}^*$ a finite multiset of receptors,
   - R_i is a finite set of evolution rules of the forms $x \rightarrow yz$ for $z \in \{\,\bar{u} \,|\, u \in V^*\,\}^*$, $x, y \in V^*$, and (x, out), with $x \in V^*$;

4. c_O is the output cell.

A tissue P system with receptors $\mathcal{T}$ is defined by a finite undirected graph γ where each node is labeled by a value in $\{1, 2, \ldots, n\}$ and corresponds in one-to-one manner to a cell C_i. The edges in the graph γ define the bonds that exist among the cells and they define the structure of the P system $\mathcal{T}$. An edge is represented as an unordered pair $\{i, j\}$, with $1 \leq i \neq j \leq n$.

Each cell C_i represents a basic functional unit of the system that is characterised by: a finite multiset of objects w_i that defines the initial content of the cell C_i, a finite multiset of receptors ρ_i that are initially assigned to the cell C_i, and a finite set of evolution rules R_i.

Each set of rules R_i contains evolutions rules of the form $x \to yz$ where x and y are finite multisets of objects whereas z is a finite multiset of receptors. A rule of this form specifies that a multiset x can be consumed inside a cell i in order to produce a new multiset of objects y and a new multiset of receptors z. Moreover, we consider communication rules of the form (x, out) that allow a cell to release a multiset x in the environment. This is mainly to simplify the operation of removing objects from the system. In fact, objects that are sent out from the cells are considered as being lost because, as usual in evolution-communication P systems, we do not provide any specific notion of environment.

A computation is then defined in the usual way by saying that, by starting from the initial configuration, a tissue P system with receptors evolves by applying the rules in a non-deterministic maximal parallel manner. Specifically, in each step of a computation, in each cell i:

- a multiset of objects x can be either transformed by a rule R_i or communicated to a cell j by means of some receptor associated with that cell j; a multiset x can be moved from a cell i to a cell j non-deterministically chosen if and only if, there exists an edge $\{i, j\}$ in γ and cell j contains a receptor $\bar{x}$;

- all the receptors that have been used for some communication purposes are removed from cell i, and all the new receptors produced by rules in R_i are added to cell i;

Notice that here the number of objects that can be communicated in a single step of communication is bounded by the overall size of the receptors that are currently assigned to the cells. In other words, receptors act as communication channels of finite capacity that can be used just once in a single step of computation and then they need to be replaced.

A configuration of a tissue P system $\mathcal{T}$ is defined as a tuple $\Sigma = (C'_1, C'_2, \ldots, C'_n)$, with $C'_i = (w'_i, \rho'_i)$ the current status of the cell i in terms of objects and receptors. A computation in $\mathcal{T}$ is said to be successful if it reaches a configuration Σ_f where no more rules can be applied and no more receptors can be used. The configuration Σ_f is called a halting one. The result of a successful computation is the Parikh vector associated with the multiset of objects that is contained in the output cell c_O in the halting configuration. The set of Parikh vectors generated by all the successful computations in $\mathcal{T}$ is denoted by $Ps(\mathcal{T})$.

We denote by $PsOTP_{n,p}(e, rec_r, out_h)$, for $e \in \{nCoo, Coo\}$, $n, p, r, h \geq 1$ the family of sets of Parikh vectors that are generated by tissue P systems with at most n cells and where, in each step of any computation, in each cell, the number of receptors is less than or equal to p. Furthermore, e specifies the type of evolution rules that are used (cooperative or non-cooperative), r specifies the maximal size for the receptors that are used inside the cells, and h specifies the maximal size for the multisets that appear in output rules. The

parameter p has been specifically introduced in order to bound the number of receptors that can be assigned to the cell-surfaces as, for each cell, this surface is supposed to be of a finite capacity. On the other hand, here we do not care about the number of different types of receptors that are used in the cells; this might represent a further research topic. As well as this, we are not concerned with the form of the graph that defines the structure of the system.

The first natural problem we have to look at regards the generative capacity of tissue P systems with a bounded number of cells and a bounded number of receptors. Here we provide some partial answers that result from this preliminary investigation.

We know from the result obtained for the basic model of P systems [16] that P systems with cooperative evolution rules and one membrane are computationally complete. This result can be immediately transferred to tissue P systems with receptors.

Theorem 2.1. $PsOTP_{1,1}(Coo, rec_1, out_1) = PsRE$.

More interesting is the case of systems with non-cooperative rules.

Theorem 2.2. $PsMAT \subseteq PsOTP_{3,*}(nCoo, rec_1, out_1)$

Proof. Consider a matrix grammar without appearance checking in binary normal form where all the matrices of the form $(X \to Y, A \to x_1x_2)$, with $x_1, x_2 \in N_2 \cup T \cup \{\lambda\}$, $Y \in N_1$, are labeled in a one-to-one manner with values 1 to k for some $k \geq 1$. Here, N_1, N_2 denote two separate alphabets of non-terminal symbols whereas, T denotes an alphabet of terminal symbols.

We construct a tissue P system $\mathcal{T}$ that contains 3 cells where: $C_1 = (XA, \lambda)$, for $(S \to XA)$ the initial matrix of the matrix grammar to be simulated, $C_2 = (\$, \lambda)$, for $\$$ a new special symbols, and $C_3 = (\lambda, \bar{f})$, for f another new special symbol. The structure of the system $\mathcal{T}$ is then given by the finite undirected graph $\gamma = (\{1, 2, 3\}, \{\{1, 2\}, \{1, 3\}\})$.

Then, for each matrix $i : (X \to Y, A \to x_1x_2)$, $1 \leq i \leq k$, and for each $1 \leq j \leq k$, we consider the following sets of rules associated with the cells:

$$
\begin{array}{lll}
R_1: & X \to p_i\bar{q}_i, & \\
& p_j \to \#, & \\
& \# \to \#, & \\
& q_j \to \bar{\dagger}_j, & \\
& \dagger_i \to x_1'x_2'Y'\$ & \\
& s' \to s & \text{for each } s \in (N_1 \cup N_2 \cup T). \\
R_2: & \$ \to \bar{p}_j, & \\
& p_i \to \dagger_i\bar{A}\bar{\$}, & \\
& \dagger_j \to \dagger_j, & \\
& A \to \bar{q}_j. &
\end{array}
$$

The simulation of a matrix $i : (X \to Y, A \to x_1x_2)$ is performed in the following way. In cell 1, the symbol X is replaced by an object p_i and a receptor $\bar{q}_i$. At the same time, the object $\$$ is replaced in cell 2 by a receptor $\bar{p}_j$. Now, if $i \neq j$ then we are forced to apply in cell 1 the rule $p_j \to \#$, which introduces the symbol $\#$ that generates an infinite computation. Otherwise, if $i = j$, we can move the object p_i from cell 1 to cell 2; here, the object p_i is replaced by an object $\dagger_i$, a receptor $\bar{A}$, and a receptor $\bar{\$}$. Next, if cell 1 does not contain any object A then the simulation of the matrix i stops but the computation does not halt because of the rule $\dagger_i \to \dagger_i$. Otherwise, an object A is moved from cell 1 to

cell 2, where it is replaced by an object q_j. Yet again, if $i \neq j$ then we obtain an infinite computation for the rule $\dagger_i \to \dagger_i$. If we have instead produced an object q_i then, this can be moved from cell 2 to cell 1 by using the receptor $\bar{q}_i$. Here we obtain a receptor $\bar{\dagger}_i$ that is used to remove the object $\dagger_i$ from cell 2. The simulation of the matrix i is completed in two further steps of computation: first we use the rule $\dagger_i \to x'_1 x'_2 Y'\$$ in cell 1, then we move the object \$ from cell 1 to cell 2 and we use a rule $s' \to s$ for each symbol s' that is present in cell 1.

In a similar way, we simulate matrices of the form $(X \to \lambda, A \to x_1 x_2)$, with $x_1, x_2 \in T$ that are used to finish a derivation in the grammar. Specifically, this is done by considering in cell 1 a rule $\dagger_i \to f$ instead of a rule $\dagger_i \to x'_1 x'_2 Y'\$$. The object f once produced is immediately moved from cell 1 to cell 3; this cell in fact contains by definition a receptor $\bar{f}$. In cell 3, we consider the rule $f \to \bar{A}_1 \dots \bar{A}_r$, for $N_2 = \{A_1, \dots, A_r\}$, and the rules $A \to A$, for each $A \in N_2$. These are used to check whether cell 1 still contains some non-terminal symbols or not. If that is the case then, the computation halts by having correctly simulated a derivation in the matrix grammar.

Notice that the number of receptors cannot be bounded because, in order to make sure that the computation correctly halts, we need to produce in cell 3 a number of receptors that depends on the size of the alphabet N_2 in the matrix grammar to be simulated. □

This result shows that tissue P systems with receptors of size 1 and non-cooperative rules are at least as powerful as matrix grammars without appearance checking. Finding an upper bound for the power of this particular family of tissue P systems remains an open problem as well as investigating the generative capacity of systems with a bounded number of receptors. On the other hand, we expect the universality to be proved for tissue P systems that use receptor of size 2.

Conjecture 2.1. $PsOTP_{n,*}(nCoo, rec_2, out_1) = PsRE$, for some $n \geq 1$.

Obviously, the exact number of membrane has to be determined and possibly an upper bound on the number of receptors has to be found.

Remark 2.2. $[_i\, x\,]_i \to [_j\, y\,]_j$ In Remark 2.1 we have already pointed out the similarities between tissue P systems with receptors and other existing models of P systems. In this respect, one may notice an obvious link between the results presented here and the results already obtained in the area of P systems, where, in order to obtain the universality, symport rules of size 2 or promoter of size 2 are always used [16], [18], [13]. However, if Conjecture 2.1 was proved then, we would get an universality result for tissue P systems with rules of size 2 even in this restricted framework based on the notion of receptors as specified in Remark 2.1. On the other hand, one can always think of getting rid of receptors and considering generalised boundary rules of the form $x\,]_i\,[\,y \to]_i\,[\,xy$ as indicated in Remark 2.1 (i.e., a multiset x can be moved from a cell i to another one non-deterministically chosen among those that contact cell i and contain a multiset y). It would be then interesting to investigate the power of these communication rules especially in the minimal case where $|x| = 1$, and $|y| = 1$ for each communication rule in the system.

2.2 Rules Activated by Signals

Let us reverse the perspective and look at signals as special objects that are produced inside a cell to influence the behaviour of its neighbouring cells. We have already mentioned

possible analogies between the signaling mechanism and the well-known notion of promoters: the possibility of using a rule inside a membrane heavily depends on the presence of particular objects called promoters. In a similar way, signal molecules can be considered as special objects that, after having been produced inside a cell, are immediately sent to the neighbouring cells where they are responsible for the activation of new sets of rules. From a biological point of view, this reflects the fact that when a receptor binds to signal molecules, these two molecules become a single active unit that produces a burst of chemical activity inside the cell where several other molecules get involved. Moreover, there are cases where the effect of a signal lasts for a while with the possibility of being reinforced by increasing for instance the concentration of the signal molecule.

Thus, we introduce a variant of tissue P systems with signals where the rules that can be used inside a cell at a time depend on the signals that have been recognized by the cell at a previous time. To this aim, we identify two separate alphabets: an alphabet of objects, and an alphabet of signals. Then, each cell gets assigned several sets of rules that are coupled with different receptors; this is represented by considering pairs of the form $\langle s, R\rangle$, with s a signal and R a finite set of evolution rules. This is to specify that the cell is always able to recognise a signal s, which is then responsible for the activation of the rules in R. A signal s is recognised by a cell by adding to its content an occurrence of a object $\bar{s}$, and rules in R can be used inside that cell if and only if, that cell contains at least one object $\bar{s}$ ($\bar{s}$ represents an active form of the signal s, which is able to initiate a response inside the target cell). Every time rules in R are used an occurrence of the object $\bar{s}$ is consumed. In this way, we are able to take in account of the number of occurrences of the signal s that has been produced at the same time. In fact, for each occurrence of the object s that has been produced in a cell at a time, we introduce an occurrence of the object $\bar{s}$ in each cell that is in direct contact with the signaling cell. In other words, we adopt a replicative model where an object s is replicated in the form $\bar{s}$ in all the neighbouring cells.

Formally, we give the following definition of tissue P systems with signals.

Definition 2.2. A tissue P system with signals is a construct:

$$\mathcal{T} = (V, S, \gamma, C_1, C_2, \ldots, C_n, c_O)$$

where:

1. V is a finite alphabet of symbols called objects;
2. S is a finite alphabet of symbols called signals, such that $S \cap V = \emptyset$;
3. $\gamma = (\{1, 2, \ldots n\}, A)$, is a finite undirected graph that defines the structure of the system where $A \subseteq \{\{i, j\} \mid 1 \leq i \neq j \leq n\}$;
4. $C_i = (w_i, \rho_i, \langle s_1, R_i^{(1)}\rangle, \langle s_2, R_i^{(2)}\rangle, \ldots, \langle s_k, R_i^{(k)}\rangle)$, with $k \geq 1$, for each $1 \leq i \leq n$, with:
   - $w_i \in V^*$ a finite multiset of objects,
   - $\rho_i \in \{\, \bar{s} \mid s \in V^* \,\}^*$ a finite multiset of signals that has been initially recognised by the cell i,
   - for each $1 \leq h \leq k$, $s_h \in S$, and $R_i^{(k)}$ is a finite set of evolution rules of the forms $x \to y$ for $x \in V^*$, $y \in (V \cup S)^*$, and (x, out), with $x \in V^*$;

5. c_O is the output cell.

As in definition 2.1, a tissue P system $\mathcal{T}$ with signals is defined as a collection of $n \geq 1$ cells, which are associated in a one-to-one manner to the nodes of a finite undirected graph γ that defines the structure of the system. For each $1 \leq i \leq n$, the initial configuration of each cell i is given by a multiset of objects w_i and a multiset of signals ρ_i that are supposed to have been already recognised by the cell i. Then, each cell i gets assigned $k \geq 1$ sets of evolution rules; each one of this set of rules is coupled with a specific signal that is responsible for the activation of that set of rules. Rules are transformation rules of the form $x \to y$ that allow a cell to consume a multiset of objects x in order to produce a new multiset y that contains both objects in V and signals in S. Signals in S cannot appear on the left side of any rule.

A tissue P system with signals $\mathcal{T}$ evolves by applying the rules in a non-deterministic maximal parallel manner as usual, but with the following restriction imposed by the presence of signal objects: in each step, in a cell i, for each pair $\langle s_h, R_i^{(h)} \rangle$, rules in $R_i^{(h)}$ can be used in cell i if and only if, cell i contains at least an occurrence of the object $\bar{s}$; the use of rules in $R_i^{(h)}$ always consumes an occurrence of $\bar{s}$. As a consequence of these rules, new objects and new signals can be produced inside cell i; objects that are not signals remain inside cell i whereas, signals are immediately propagated to the neighbouring cells. Specifically, in each step of a computation, for each signal $s \in S$ that is produced inside a cell i by means of some rule, an object $\bar{s}$ is introduced in each cell j that contains at least a pair $\langle s, R \rangle$ and such that there exists an edge $\{i, j\}$ in γ.

Yet again we say that the result of a successful computation (i.e., a computation that reaches a halting configuration) is the Parikh vector associated with the multiset of objects in V that is contained in the output cell c_O in the halting configuration. The set of Parikh vectors generated by all the successful computation in $\mathcal{T}$ is denoted by $Ps(\mathcal{T})$. Then, we denote by $PsOTP_{n,s,k}(f)$, with $n, s, k \geq 1$, $f \in \{Coo, nCoo\}$, the family of Parikh vectors of natural numbers generated by tissue P systems with signals where: n is the number of cells, s is the cardinality of the alphabet S, and k is the maximum number of different sets of rules that are associated to a cell.

In order to clarify the notion of tissue P system with signals, we present an example of system with 2 cells that provides a first insight into the power of this variant of P systems.

Example 2.1. Consider a tissue P system $\mathcal{T} = (V, S, \gamma, C_1, C_2, c_O)$ where:

$V = \{a, b, c\}$,
$S = \{s_1, s_2, s_3\}$,
$\gamma = (\{1, 2\}, \{\{1, 2\}\})$
$C_1 = (ab, \lambda, \langle s_1, \{a \to aa, b \to b\, s_3\}\rangle, \langle s_2, \{(b, out)\}\rangle)$,
$C_2 = (c, \bar{s_3}, \langle s_3, \{c \to c\, s_1, c \to s_2\}\rangle)$,
$c_O = 1$.

A computation in $\mathcal{T}$ gets started in cell 2 where we can use either a rule $c \to c\, s_1$ or a rule $c \to s_2$ by consuming the unique occurrence of the object $\bar{s_3}$. If we apply the rule $c \to c\, s_1$ inside cell 2 then, we obtain an object c and a signal s_1, which is immediately recognised by cell 1 that gets assigned an object $\bar{s_1}$. In this way, the system $\mathcal{T}$ reaches a configuration:

$$((ab, \bar{s_1}), (c, \lambda)).$$

At this point, we can use the rules $a \to aa$, $b \to b\, s_3$ in cell 1 that produce the configuration:

$$((aab, \lambda), (c, \bar{s_3})).$$

Yet again, if we use the rule $c \to c\, s_1$ in cell 2 then we obtain the configuration:

$$((aab, \bar{s_1}), (c, \lambda)).$$

Next, we have to reuse the rules $a \to aa$, $b \to b\, s_3$ in cell 1 by producing the configuration

$$((aaaab, \lambda), (c, \bar{s_3})).$$

Now, it is obvious that this process can be iterated for an arbitrary number of times. At any moment, we can stop a computation by using the rule $c \to s_2$ inside the cell 2, which introduces in cell 1 an object $\bar{s_2}$. In presence of the object $\bar{s_2}$, we have to apply the rule (b, out) in cell 1, and the computation halts.

Now it appears clearly that any successful computation in $\mathcal{T}$ halts in a configuration where the output cell contains a multiset of the form $(a, 2^m)$, for some $m \geq 1$.

At this point, we can investigate the generative capacity of tissue P systems with signals. We know from [16] that P systems with promoters and non-cooperative rules are computationally complete. Moreover, Example 2.1 shows that this variant of tissue P systems with non cooperative rules are quite powerful: systems with 2 cells are able to generate Parikh vectors that are not in $PsMAT$ (it is known that all one-letter languages in MAT are regular). Therefore, we expect the following conjecture to be proved:

Conjecture 2.2. $PsOTP_{2,*,*}(nCoo) = PsRE$.

Obviously, we face the problem of finding a bound for the number of different signals and a bound for the number of different sets of rules per cell. On the other hand, the universality for systems with 2 cells and cooperative rules can be easily obtained as a straightforward adaption of the universality of the basic model of P system when cooperative rules are considered [16].

3 Population P Systems: Cell Division

Cell division is the fundamental mechanism of tissue renewal in multicellular organisms. Cell division allows cells to reproduce by always generating two new cells from a single cell. These cells can then be attached to the existing part of the tissue by means of both cell movement and bond making (two cells get in touch somehow and tightly adhere each other). The importance of cell division in multicellular organisms is much more evident during the process of embryo-genesis when a functionally complete organism is developed by starting from a single cell. During this process, cells proliferate and become specialized in many different ways. Cells differentiation leads to specialized cells that will remain different thereafter and will perform distinct functions in distinct tissues.

More insights on cell behaviour in tissues are provided by the work done in [19] where a first simulation model for the activity of cells in the epidermis is proposed. In that model, various types of cells are identified that can appear in the system in different states at different stages. Typically, a population of stem cells is initially specified, which are able to produce cells that can divide only for a finite number of times producing offsprings; these cells will then divide as many times as they can and eventually they will differentiate into specialized cells that are not able to divide anymore.

Here, we want to consider for P systems some of these basic features of cell behaviour in tissues by introducing a notion of population P systems with cell division where the

structure of the system dynamically changes. In fact, cell division is defined as an operation that allows new cells to be introduced in the system, which can then form new links with the existing ones; this operation modifies the graph that defines the structure of the system by inserting new nodes and altering the sets of edges in the graph. As well as this, an operation of cell differentiation and one of terminal cell differentiation are considered, which make possible to change the type of the cells in the system by varying in this way the set of rules that can be used by a given cell.

From a more general point of view, the operations considered here can be interpreted in various ways. Cell division may be regarded as a generic operation that allows new individuals to be introduced in the current population. Cell differentiation may be considered as a change in the role of an individual in the population that turns to perform new tasks. Finally, terminal cell differentiation may be associated with an event that give some advantage to an individual that will then remain in the system in a more stable form.

Formally, we give the following definition of population P systems with cell division.

Definition 3.1. A population P system with cell division is a construct:

$$\mathcal{P} = (V, K, T, \gamma, C_1, .., C_n, R, \alpha)$$

where:

1. V is a finite alphabet of symbols called objects;
2. K is a finite alphabet of labels for the cells, which define different types of cells;
3. T is a finite alphabet of terminal labels such that, $T \cap K = \emptyset$;
4. $\gamma = (\{1, 2, \dots n\}, A)$, is a finite undirected graph that defines the structure of the system where $A \subseteq \{\{i, j\} \mid 1 \leq i \neq j \leq n\}$;
5. $C_j = (w_j, i_j)$, for each $1 \leq j \leq n$, with:
   - $w_j \in V^*$ a finite multiset of objects,
   - $i_j \in K$ a label that defines the type of the cell j,
6. R is a finite set of rules of the following forms:
   (a) $[_i\, x \to y\,]_i$, with $x \in V^*$, $y \in (V \times \{here, go, out\})^*$, and $i \in K$ (evolution rules),
   (b) $[_i\, x\,]_i \to [_i\, y\,]_i [_i\, z\,]_i$, with $x, y, z \in V^*$, and $i \in K$ (cell division rules),
   (c) $[_i\, x\,]_i \to [_j\, y\,]_j$, with $x, y \in V^*$, and $i \neq j \in K$ (cell differentiation rules),
   (d) $[_i\, x\,]_i \to [_t\, y\,]_t$, with $x, y \in V^*$, $t \in T$, and $t \in T$ (terminal cell differentiation rules);
7. $\alpha \subseteq (V^* \times K) \times (V^* \times K)$ is a binary relation that defines a notion of similarity between two cells.

A population P system $\mathcal{P}$ with cell division is defined as an initial population of cells where each cell gets assigned a finite multiset of objects and a label in K that defines its type. Each cell corresponds in a one-to-one manner to a node in the graph γ that defines the structure of the system by specifying the existing bonds among the cells. The P system $\mathcal{P}$ evolves according to the rules in R, which contains rules for modifying and

communicating the objects placed inside the cells, for dividing a cell, for differentiating a cell by changing its label, and for terminally differentiating a cell by changing its label into a terminal one.

Now, consider a configuration of a population P system $\mathcal{P}$ at a time, which is given by a tuple

$$\Sigma = (C'_1, C'_2, \ldots, C'_m, \gamma')$$

where $C'_1, C'_2, \ldots, C'_m$ represents the current population of cells in the P system $\mathcal{P}$, and $\gamma' = (\{1, 2, \ldots m\}, A)$, with $A \subseteq \{\{i, j)\} \,|\, 1 \leq i \neq j \leq m\}$, is the graph that defines the current structure of the P system $\mathcal{P}$, for some $m \geq 1$. The meaning of the rules in R is given below.

Evolution Rules. When a rule $[_i\, x \rightarrow y\,]_i$ is used in a cell C'_h of type i, the multiset x is replaced by a multiset y of pairs (a, t) with $a \in V$, $t \in \{here, go, out\}$. The target t is meant to specify the place where the object a has to be moved: *here* means the object has to stay in cell C'_h; *go* means the object has to be moved into one cell C'_p non-deterministically chosen such that there exists and edge $\{h, p\}$ in γ'; *out* means the object has to be removed from the system by releasing it in the extracellular space. As usual, evolution rules are applied in maximal parallel non-deterministic manner. The objects that are produced by means of these rules are immediately moved according to the corresponding target and they will become available in the respective places by starting from the next step. The objects that are released in the extracellular space are considered as being lost.

Cell Division Rules. A cell $C'_h = (w, i)$ can be divided by means of a rule

$$[_i\, x\,]_i \rightarrow [_i\, y\,]_i [_i\, z\,]_i \tag{2}$$

if and only if, it contains at least an occurrence of the multiset x. Thus, for each cell $C'_h = (w, i)$ that divides by means of rules (2), the following holds:

- two new cells $C'_{h'} = (w', i)$, and $C'_{h''} = (w'', i)$ are created where w' and w'' are the same as w except for the multiset x, which is replaced in w' by y, and in w'' by z; the node h in γ' is replaced by two nodes h' and h''; for each edge of the form $\{h, r\}$ that is removed from the graph two new edges $\{h', r\}$, $\{h'', r\}$ are introduced; an edge $\{h', h''\}$ is also created, which creates a bond among the two new cells; these operations are performed in parallel at the same time for all the cells that divide in the same step of a computation; thus, at the end of this stage, all the pairs of cells that result from the division of a single cell will be bonded to all the cells (or to the offspring of these cells) that formerly were bonded with the mother cell; we call this new graph γ'';

- for each node p in γ'' that has been produced by cell division, we the subset of edges $E_p = \{\, \{p, r\} \,|\, (C'_p, C'_r) \in \alpha$ or $(C'_r, C'_p) \in \alpha\,\}$ from γ''; we remove then all the edges $\{p, q\} \notin E_p$ from γ'' (i.e., a new cell will retain bonds with all the cells in its neighbourhood that are similar to it according to the relation α).

- the nodes and the edges in the graph that results from the process of forming new bonds are eventually renamed in a one-to one manner with labels in $\{1, 2, \ldots, m'\}$, where m' is the number of nodes in the resulting graph.

Cell Differentiation Rules. In presence of a particular multiset x a cell of type i can differentiate into a cell of type j by means of a rule $[_i\, x\,]_i \rightarrow [_j\, y\,]_j$, which changes the label the cell gets assigned and replaces the multiset x with the multiset y.

Terminal Cell Differentiation Rules. In presence of a particular multiset x a cell of type i can differentiate into a terminal one by means of a rule $[_i\, x\,]_i \rightarrow [_t\, y\,]_t$, which turns the label the cell gets assigned into a label in T and replaces the multiset x with the multiset y. A cell that gets assigned a label in T cannot evolve anymore as these labels are not involved in any rule.

Moreover, only one rule per cell that is either a division, a differentiation, or a terminal differentiation one can be used in a step of a computation.

The result of a successful computation is the Parikh vector associated with the multiset of objects situated in the terminal cells.

Thus, we have got a variant of population P systems where the graph defining the structure of the system may change during a computation through the use of cell division rules that introduce new cells in the system and new links among the cells in the system. As well as this, the objects placed inside the cells can be modified and communicated from a cell to another one by means of some evolution rules, and cells can change the labels they get assigned by means of some differentiation rules. The use of differentiation rules results in a change of the rules that can be applied in a cell. It appears clearly that the rules considered here are inspired by what has been done so far in the framework of P systems with active membranes [16]: a class of P systems where the membrane structure is modified by an operation called membrane division.

Therefore, the first obvious challenge is to investigate the generative capacity of population P systems with cell division by comparing it with the generative power of P systems with active membranes; the efficiency of population P systems in solving NP complete problems has to be investigated as well. Some variants of this model are presented in the sequel.

Remark 3.1. As we have already said, the relation of similarity α is considered as being as general as possible without any specific restrictions: one can express any sort of relation between the content of two cells in terms of the forms of the multisets they contain. For instance the version provided by Definition 3.1 may be replaced by a condition that specifies that two cells C_p, C_q are similar if and only if, the strings that represent the respective multisets of objects, w_p, w_q belong to a particular (infinite) language $L_1 \times L_2$. Some simpler specifications for the relation α may be of interest as well. In this respect, an easy choice is to consider α as a finite set of connecting rules defined either by pairs of labels or by pairs of finite multisets of objects. Furthermore, it may be interesting to identify notions of similarity that bear some sort of biological inspiration or relevance. For instance, the model proposed in [19] considers the process of forming new bonds as depending on a value of probability that a cell gets assigned. Specifically, a random number between 0 and 1 is generated and bonds will be formed with cells whose probability value is greater than this number. The values of probability associated with cells are continually updated.

Remark 3.2. Another variant of these systems may be considered by evaluating the graph structure associated with after each computation step by adding or removing bonds between any two cell components C_p, C_q depending on whether $(C_p, C_q) \in \alpha$ or $C_p, C_q) \notin$

α, respectively. As well as this, a specific operation of cell death of the form $[_i\, x\,]_i \rightarrow \dagger$ can be considered, which causes the removal of a cell of type i from the system.

Remark 3.3. Apart from membrane division, another important feature of P systems with active membranes is membrane polarisation, which is specified as a further parameter that is assigned to membranes. Such a feature is not considered in population P systems with cell division according to Definition 3.1. In fact, this is coherent with the current trend of research in P systems with active membranes of removing polarisation and focusing on the operation of changing the labels of the membranes [3], [17]. In particular, it is shown in [3] that P systems with rules for changing the labels are as powerful as P systems that use membrane polarisation: they are computationally complete. This important feature of changing the label is captured here by means of specific differentiation rules, which are the unique rules that are allowed to do so. Thus, with respect to what was done in [3], we consider for population P systems a more restrictive framework where neither cell division rules nor communication rules are able to change the label of any cell. Furthermore, population P systems with cell division adopt a weaker form of communication where the target cell for an object is always non-deterministically chosen. These observations suggests a first research topic that concerns the power of P systems where the operation of changing the labels is restricted to specific differentiation rules and where a non-deterministic form of communication is adopted. As a first step in this direction, we show here how to express the feature of membrane division for elementary membranes in the context of population P systems with cell division.

Consider a population P system $\mathcal{P}$ where the graph γ is a tree. We place a distinct object d inside each one of the cells that corresponds to nodes that have some leafs among the children . Next, we assign an object $\bar{d}$ to each cell that corresponds to a leaf. Then, we consider cell division rules of the form:

$$[_i\, \bar{d}\, a\,]_i \rightarrow [_i\, \bar{d}\, b\,]_i [_i\, \bar{d}\, c\,]_i,$$

which specify that only cells that correspond to a leaf in the tree γ (i.e. elementary membranes) can divide; these are the unique cells that contain some object $\bar{d}$. For what concerns the process of bond making, we can consider a simple relation α that states that a cell $C_1 = (w_1, i_1)$ is similar to a cell $C_2 = (w_2, i_2)$ if and only if, w_1 contains an object d and w_2 contains the corresponding object $\bar{d}$, or vice versa. In this way, when a cell that contains an object $\bar{d}$ divides, this will always form a bond with a cell in its neighbourhood that contains an object d; this cell is unique and corresponds to its parent node.

At this point, by going alongside this direction, we can conjecture that we can get for tissue P systems with cell division the full power of P systems that use rules for changing the labels and membrane division for elementary membranes. Moreover, this can be done by considering α as defined by a finite set of pairs (a, b), with $a, b \in V$; these specify that two cells irrespective of their labels can be linked each other if they contain respectively an object a and an object b.

Remark 3.4. An alternative way of getting an output from P systems where the structure changes during a computation has been considered in [4] where, instead of considering the output of a computation as a single entity, the output is given by catenating the content or the labels of each region of the whole configuration reached by the system at the end of a computation. Specifically, three output modes have been considered in [4] for generating languages by means of P systems with active membranes: visiting the tree associated with

the membrane structure, sending out the objects, and collecting the traveller traces. Now, it appears clearly that there exists the possibility of reconsidering this approach in the context of population P systems with cell division where the structure of the system is defined as an arbitrary graph. Moreover, we have introduced in definition 3.1 a distinct alphabet of terminal labels T that are used to identify cells that result from the application of terminal differentiation rules; cells that get assigned a label in T cannot evolve anymore by means of any rule. This feature of population P systems might be used to get a specific output from a computation by considering for instance the labels of terminally differentiated cells in the order they are produced. In a similar way, we can think of using the traces of the traveller but restricted to the case of terminal labels. We again face the challenge of reconsidering the results previously obtained in [4] in this framework based on population P systems.

Remark 3.5 (The Gradient Rule). As we have already mentioned, cell division plays a fundamental role in the process of developing a multicellular organism from a single cell as well as cell differentiation that allows cells to specialise in many different ways. The process of cell differentiation can be regarded as a switching process where some genes become active and others become inactive. This regulation mechanism of the gene expression is heavily dependent on interactions and cooperations among different cells that make possible variations on the concentration of chemicals inside the cells while the organism is developing from the embryo. In this way, cells tend to become regionally determined: different regions are formed where cells obtain different concentrations of specific chemicals and follow separate differentiating paths [2]. In some important cases, one can observe the formation of chemical gradients where special signal molecules diffuse out from a localised source and their concentration increase as one moves away from the source. Cell at different distances from the source are thus driven to behave in different ways according to the concentration of the signal molecules. A typical example is given by the development of *Drosophila* where chemical gradients were recognised for the first time as being responsible for the initial segmentation of the body into head, thorax and abdominal region as well as the dorsal-ventral axis determination. [11].

In Natural Computing area, developmental biology has been constituting a major source of inspiration for the definition and implementation of shape languages and visualising tools for systems with a dynamical structure, such as, L systems approaches [12], MGS language [10], and Amorphous Computing [1], [14]. In this latter approach, there is an interest in understanding how global shapes can be generated by means of local interactions at the level of simple computational units. More precisely, a system is defined as a collection of cells that are able to interact locally by means of various primitives. Among these, specific primitives for manipulating gradient values have been identified. Each cell can create a gradient, communicate a gradient value to its neighbours, and receive a gradient value from its neighbours. A cell always stores the gradient with the minimal value among those it has received. In this way, the gradient value is used to have a rough estimation of the distance from a possible source of information.

Here, we present a straightforward model of chemical gradient for population P systems with cell division that is based on varying the number of occurrences of a special object g inside the cell. Specifically, let $\mathcal{P}$ be a tissue P system with cell division. We identify in the graph γ that defines the structure of the system a subset of source cells that will not contain any occurrence of the object g. Each cell of type i that can be reached in n steps from a source cell of type i through the graph γ gets assigned a multiset g^n; if a cell can be reached from several different sources then, we assign to the cell the multiset with

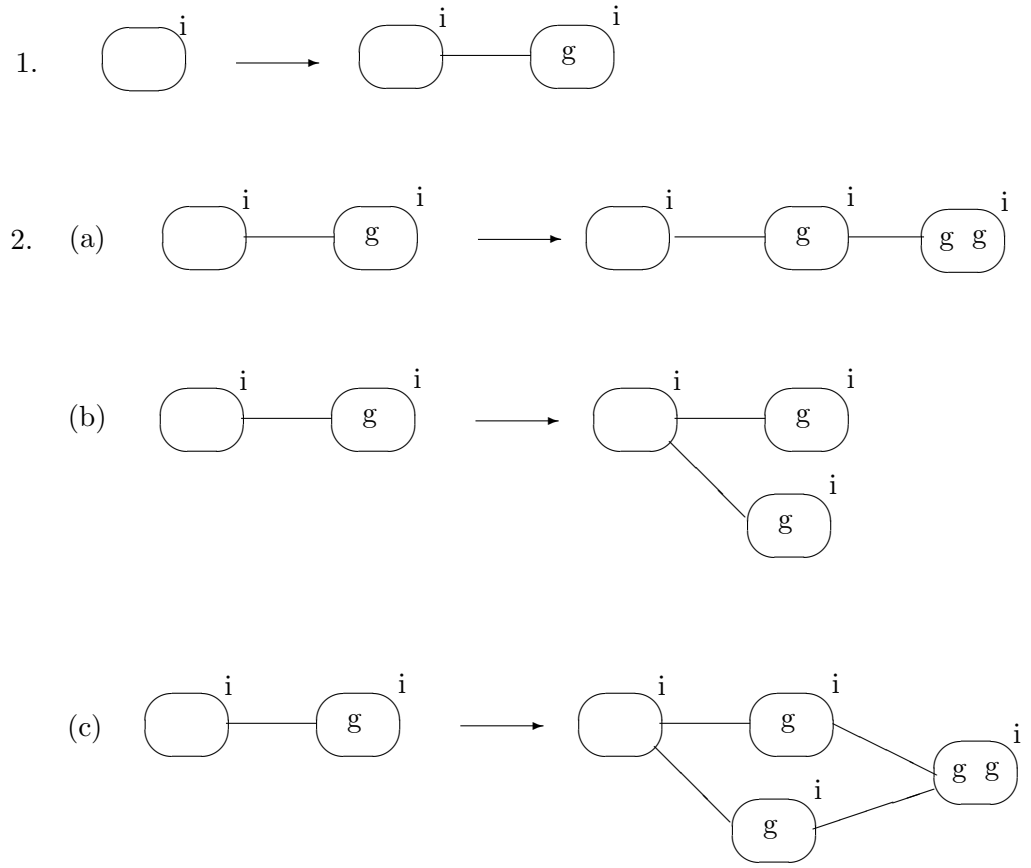

Figure 2: *A source cell of type i divides by using either a rule (a) $[_i\, d\,]_i \to [_i\,]_i\,[_i\, d\, g\,]_i$, or a rule (b) $[_i\, d\,]_i \to [_i\, d\,]_i\,[_i\, g\,]_i$, or a rule (c) $[_i\, d\,]_i \to [_i\, d\,]_i\,[_i\, d\, g\,]_i$; (a) the new cell with the greatest gradient value is given the chance to divide (i.e., the system develops alongside a line); (b) the new cell with the lowest gradient value is given the chance to divide (i.e., cells irradiate from a common origin point); (c) both the new cells are given the chance of dividing.*

the minimal number of occurrences of g. Next, we consider cell division rules of the form $[_i\, x\,]_i \to [_i\, y\,]_i[_i\, z\, g\,]_i$, with $|x|_g = |y|_g = |z|_g = 0$ (i.e., a cell that divides increases by 1 the number of objects g in one of the daughter cell). As well as this, we specify the object g can never be modified by any rule.

At this point, we can state the gradient rule that drives the process of bond making: a cell of type i that contains a multiset g^n can form bonds only with cells of type i that contain either a multiset g^{n-1} or a multiset g^{n+1}. This is the way the relation α is defined. Figure 2 illustrates an example of two division steps that are obtained by starting from a single source cell.

Cases where the operation of cell differentiation is used are more problematic because a cell of type i can differentiate into a cell of type j, which may be the first cell of type j in the system. In that case, when the cell of type j divides, the new cells cannot form bonds with any of the preexisting cells and the graph will result to be unconnected. In order to

avoid this, one can think of refining the relation α in such a way that specifies how to relate cells of different types. Moreover, the gradient value can be only modified by cell division rules and cells that stop dividing cannot have their gradient value changed anymore even the surrounding context changes. For instance, if a source cell of type i disappears then, the other cells lose all the information about the distance from the source. Anyway, we can only formulate as an interesting research topic to look at how the gradient rule emerges from interactions among cells and how a full-developed P system can be produced by starting from an initial population of few cells.

4 Final Remarks

This paper, mostly reporting work under development, refers to tissue P systems and proposes two variants of this model using, in the framework of an evolution-communication paradigm, receptors and signal objects. The last part introduces a population P system model which relies on cell division and involves a dynamic structure of the considered skeleton framework which evolves according to some prescribed rules. A list of various problems is mentioned and a strong biological motivation is provided. A number of results are still under scrutiny and many more could emerge later on. In a forthcoming paper we have to address in more depth the problems related to tissue P systems using receptors and signals and later on the wealth of variants that may be considered in relationship with yet very crude concept of population P systems.

References

[1] Abelson, H., Allen, D., Coore, D., Hanson, C., Homsy, G., Knight, F., T., Nagpal, R., Rauch, E., Sussman, J., G., Weiss, R. (2001). Amorphous Computing. *Communication of the ACM*, **43**, 1.

[2] Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P. (2002). *The Molecular Biology of The Cell. Fourth Edition*. Garland Publ. Inc., London.

[3] Alhazov, A., Pan, L. Păun, Gh. (2003). Trading Polarization for Labels in P Systems. Submitted.

[4] Bernardini, F., Gheorghe, M. (2003). Languages Generated by P Systems with Active Membrane, *New Generation Computing*, to appear.

[5] Bernardini, F., Manca, V. (2003). P Systems with Boundary Rules. In [18], 107–118.

[6] Cavaliere, M., (2003). Evolution Communication P Systems. In [18], 134–145.

[7] Csuhaj-Varju, E., Vaszil, G. (2003). P Automata or Purely Communicating Accepting P Systems. In [18], 219–233.

[8] Dassow, J., Păun, Gh. (1989). *Regulated Rewriting in Formal Languages*. Springer-Verlag, Berlin, Heidelberg, New York.

[9] Freund, R., Martín-Vide, C., Obtulowicz, A., Păun, Gh. (2002). On Three Classes of Automata-Like P Systems. *Proc. DLT 2003*, Szeged, Hungary, LNCS 2710, Springer, 2003, 292–303.

[10] Giavitto, J., Michel, O. (2001). MGS: A Rule-Based Programming Language for Complex Objects and Collection. *Electronic Notes in Theoretical Computer Science*, **59**, Elsevier Science Publisher.

[11] Lawrence, P.A. (1991). *The Making of a Fly: The Genetics of Animal Design*. Blackwell Scientific, Oxford.

[12] Lindenmayer, A., Prusinkiewicz, P. (1990). *The Algorithmic Beauty of Plants*. Springer-Verlag, Berlin, Heidelberg, New York.

[13] Martín-Vide, C., Mauri, G., Păun, Gh., Rozenberg, G., Salomaa, A., eds. (2004). *Membrane Computing. International Workshop, WMC 2003, Tarragona, Spain, July 2003. Revised Papers. Lecture Notes in Computer Science*, **2933**, Springer-Verlag, Berlin, Heidelberg, New York.

[14] Nagpal, R. (2001). *Programmable Self-Assembly: Constructing Global Shape using Biologically Inspired Local Interactions and Origami Mathematics*. PhD Thesis, Massachusetts Institute of Technology.

[15] Păun, Gh. (2000). Computing with Membranes. *Journal of Computer and System Sciences*, **61**, 1, 108–143.

[16] Păun, Gh. (2002). *Membrane Computing. An Introduction*. Springer-Verlag, Berlin, Heidelberg, New York.

[17] Păun, Gh. (2004). Further Open Problems in Membrane Computing. Available online at http://psystems.disco.unimib.it/.

[18] Păun, Gh., Rozenberg, G., Salomaa, A., Zandron, C., eds. (2003). *Membrane Computing. International Workshop, WMC-CdeA 02, Curtea de Arges, Romania, August 19-23, 2002. Revised Papers. Lecture Notes in Computer Science*, **2597**, Springer-Verlag, Berlin, Heidelberg, New York.

[19] Walker, D., Holcombe, M., Southgate, J., MacNeil, S., Smalwood, R. (2004). Agent-Based Modeling of The Social Behaviour of Cells. *Biosystems*, to appear.