# EP-colonies: Micro–Organisms in a Cell-like Environment

Erzsébet Csuhaj-Varjú

Computer and Automation Institute
Hungarian Academy of Sciences
Kende utca 13–17, H-1111 Budapest, Hungary
E-mail: csuhaj@sztaki.hu

**Summary.** The aim of this note is to introduce a model for describing populations of extremely simple organisms which live in and interact with a dynamically changing cell-like environment. In addition to the definition of the notion, we present examples illustrating some possible properties of these constructs and propose some bio-inspired problems to study.

## 1 Introduction

P systems or membrane systems are unconventional models of computing, in most cases as powerful as the Turing machines. The notion – abstracted from the structure and the functioning of the living cell – was introduced by Gh. Păun [4]. A P system is a membrane architecture which separates regions containing multisets of objects. These objects correspond to chemicals (chemical ingredients), which – representing chemical reactions – might change and move among the different regions, according to rules associated to the regions. As the reader might observe, a P system is a rather complex object, and this is why it is suitable both for computing and modelling biological or bio-chemical phenomena. The model, where the P system communicates with an environment through input and output objects, is called P automaton [2]. Certain variants of P automata are computationally complete, moreover it can be shown that collections of extremely simple P automata-like constructs can also form computational devices of the power of the Turing machines. Such constructs are the so-called P colonies [3], a class of abstract devices composed from as simple as possible P automata-like systems acting and evolving in a shared environment. The members of the P colony, the cells, are represented by a bounded number of objects (and this number does not change during the computation) and a finite number of programs which are tuples of rules of the form $a \to b$ or $c \longrightarrow d$, where $a, b, c, d$ are objects. The rule $a \to b$ means that an object $a$ in the cell is changed for an object $b$, i.e., the chemical changes according to rules of a chemical reaction. The rule $c \longrightarrow d$ means that an object

$c$ found inside the cell is exchanged with an object $d$ found outside, in the environment, that is, $d$ is input the cell and $c$ is output. The latter rule describes the communication between the cell and its environment and each program contains only one communication rule. The cells are able to scrutinize their environment, that is, to observe the presence or absence of a certain object. The cell functions by changing its states which is realized by performing one of its programs. The application of a program means a parallel application of a rule for each object in the cell. A P colony functions in such manner that at any moment of time each enabled cell changes its states. Those cells which are not able to perform a program remain inactive. Both at the beginning and during the whole functioning of the system the environment is supposed to be an infinite supply of a special object, called the environmental object. After starting, the evolution of the collection of the cells, the P colony, continues as long as no cell can be found which is able to apply any of its programs. In this case we say that the computation halts. The result of a halting computation is the number of copies of a distinguished object present in the environment. It is easy to observe that the role of the environment is essential in these constructs, it is a medium both for communication and for storing an unbounded quantity of information. In [3] it was shown that P colonies are as powerful as the Turing machines. The large computational power of P colonies is due to both these features of the environment and the capability of the cells to check the presence of the objects in the environment.

One natural idea which immediately emerges examining the concept of P colonies is the following: What about constructs where the environment is bounded but dynamically changing not only due to the actions of the cells but to its own rules as well? For example, what about P colonies where the actual objects in the environment are generated in some manner? Continuing this line of ideas, what about constructs, where the structured environment is a P system and the regions might contain usual objects, which correspond to chemicals, and other objects which correspond to cells in a P colony? These latter constructs can be considered as models of cells with very simple micro-organisms inside the regions. How these micro-organisms influence the behavior of the cell? What can we say about these systems both from computational point of view and from the point of view of modelling biological phenomena? We note that this model is a related concept to the so-called eco-grammar system [1], a formal language theoretic model of a collection of agents which are in interaction with their dynamically changing environment. The theory of eco-grammar systems forms a subfield of grammar systems theory, an area in formal language theory providing models for describing multi-agent systems.

In this note, we introduce the new model, called EP colony, where letter E refers both to evolution and environment, and then we illustrate the notion by some very simple examples. At the same time, we propose topics, inspired by biological phenomena, for future research.

## 2 The Concept

We suppose that the reader is familiar with the basic elements of formal languages, thus we introduce here in a formal way only our computing model and present a few basic notions.

An *alphabet* is a finite and non-empty set of abstract symbols. For an alphabet $A$ we denote by $A^*$ the set of all strings of symbols from $A$, including the empty string. A *multiset* over an alphabet $A$ is a mapping from $A$ to $\mathbf{N}$, the set of natural numbers; we represent a multiset by a string from $A^*$ (the number of occurrences of a symbol $a \in A$ in a string $w$ represents the multiplicity of $a$ in the multiset represented by $w$; thus, all strings obtained by permuting symbols in the string $w$ represent the same multiset).

We present the basic notions of membrane computing; the interested reader may find more detailed information on the theory of P systems in the monograph [5]. A P system is a structure of hierarchically embedded membranes, each having a label and enclosing a region containing a multiset of objects and possibly other membranes. The out-most membrane which is unique and usually labelled with 1, is called the skin membrane. The membrane structure is represented by a sequence of matching parentheses where the matching pairs have the same label as the membranes they represent. If $x \in \{[_i, ]_i \mid 1 \leq i \leq n\}^*$ is such a string of matching parentheses of length $2n$, denoting a structure where membrane $i$ contains membrane $j$, then $x = x_1\ [_i\ x_2\ [_j\ x_3\ ]_j\ x_4\ ]_i\ x_5$ for some $x_k \in \{[_l, ]_l \mid 1 \leq l \leq n,\ l \neq i, j\}^*,\ 1 \leq k \leq 5$. If membrane $i$ contains membrane $j$, and there is no other membrane, $k$, such that $k$ contains $j$ and $i$ contains $k$ ($x_2$ and $x_4$ above are strings of matching parentheses themselves), then we say that membrane $i$ is the parent membrane of $j$, denoted by $i = parent(j)$, and at the same time, membrane $j$ is one of the child membranes of $i$.

By the contents of a region we mean the multiset of objects which is contained by the corresponding membrane excluding those objects which are contained by any of its child membranes.

Now we introduce the computing device we consider in this paper.

**Definition 1.** *An EP colony is a construct*

$$\Pi = (V_E, V_A, \mu, R_1, \ldots, R_n, C_1, \ldots, C_s, w_1, \ldots, w_n),\ n \geq 1, s \geq 1,$$

*where:*

- $V_E$ *is an alphabet, the alphabet of environmental objects in* $\Pi$*;*
- $V_A = \{C_1, \ldots, C_s\} \cup \{(, )\}$*, where* $\{C_1, \ldots, C_s\}$ *is the alphabet of the names of the micro-organisms (the types of the micro-organisms) in* $\Pi$*, and symbols* ( *and* ) *are separators;* $V_E$ *and* $V_A$ *are disjoint alphabets.*
- $\mu$ *is a membrane structure with* $n$ *membranes;*

- *$R_i$, $1 \leq i \leq n$, is the set of rules of the $i$-th region (membrane); each rule is of the form $u \to v(tar)$, where $u$ and $v$ are multisets over $V_E$ and $tar \in \{here, in, out\}$;*
- *$C_j = (O_j, P_j)$, $1 \leq j \leq s$, is a micro-organism, where*
  - *$O_j$ is a finite multiset over $V_E$ (the initial state of the micro-organism), and*
  - *$P_j$ is a finite set $\{p_{j,1}, \ldots, p_{j,k_j}\}$, $k_j \geq 1$, of programs. Each program $p_{j,l}$ is an $s_{j,l}$-tuple of rules of the form $(r_{j,l,1}, \ldots, r_{j,l,s_{j,l}})$, $s_{j,l} \geq 1$, with $s_{j,l} = card(Q_j)$, where $r_{j,l,1}$ is of the form $c \longrightarrow d$ (one object exchange with the environment), and rules $r_{j,l,2}, \ldots, r_{j,l,s_{j,l}}$ are of the form $a \to b$ (evolution rule). Objects $a, b, c, d$ are elements of $(V_E \cup V_A) - \{(, )\}$.*
- *$w_i$ is the initial contents (multiset) of the $i$-th membrane, with $w_i = v(O_{i_1}) \ldots (O_{i_{s_i}})$, where $v$ is a multiset of objects over $V_E$ and $O_{i_k}$, for $1 \leq k \leq s_i$, are initial states of micro-organisms.*

The initial state of $\Pi$ is the $n$-tuple $(w_1, \ldots, w_n)$, where $w_i$, $1 \leq i \leq n$, is the initial state of region $i$.

We define the state of the EP colony in an analogous manner. It is an $n$-tuple $(u_1, \ldots, u_n)$, where $u_i$ is the state of the $i$-th region and $u_i = v(H_{i_1}) \ldots (H_{i_{r_i}})$, $r_i \geq 0$, where $v$ is a multiset of objects from $V_E$ and $H_{i_k}$, $1 \leq k \leq r_i$ is a multiset of objects over $(V_E \cup V_A) - \{(, )\}$, the state of a micro-organism found in region $i$. We note that the number of the micro-organisms found in a region can change under the functioning.

Notice that any micro-organism can be considered as a small P system with only one membrane; we use the separator symbols ( and ) only for technical reasons.

Thus, the regions of the EP colony might contain two types of objects: firstly, usual chemicals, represented by elements of $V_E$. Multisets of these objects might undergo chemical reactions, described by the rules $u \to v(tar)$. In this case, a multiset of objects $u$ changes for a multiset of objects $v$ and then $v$ either remains in the region ($tar = here$), or leaves the region in some direction ($tar = in$ or $tar = out$). These rules correspond to standard rules of P systems. Secondly, there might be micro-organisms in the region, which can change their states. Under a change of their state, these micro-organisms perform a rule for any object they have. When they apply the communication rule, then they exchange an object (this can be the same object) with their environment which is now the contents of the region. When they perform an evolution rule, then they rewrite (possibly identically) an object they have. Notice an important point in the definition: the rules of the micro-organism are allowed to manipulate (create, delete, exchange) micro-organisms as follows: if a communication rule $a \longrightarrow C$ is used, where $C$ is a letter from $V_A$, that is, the name of a micro-organism, then a new micro-organism gets birth in the region (with the initial contents of its type); if a communication rule $C \longrightarrow a$ is used, where $C$ is a letter from $V_A$, then a micro-organism of type $C$ disappears (temporarily) from the region. A rule $a \to C$ with $C$ as above and $a \in V_E$ indicates the potential birth of a micro-organism in another micro-organism, while a rule $C \to a$ refers to that the micro-organism is extinct, and

the rule $C_1 \rightarrow C_2$ with $C_1, C_2 \in V_A$ means that the type of the potential micro-organism is changed. Thus, the rules make possible to change the number of the micro-organisms in the regions.

After this explanation we define the way of the functioning of EP colonies. Since the formal details are rather complicated, we present an informal explanation.

The EP system $\Pi$ functions by changing its states as follows: at every direct change of the state each enabled micro-organism (a maximal number of enabled micro-organisms) in any region must perform one of its programs. In the first case the micro-organisms function in a so-called strongly competitive manner, while in the second case they work in a weakly competitive manner. The micro-organism is called enabled if it is able to perform one of its programs. In the case of strong competition, if there is no sufficient environmental symbol or micro-organism in the region to process for every enabled micro-organism, then the system is blocked. In the case of weak competition, those micro-organisms which are not able to perform a program remain inactive and then they either remain in the region or leave the region in some direction, chosen non-deterministically. The environmental objects in the region which were not affected by the micro-organisms are developed and communicated by applying the rules of the corresponding region. These rules are applied in a maximal parallel manner, that is, as many rules are applied in parallel as it is possible: if a rule can be applied in the environment, then it must be applied.

The sequences of direct changes of the state define the computations in the EP system. We can define the result of the computation through the concept of halting and thus by a halting state. Halting can be considered in different manners: One possibility is, for example, when there is no micro-organism to perform any program in any of the regions. The second (stronger) condition is: if no action is possible to be performed in the systems, i.e., there is no enabled micro-organism and no enabled rule in any region. (A rule of a region is enabled if it can be performed.)

While the above variants define the result of the computation by halting, we can define it through equilibrium as well, that is, as states which change for themselves.

## 3 Topics for Research

EP colonies as P colonies are rather complex constructs, both in their structure and their functioning, thus their large computational power is expected. In [3], the computational completeness of P colonies was proved, by different simulations of register machines (which is known to be of universal power). To obtain this large computational power, at any step of the computation a certain but bounded number of new copies of the so-called special environmental symbol was used. (Remember that the environment in this case contains an infinite number of copies of this symbol). Using some trivial generative rules and a special new so-called marker symbol, we can easily produce this amount of environmental symbols at each step. Thus, it is easy to demonstrate that EP colonies, even with one region

and with a limited number of micro-organisms (and this number does not change under functioning) are as powerful as the register machines. Therefore, if we would like to study EP colonies from the point of view of computation, we should find natural and reasonable restrictions of the model.

However, one interesting problem to study is, whether or not the two different functioning modes are equivalent. Namely, to decide whether or not the strong competition and the weak competition lead to different computational powers, and which are the restrictions of the model that imply difference. The answer to this question can also be related to the following problem. How long a micro-organism can remain alive, can resist in an inactive state? Supposing that inactive states are possible, has this fact any influence on the behavior of the system?

The following simple example demonstrates some differences between the two functioning modes.

*Example 1.* Instead of the formal details, we describe the EP colony in an informal manner. Suppose that it has one membrane and three identical micro-organisms. The rules of the membrane are $a \to a$, $a \to b$, $b \to a$, $b \to b$, and the program of the micro-organisms is $(a \leftrightarrow b, b \to a)$, and they have $ab$ as initial state. Assume the EP colony having the initial state $(b^6(ab)(ab)(ab))$. In the case of strongly competitive functioning the EP colony is able to work if and only if at any step of the functioning at least 3 copies of object $b$ are present in the environment; if this is not the case, then the functioning of the EP colony gets blocked. In the case of the weak competition among the micro-organisms, the system can continue its work for any number of occurrences of $b$.

While in the case of P colonies the research emphasis was put on the computational power of the system, in order to demonstrate that very simple computational devices with indirect communication can be of large computational power, in the case of EP colonies the properties of their state sequences are in the focus of interest: for example, can we detect periodicity with respect to some properties in the state sequence or the EP system is with aperiodic behavior in certain properties. For example, is there some periodicity in the number of micro-organisms or in the number of micro-organisms of the same type occurring in a certain region during the functioning of the system? Or, if we associate patterns to the population of the micro-organisms, can we identify some well-known patterns under the functioning?

EP systems provide the possibility of defining unconventional complexity measures. For example, in the case of weak competitive functioning, important measures can be the so-called migration complexity measures: the number of micro-organisms leaving the region where they are present (these micro-organisms migrate from one region into another one), the number of micro-organisms leaving a certain region at a change of a state or in the course of the functioning, etc. Analogously, we can examine these parameters with respect to a certain type of micro-organisms. Interesting questions are, what about EP systems with constant, linearly bounded, or polynomially bounded migration complexities?

The following simple example demonstrates the behavior of an EP colony under migration.

*Example 2.* As in the previous case, we present only the necessary formalism. Let the EP colony have two membranes, namely the membrane structure $[[\ ]_2]_1$, and one micro-organism. Let the rules of the regions be defined as follows: region 1: $a \to a$, $a' \to a''a''$, $a'' \to a$ and region 2: $d \to d^2, c \to c$.

Let the micro-organism have the following programs:

- $(a' \leftrightarrow a, a \to a', c \to c)$,
- $(c \leftrightarrow d, a' \to a'', a \to a)$,
- $(a'' \leftrightarrow a, a' \to a'', d \to d)$,
- $(d \leftrightarrow c, a'' \to a', a \to a)$.

Suppose that the micro-organism starts working from the initial state $a'ac$ and let the EP colony has the initial state $(a(a'ac), d)$. (The micro-organism is in region 1.) Moreover, suppose that when the micro-organism has no applicable program in the region, then it moves to the other region.

Then, for any $k$, there are periods of the computation of length $2^k$ such that the micro-organism does not leave the first region during this period.

Similarly to the migration complexity, other interesting complexity measures can be the so-called infection size and the so-called infection level of the EP colony. The size of the infection of a region (of the EP colony) is the number of micro-organisms found in the region (in the system); while the infection level is the number of the micro-organisms divided by the number of environmental objects or by the total number of objects in the region (or in the system). Notice that these measures are of particular interest, they might make it possible to define the notion of an epidemic, i.e., an increasing number of micro-organisms under functioning, or resistance, i.e., an (almost) constant number of micro-organisms during the whole computation.

We illustrate these notions by a simple example.

*Example 3.* Instead of the formal details, we give the example in a rather informal way. Suppose that the EP colony has one membrane and $k$ micro-organisms. Let the region have the following rules: $a \to a^2$, $b_i \to a^2$, $1 \leq i \leq k$.

Let micro-organism $C_i$, $1 \leq i \leq k$, be given with the following program: $(a \leftrightarrow b_i, a \to b_i)$, and let its initial state $ab_i$.

Suppose that the initial state of the EP colony is $a^{2k+1}(ab_1)\dots(ab_k)$.

Then, at any moment of time, objects $b_i$, $1 \leq i \leq k$, are present in the region and the number of copies of object $a$ is exponentially increasing. The infection size is constant and the infection level converges to the zero.

EP colonies offer a wide variety of problems for study. The model is sufficiently complex not only for computing but realizing interesting behavioral patterns as well. Especially interesting candidates for future research are the extensions to the models defined over two-dimensional or three-dimensional objects. These and similar questions are waiting for investigation.

## References

1. E. Csuhaj-Varjú, A. Kelemenová, J. Kelemen, Gh. Păun: Eco-grammar systems – A grammatical framework for life-like interactions. *Artificial Life*, 3 (1997), 1–28.
2. E. Csuhaj-Varjú, Gy. Vaszil: P automata. In *Pre-Proceedings of the Workshop on Membrane Computing WMC-CdeA 2002* (Gh. Păun, C. Zandron, eds.), Curtea de Argeş, Romania, August 19-23, 2002. Pub. No. 1 of MolCoNet-IST-2001-32008 (2002) 177–192, and also in *Membrane Computing. Lecture Notes in Computer Science* (Gh. Păun, G. Rozenberg, A. Salomaa, C. Zandron, eds.), LNCS 2597, Springer-Verlag, Berlin, 2003, 219–233.
3. J. Kelemen, A. Kelemenová, Gh. Păun: Preview of P colonies: A biochemically inspired computing model. In: *Workshop and Tutorial Proceedings. Ninth International Conference on the Simulation and Synthesis of Living Systems (Alife IX)* (M. Bedau et al., eds.), Boston Mass., 2004, 82–86.
4. Gh. Păun: Computing with membranes. *Journal of Computer and System Sciences*, 61 (2000), 108–143.
5. Gh. Păun: *Membrane Computing. An Introduction.* Springer-Verlag, Berlin, 2002.