A characterization of symport/antiport P systems through Information Theory ⋆ (*extended abstract*)

José M. Sempere

Departamento de Sistemas Informáticos y Computación.
Universidad Politécnica de Valencia.
email: jsempere@dsic.upv.es

Abstract. In this work we analyze communication P systems under the framework of Information Theory. Given a cell-like P system with communication and evolution rules, we analyze the amount of information that it holds as the result of symbol movements across the membranes. Hence, *antiport* rules can be defined as bidirectional information channels, while *symport* and evolution rules can be viewed as unidirectional information channels. Under this approach, we propose some results about the information of a P system and its entropy.

Keywords: Communication P systems, Information Theory, Entropy

Some basic concepts

P systems were introduced as a computational model inspired by the information and biochemical product processing of living cells through the use of membrane communication. In most of the works about P systems, information is represented as multisets of symbol/objects which can interact and evolve according to predefined rules. From the beginning, the most important component of the system has been the kind of rules that it holds. There have been different proposals to define the rules of the system such as evolution rules, communication rules, active rules to create/dissolve membrane structures, active rules with polarization, and so on and so forth.

Here, we pay our attention to the following fact: the rules of a P system produce/consume new symbols in different regions of the system, so they can be considered information regulators that act over a region, which can be considered information senders and receivers in a pure communication system. Hence, the behavior of the P system can be analyzed with Information Theory tools. Mainly, we can analyze any P system through the characterization of the entropies at every region according to its membrane structure and rules.

⋆ Work partially supported by the Spanish Ministry of Economy and Competitiveness under EXPLORA Research Project SAF2013-49788-EXP

We will introduce basic concepts related to multisets, Information Theory and P systems. We suggest to the reader the references [3, 4] to introduce membrane computing, and the books [2, 5] to introduce Information Theory. We will provide some definitions from multiset theory as exposed in [7].

Let D be a set. A multiset over D is a pair $\langle D, f\rangle$ where $f : D \longrightarrow \mathbb{N}$ is a function. If $A = \langle D, f\rangle$ and $B = \langle D, g\rangle$ are two multisets then the removal of multiset B from A, denoted by $A \ominus B$, is the multiset $C = \langle D, h\rangle$ where for all $a \in D$ $h(a) = max(f(a) - g(a), 0)$, and their sum, denoted by $A \oplus B$, is the multiset $C = \langle D, h\rangle$, where for all $a \in D$ $h(a) = f(a) + g(a)$. We will say that $A = B$ if the multiset $(A \ominus B) \oplus (B \ominus A)$ is empty that is $\forall a \in D$ $f(a) = 0$.

The size of any multiset M, denoted by $|M|$ will be the number of elements that it contains.

We can suggest to the reader the books [2, 5] and the classical work by C.E. Shannon [6] in order to have a full view of Information Theory.

An *information source* is defined by the tuple (S, P) where S is an alphabet and P a probability distribution over S. A cornerstone of Information Theory is the concept of *entropy* which is attached to information sources. The entropy of an information source I, with an alphabet S and probability distribution $P : S \rightarrow [0, 1]$ is defined as

$$H(I) = -\sum_{a \in S} P(a) \cdot \log_2 P(a)$$

Observe that we are working with trivial codes where the alphabet of an information source is its encoding. In addition, we have fixed the base 2 for the logarithm, so the information entropy is described in bits. The change from a binary base to a different one can be easily carried out in a logarithm base change.

A *cell-like P* system of degree m with *communication* rules is a construct

$$\Pi = (V, \mu, w_1, \cdots, w_m, (R_1, \rho_1), \cdots, (R_m, \rho_m), i_0),$$

where:

- V is an alphabet (the *objects*)
- μ is a membrane structure consisting of m membranes
- w_i, $1 \leq i \leq m$, is a string representing a multiset over V associated to the region i
- R_i, $1 \leq i \leq m$, is a finite set of rules of the form (u, v) with $u \neq \lambda$ and $v \neq \lambda$ (evolution rules), $(u, out; v, in)$ with $u \neq \lambda$ and $v \neq \lambda$ (antiport rule) and (x, out) or (x, in) with $x \neq \lambda$ (symport rule). The strings u, v and x are defined over the alphabet V.
- i_0 is a number between 1 and m and it specifies the *output* membrane of Π (in the case that it equals to ∞ the output is read outside the system).

Observe that, in the previous definition, we have omitted an output alphabet, a catalyst alphabet and dissolution rules. In addition, we have omitted priorities in the rule sets and other communication rules with explicit address. The main reason is that we want to establish a preliminary analysis with the most simple systems.

The entropy of a P system

We will define the entropy of a P system by analyzing how the multisets at every region evolve according to the rules of the system. First, we define the entropy of the multisets of the regions and, then, the entropy of a P system.

Definition 1. *(***self-referred entropy of a multiset***). Let us consider a multiset $A = \langle D, f \rangle$. The self-referred entropy of A is defined as*

$$H_s(A) = -\sum_{a \in D} fr(a) \cdot log_2 fr(a)$$

where $fr(a) = \frac{f(a)}{|D|}$.

Observe that the self-referred entropy of a multiset is a static concept given that we have substitute the probability distribution by the frecuency of appearance of every object at the region ($fr(a)$). Observe that there is no external probability distribution over the rules and the objects.

In the following, we analyze the evolution of self-referred entropies according to the system computations.

Definition 2. *Let Π be a P system of degree m and $c_t = (\mu, w_1^t, \cdots, w_m^t)$ be a configuration of the system during a computation at time t. Then*

1. *The absolute entropy of Π at time t is $H^t_{abs}(\Pi) = \sum_{1 \leq i \leq m} H_s(w_i^t)$*
2. *The maximal entropy of Π at time t is $H^t_{max}(\Pi) = max\{H_s(w_1^t), \cdots, H_s(w_m^t)\}$*

The question about the computation of the entropy of a P system is completely based on the calculation of the different multisets at every region, according to the rules that affect to that region. Hence, at time t the multiset w_i^t will evolve, in the next transition, to the multiset $w_i^t \ominus l(R, i) \oplus r(R, i)$, where $l(R, i)$ is a multiset based on the left hand side of the rules that affect to the region i, and $r(R, i)$ is a multiset based on the right hand side of the rules that affect to the region i.

In this work we will overview, among others, the following questions and aspects :

1. What is the relationship between the kind of rules at every region (symport, antiport or evolution) and the evolution of its entropy ?
2. What is the definition of confluence under an information theory point of view ?
3. How does the operational mode (i.e. maximal/minimal parallelism) affect to entropy ?
4. How is the entropy defined in a stochastic/probabilistic P system ?
5. What is the definition of the entropy of a P system, if the external output is defined ?

References

1. C. Calude, Gh. Păun, G. Rozenberg and A. Salomaa, *Multiset Processing* LNCS 2235. Springer. 2001.
2. T.M. Cover and J.A. Thomas. *Elements of Information Theory* John Wiley & Sons, Inc. 1991
3. Gh. Păun. *Membrane Computing. An Introduction.* Springer. 2002.
4. Gh. Păun, G. Rozenberg and A. Salomaa (editors). *The Oxford Handbook of Membrane Computing.* Oxford University Press. 2010.
5. S. Roman. *Introduction to Coding and Information Theory.* Springer. 1997.
6. C.E. Shannon. A Mathematical Theory of Communication. *The Bell System Technical Journal*, Vol. 27, pp 379-423, 623-656, July, October, 1948.
7. A. Syropoulos. *Mathematics of Multisets.* In [1] pp 347-358.