Spiking Neural P Systems:  
A Short Introduction and New Normal Forms

Linqiang Pan¹,², Gheorghe Păun²,³, Mario J. Pérez-Jiménez²

¹ Department of Control Science and Engineering  
Huazhong University of Science and Technology  
Wuhan 430074, Hubei, China  
lqpan@mail.hust.edu.cn, lqpan@us.es  

² Department of Computer Science and Artificial Intelligence  
University of Sevilla  
Avda. Reina Mercedes s/n, 41012 Sevilla, Spain  

³ Institute of Mathematics of the Romanian Academy  
PO Box 1-764, 014700 Bucharest, Romania  
george.pau@imar.ro, gpaun@us.es  

Summary. Spiking neural P systems are a class of P systems inspired from the way the neurons communicate with each other by means of electrical impulses (called “spikes”). In the few years since this model was introduced, many results related to the computing power and efficiency of these computing devices were reported. The present paper quickly surveys the basic ideas of this research area and the basic results, then, as typical proofs about the universality of spiking neural P systems, we present some new normal forms for them. Specifically, we consider a natural restriction in the architecture of a spiking neural P system, to have neurons of a small number of types (i.e., using a small number of sets of rules). We prove that three types of neurons are sufficient in order to generate each recursively enumerable set of numbers as the distance between the first two spikes emitted by the system; the problem remains open for accepting SN P systems. The paper ends with the complete bibliography of this domain, at the level of April 2009.

1 Introduction

Spiking neural P systems (SN P systems, for short) were introduced in [32] in the aim of defining computing models based on ideas specific to spiking neurons, currently much investigated in neural computing (see, e.g., [21], [42], [43]). The resulting models are a variant of tissue-like and neural-like P systems from membrane computing – we refer to [56] for basic information in membrane computing, [71] for a comprehensive presentation, and to the web site [81] for the up-to-date information.

In short, an SN P system consists of a set of neurons placed in the nodes of a directed graph and sending signals (spikes, denoted in what follows by
the symbol $a$) along synapses (arcs of the graph). Thus, the architecture is that of a tissue-like P system, with only one kind of objects present in the cells. The objects evolve by means of spiking rules, which are of the form $E/a^c \rightarrow a; d$, where $E$ is a regular expression over $\{a\}$ and $c, d$ are natural numbers, $c \geq 1, d \geq 0$. The meaning is that a neuron containing $k$ spikes such that $a^k \in L(E), k \geq c$, can consume $c$ spikes and produce one spike, after a delay of $d$ steps. This spike is sent to all neurons to which a synapse exists outgoing from the neuron where the rule was applied. There also are forgetting rules, of the form $a^s \rightarrow \lambda$, with the meaning that $s \geq 1$ spikes are forgotten, provided that the neuron contains exactly $s$ spikes. We say that the rules “cover” the neuron, all spikes are taken into consideration when using a rule. The system works in a synchronized manner, i.e., in each time unit, each neuron which can use a rule should do it, but the work of the system is sequential in each neuron: only (at most) one rule is used in each neuron. One of the neurons is considered to be the output neuron, and its spikes are also sent to the environment. The moments of time when a spike is emitted by the output neuron are marked with 1, the other moments are marked with 0. This binary sequence is called the spike train of the system – it might be infinite if the computation does not stop.

The result of a computation is encoded in the distance between consecutive spikes sent into the environment by the (output neuron of the) system. In [32] only the distance between the first two spikes of a spike train was considered, then in [66] several extensions were examined: the distance between the first $k$ spikes of a spike train, or the distances between all consecutive spikes, taking into account all intervals or only intervals that alternate, all computations or only halting computations, etc.

Systems working in the accepting mode were also considered: a neuron is designated as the input neuron and two spikes are introduced in it, at an interval of $n$ steps; the number $n$ is accepted if the computation halts.

Both in the generating and the accepting case, SN P systems were proved to be computationally complete (equivalent with Turing machines; we also say that SN P systems are “universal”: the equivalence with Turing machines is constructive, hence starting the proof of equivalence from universal Turing machines, or from equivalent universal devices, directly lead to universal SN P systems).

Recently, SN P systems were also used in order to devise (theoretical) ways to solve computationally hard problems in a feasible (polynomial) time. This is usually achieved in membrane computing by means of tools which allow producing an exponential working space in a linear time; the standard way to do it is membrane division. In the SN P area, a different strategy was first explored: with inspiration from the fact that the brain consists of a huge number of neurons out of which only a small part are used, in [9] one address computationally hard problems by assuming that an arbitrarily large SN P system is given “for free”, pre-computed, with a structure as regular as possible, and without spikes inside; solving a problem starts by
introducing spikes in certain neurons (in a polynomially bounded number of neurons a polynomially bounded number of spikes are introduced); then, by moving spikes along synapses, the system self-activates, and a specific output provides the answer to the problem. This was illustrated in [9] for SAT. Then, both ways to produce this “pre-computed” resource during the computation were considered ([74]), and rules for dividing neurons ([53]), with inspiration from the observation that there are so called neural stem cells, which divide repeatedly, producing new neurons ([18]). The research in this last direction is just started, and further progresses are expected.

The present survey is a brief one, many further directions of research were explored (asynchronous SN P systems, inhibitory spikes and/or synapses, using the rules in a parallel way, Hebbian learning for SN P systems, modeling certain neurophysiological processes, and so on); the bibliography which closes the paper can be a good source of information for the reader.

In order to have a sort of a case study of research in this area, we end the paper with some results (obtained during the Seventh Brainstorming Week on Membrane Computing, Sevilla, February 2-6, 2009), concerning a new normal form for SN P systems. Details will be given in Section 6.

2 Prerequisites

We assume the reader to have some familiarity with (basic elements of) language and automata theory, e.g., from [73], as well as with basics of membrane computing, e.g., from [56], [71], and [81], and we introduce here only a few notations, as well as the notion of register machines, used in the proofs from Section 6 (based on paper [51]).

For an alphabet $V$, $V^*$ denotes the set of all finite strings of symbols from $V$, the empty string is denoted by $\lambda$, and the set of all nonempty strings over $V$ is denoted by $V^+$. When $V = \{a\}$ is a singleton, then we write simply $a^*$ and $a^+$ instead of $\{a\}^*, \{a\}^+$. For a regular expression $E$ we denote by $L(E)$ the regular language identified by $E$.

By $NFIN, NREG, NRE$ we denote the families of finite, semilinear, and Turing computable sets of (positive) natural numbers (number 0 is ignored); they correspond to the length sets of finite, regular, and recursively enumerable languages, whose families are denoted by $FIN, REG, RE$. We also invoke below the family of recursive languages, $REC$ (the languages with a decidable membership).

A register machine (see, e.g., [47]) is a construct $M = (m, H, l_0, l_h, I)$, where $m$ is the number of registers, $H$ is the set of instruction labels, $l_0$ is the start label (labeling an ADD instruction), $l_h$ is the halt label (assigned to instruction $\text{HALT}$), and $I$ is the set of instructions; each label from $H$ labels only one instruction from $I$, thus precisely identifying it. The instructions are of the following forms:
4. L. Pan, Gh. Păun, M.L. Pérez-Jiménez

- \( l_i : (\text{ADD}(r), l_j, l_k) \) (add 1 to register \( r \) and then go to one of the instructions with labels \( l_j, l_k \)),
- \( l_i : (\text{SUB}(r), l_j, l_k) \) (if register \( r \) is non-empty, then subtract 1 from it and go to the instruction with label \( l_j \), otherwise go to the instruction with label \( l_k \)),
- \( l_h : \text{HALT} \) (the halt instruction).

A register machine \( M \) computes (generates) a number \( n \) in the following way: we start with all registers empty (i.e., storing the number zero), we apply the instruction with label \( l_0 \) and we proceed to apply instructions as indicated by the labels (and made possible by the contents of registers); if we reach the halt instruction, then the number \( n \) stored at that time in the first register is said to be computed by \( M \). The set of all numbers computed by \( M \) is denoted by \( N(M) \). It is known that register machines compute all sets of numbers which are Turing computable, hence they characterize \( NRE \).

Without loss of generality, we may assume that in the halting configuration, all registers different from the first one are empty, and that the output register is never decremented during the computation, we only add to its contents.

We can also use a register machine in the accepting mode: a number is stored in the first register (all other registers are empty); if the computation starting in this configuration eventually halts, then the number is accepted. Again, all sets of numbers in \( NRE \) can be obtained, even using deterministic register machines, i.e., with the ADD instructions of the form \( l_i : (\text{ADD}(r), l_j, l_k) \) with \( l_j = l_k \) (in this case, the instruction is written in the form \( l_i : (\text{ADD}(r), l_j) \)).

**Convention:** when evaluating or comparing the power of two number generating/accepting devices, number zero is ignored.

3 Spiking Neural P Systems

We introduce here the SN P systems in the standard form (with non-extended rules):

An SN P system of degree \( m \geq 1 \) is a construct of the form

\[
\Pi = (O, \sigma_1, \ldots, \sigma_m, \text{syn, in, out}),
\]

where:

1. \( O = \{a\} \) is the singleton alphabet (\( a \) is called spike);
2. \( \sigma_1, \ldots, \sigma_m \) are neurons, of the form

\[
\sigma_i = (n_i, R_i), 1 \leq i \leq m,
\]

where:

a) \( n_i \geq 0 \) is the initial number of spikes contained in \( \sigma_i \);
b) \( R_i \) is a finite set of rules of the following two forms:
Spiking Neural P Systems

1. $E/a^c \rightarrow a; d$, where $E$ is a regular expression over $a$, $c \geq 1$, and $d \geq 0$;
2. $a^s \rightarrow \lambda$, for some $s \geq 1$, with the restriction that for each rule $E/a^c \rightarrow a; d$ of type (1) from $R_i$, we have $a^s \notin L(E)$;
3. $\text{syn} \subseteq \{1, 2, \ldots, m\} \times \{1, 2, \ldots, m\}$ with $(i, i) \notin \text{syn}$ for $1 \leq i \leq m$ (synapses between neurons);
4. $\text{in}, \text{out} \in \{1, 2, \ldots, m\}$ indicate the input and output neurons, respectively.

The rules of type (1) are firing (we also say spiking) rules, and they are applied as follows. If the neuron $\sigma_i$ contains $k$ spikes, and $a^k \in L(E), k \geq c$, then the rule $E/a^c \rightarrow a; d$ can be applied. The application of this rule means removing $c$ spikes (thus only $k - c$ remain in $\sigma_i$), the neuron is fired, and it produces a spike after $d$ time units (a global clock is assumed, marking the time for the whole system, hence the functioning of the system is synchronized). If $d = 0$, then the spike is emitted immediately, if $d = 1$, then the spike is emitted in the next step, etc. If the rule is used in step $t$ and $d \geq 1$, then in steps $t, t + 1, t + 2, \ldots, t + d - 1$ the neuron is closed (this corresponds to the refractory period from neurobiology), so that it cannot receive new spikes (if a neuron has a synapse to a closed neuron and tries to send a spike along it, then that particular spike is lost). In the step $t + d$, the neuron spikes and becomes again open, so that it can receive spikes (which can be used starting with the step $t + d + 1$).

The rules of type (2) are forgetting rules and they are applied as follows: if the neuron $\sigma_i$ contains exactly $s$ spikes, then the rule $a^s \rightarrow \lambda$ from $R_i$ can be used, meaning that all $s$ spikes are removed from $\sigma_i$.

If a rule $E/a^c \rightarrow a; d$ of type (1) has $E = a^c$, then we will write it in the simplified form $a^c \rightarrow a; d$.

In each time unit, if a neuron $\sigma_i$ can use one of its rules, then a rule from $R_i$ must be used. Since two firing rules, $E_1/a^{c_1} \rightarrow a; d_1$ and $E_2/a^{c_2} \rightarrow a; d_2$, can have $L(E_1) \cap L(E_2) \neq \emptyset$, it is possible that two or more rules can be applied in a neuron, and in that case, only one of them is chosen nondeterministically. Note however that, by definition, if a firing rule is applicable, then no forgetting rule is applicable, and vice versa.

Thus, the rules are used in the sequential manner in each neuron, but neurons function in parallel with each other.

The initial configuration of the system is described by the numbers $n_1, n_2, \ldots, n_m$ of spikes present in each neuron. During a computation, the system is described both by the numbers of spikes present in each neuron and by the state of each neuron, in the open-closed sense. Specifically, if a neuron is closed, we have to specify the number of steps until it will become again open, i.e., the configuration is written in the form $\langle p_1/q_1, \ldots, p_m/q_m \rangle$; the neuron $\sigma_i$ contains $p_i \geq 0$ spikes and will be open after $q_i \geq 0$ steps ($q_i = 0$ means that the neuron is already open).

Using the rules as suggested above, we can define transitions among configurations. Any sequence of transitions starting in the initial configuration is...
called a computation. A computation halts if it reaches a configuration where all neurons are open and no rule can be used. With any computation, halting or not, we associate a spike train, the binary sequence with occurrences of 1 indicating time instances when the output neuron sends a spike out of the system (we also say that the system itself spikes at that time).

In [32], with any spike train containing at least two spikes, the first two being emitted at steps $t_1, t_2$, one associates a result, in the form of the number $t_2 - t_1$; we say that this number is computed by $\Pi$. The set of all numbers computed in this way by $\Pi$ is denoted by $N_2(\Pi)$ (the subscript indicates that we only consider the distance between the first two spikes of any computation; note that 0 cannot be computed, that is why we disregard this number when estimating the computing power of any device).

This idea was extended in [66] to several other sets of numbers which can be associated with a spike train: taking into account the intervals between the first $k$ spikes, $k \geq 2$ (direct generalization of the previous idea), or between all intervals; only halting computations can be considered or arbitrary computations; an important difference is between the case when all intervals are considered and the case when the intervals are taken into account alternately (take the first interval, ignore the next one, take the third, and so on); the halting condition can be combined with the alternating style of defining the output.

The result of a computation can be defined also as usual in membrane computing, as the number of spikes present in the output neuron in the end of a computation – we have then to work with halting computations. It is also possible to consider SN P systems working in the accepting mode: we start the computation from an initial configuration, and we introduce in the input neuron two spikes, in steps $t_1$ and $t_2$; the number $t_2 - t_1$ is accepted by the system if the computation eventually halts.

Then, the spike train itself can be considered as the result of a computation. The halting computations will thus provide finite strings over the binary alphabet, the non-halting computations will produce infinite sequences of bits. If also an input neuron is provided, then a transducer is obtained, translating input binary strings into binary strings.

### 4 Example

We illustrate the previous definition with only one example; further constructions of SN P systems will be examined in Section 6. The system is given in a graphical form in Figure 1, following the standard way to pictorially represent a configuration of an SN P system, in particular, the initial configuration. Specifically, each neuron is represented by a “membrane” (a circle or an oval), marked with a label and having inside both the current number of spikes (written explicitly, in the form $a^n$ for $n$ spikes present in a neuron) and the evolution rules; the synapses linking the neurons are represented by arrows;
L ∈ FIN, L ⊆ B⁺, we have L{1} ∈ LSNP₁(rule⁺, cons⁺, forg⁺, bound⁺), and if L = {x₁, x₂, ..., xₙ}, then we also have \{0^i xᵢ | 1 ≤ i ≤ n\} ∈ LSNP₁(rule⁺, cons⁺, forg⁺, bound⁺).

(ii) The family of languages generated by finite SN P systems is strictly included in the family of regular languages over the binary alphabet, but for any regular language L ⊆ V⁺ there is a finite SN P system \( \Pi \) and a morphism \( h : V⁺ → B⁺ \) such that \( L = h^{-1}(L(\Pi)) \). Spiking Neural P Systems.

(iii) LSNP₁(rule⁺, cons⁺, forg⁺) ⊆ REC, but for every alphabet V = \{0, 1\}, besides the fact the output neuron will be identified by its symbol 0, it is also suggestive to draw a short arrow which exits from it, pointing to the environment. For any language \( L \subseteq V⁺, L \subseteq RE, \) there is an SN P system \( \Pi \) such that \( L = \text{Fin}(1 \Rightarrow (h(\Pi))) \).

These results show that the language-generating power of SN P systems is rather eccentric; on the one hand, finite languages \( \{0, 1\} \) cannot be generated, on the other hand, we can represent any RE language as the direct morphic image of an inverse morphic image of a language generated in this way. This eccentricity is due mainly to the restricted way of generating strings, with one symbol added at each computation step. This restriction does not appear in the case of extended spiking rules, of the form \( E/α^c → α^d; k \) that of \( p \geq 1 \) spikes can be produced when consuming \( c \) spikes of the neuron (we assume that \( c \geq p \)). In this case, a language can be generated by associating the symbol \( b_i \) with a step when the output neuron sends out \( i \) spikes, with an important decision to take in the case \( i = 0 \): we can either consider \( b_0 \) as a separate symbol, or we can assume that emitting 0 spikes means inserting \( λ \) in the generated string. Thus, we both obtain strings over arbitrary alphabets, not only over the binary one, and, in the case where we ignore the steps when no spike is emitted, a considerable freedom is obtained in the way the computation proceeds. This latter variant (with \( λ \) associated with steps when no spike exits the system) is considered below.

We denote by \( LSN^c P_m(rule⁺, cons⁺, prod⁺) \) the family of languages \( L(\Pi) \), generated by SN P systems \( \Pi \) using extended rules, with at most \( m \) neurons, each neuron having at most \( k \) rules, each rule consuming at most \( p \) spikes and producing at most \( q \) spikes. Again, the parameters \( m, k, p, q \) are replaced by * if they are not bounded.

The next counterparts of the results from Theorem 2 were proved in [10].

**Theorem 3.** (i) \( FIN = LSN^c P_1(rule⁺, cons⁺, prod⁺) \) and this result is sharp in the sense that \( LSN^c P_2(rule⁺, cons⁺, prod⁺) \) contains infinite languages.

(ii) \( LSN^c P_3(rule⁺, cons⁺, prod⁺) \) ⊆ REG ⊆ \( LSN^c P_4(rule⁺, cons⁺, prod⁺) \); the second inclusion is proper, because \( LSN^c P_3(rule⁺, cons⁺, prod⁺) \) contains non-regular languages; actually, the family \( LSN^c P_5(rule⁺, cons⁺, prod⁺) \) contains non-semilinear languages.

(iii) \( RE = LSN^c P_i(rule⁺, cons⁺, prod⁺) \).

**6 New Normal Forms**

A neuron \( σ_i \) (in the initial configuration of an SN P system) is characterized by \( n_i \), the number of spikes present in it, and by \( R_i \), its associated set of rules. An SN P system is said to be in the \( kR \)-normal form, for some \( k \geq 1 \), if there are at most \( k \) different sets \( R_1, ..., R_k \) of rules used in the \( m \) neurons of the system. An SN P system is said to be in the \( knR \)-normal form, for some \( k \geq 1 \),
if there are at most $k$ different pairs $(n_1, R_1), \ldots, (n_k, R_k)$ describing the $m$ neurons of the system.

We denote by $N_{\alpha}SNP_{k}(\beta)$ the families of all sets $N_{\alpha}(II)$ computed by SNP systems in the $k\beta$-normal form, for $\alpha \in \{2, gen, acc\}$, $\beta \in \{R, nR\}$, and $k \geq 1$, without forgetting rules, but with the spiking rules using the delay feature (we do not restrict the number of neurons, that is why SNP has the subscript $\ast$).

6.1 A 3R-Normal Form Result

We are going now to prove one the new normal form result mentioned in the Introduction: SNP systems with only three different sets of rules are universal when generating numbers encoded in the first two spikes of the spike train.

**Theorem 4.** $NRE \subseteq N_2SNP_{3}(3R)$; the converse inclusion is straightforward and (3R) can be simulated in the first two spikes of the spike train.

**Proof.** We show that $NRE \subseteq N_2SNP_{3}(3R)$; the reverse inclusion is straightforward and (3R) can be simulated in the first two spikes of the spike train.

We construct an SNP system $\Pi = (\mathcal{M}, l)$ with the initial label, all neurons are initially empty, with the exception of the neuron associated with the label $l$, which contains 2 spikes. Each register $M$ can be simulated already when proving that a class of SNP systems is universal.

We show that

$$R_1 = \{a(aaa)^*/a \rightarrow a; 0, \; a \rightarrow a; 1\},$$

$$R_2 = \{aa \rightarrow a; 0, \; a \rightarrow a; 1\}.$$ 

The $ADD$ module simulates Church-Turing (3R) in a set of neurons $l = \sigma_1, l_2, \ldots, l_9$. The neurons $l_1, l_2, \ld{4}$ are output neurons, which contain $2$ spikes initially, and the neurons $l_3, l_4, \ld{9}$ are control neurons, which simulate the rules $R_1, R_2$.

We construct an SNP system $\Pi = (\mathcal{M}, \sigma)$ with the initial label, all neurons are initially empty, with the exception of the neuron associated with the label $\sigma$, which continues having one spike. The simulation of the ADD instruction is correctly completed.

Fig. 3: Module $ADD$, simulating $l_i : (ADD(r), l_j, l_k)$

We consider the following three sets of rules:

**The ADD module simulates Church-Turing (3R).** We consider a $\mathcal{R}$-machine $M$ as in Figure 3. No the property can be applied in the system $\sigma$, by the construction of the $\mathcal{R}$-system, the universal, specific $\mathcal{R}$-machine $M$ is constructed using $\mathcal{R}$ and $SU$ to simulate the instruction of which all spike is in one and the same neuron, which provides the result (the final spike is sent to the next neuron).

Specifically, neurons $l_1, l_2, \ld{9}$ simulate $ADD$ and $SUB$ instructions (in the instruction $i$, which will spike in the second neuron, which provides the result (the final spike is sent to the next neuron).

We will show that neuron $l_{14}$ can non-deterministically choose either rule to use as both of them are non-deterministically enabled in the initial configuration. The neuron in charge of the neuron $n_1$ simulates the neuron $n_1$ in the first neuron, but it can only simulate the neuron $n_1$, which contains $2$ spikes initially, with the exception of the neuron associated with the label $l_j$.

The neuron $l_4$ receives one spike and $l_{12}$ continues having one spike. The neuron $l_9$ contains now a number of spikes of the form $3n + 3$, for some $n \geq 0$ (initially we had two spikes here, hence $n = 0$) and no rule is enabled. In the next step, this neuron receives one further spike, and the first rule is fired (the number of spikes is now $3(n + 1) + 1$). All neurons $l_j$ and $l_{111}, l_{112}$ receive one spike. The last two neurons send back to $l_{10}$ one spike each, hence the number of spikes in this neuron will be again congruent with 2 modulo 3, as at the beginning. Thus, the neuron associated with the label $l_j$ has been activated.

If neuron $l_{14}$ uses the rule $a \rightarrow a; 1$, then $l_{12}$ receives two spikes at the same (after one time unit) time and this branch remains idle, while neurons $l_{18}, l_{110}, l_{113}, l_{114}$ behave like neurons $l_7, l_9, l_{111}, l_{112}$, and eventually $l_{13}$ is activated and the number of spikes from $l_{10}$ returns to the form $3s + 2$, for some $s \geq 0$.

The simulation of the ADD instruction is correctly completed.
The study of SN P systems is rather recent, but it is already clear that this area is worth investigating, not only because of the importance of neural computing based on spiking, but also because of the mathematical (computability) interest and the possible applications. The reader is advised to watch [81] for future developments in this area.

Acknowledgements

The work of L. Pan was supported by National Natural Science Foundation of China (Grant Nos. 60674106, 30870826, 60703047, and 60833113), Program for New Century Excellent Talents in University (NCET-05-0612), Ph.D. Programs Foundation of Ministry of Education of China (20060487014), Chenguang Program of Wuhan (200750731262), HUST-SRF (20072015A), and Natural Science Foundation of Hubei Province (2008CDB113 and 2008CDB180). The work of the last two authors was supported by Project TIN2006-13452 of the Ministerio de Educaci´on y Ciencia of Spain and by Project of Excellence with Investigador de Reconocida Valia, from Junta de Andalucia, grant P08 – TIC 04200.

Note: The bibliography which follows includes several papers on SN P systems published first in the proceedings volumes of the Brainstorming Week on Membrane Computing, with indications of the form “BWMC2007” with the obvious meaning, then specifying the pages. The four brainstorming volumes cited in this way are:

References


81. The P Systems Website: http://ppage.psystems.eu.