11 Years of P-Lingua: A Backward Glance

Ignacio Pérez-Hurtado, David Orellana-Martín, Miguel Á. Martínez-del-Amor, Luis Valencia-Cabrera, Agustín Riscos-Núñez, Mario J. Pérez-Jiménez

Research Group on Natural Computing
Universidad de Sevilla
Avda. Reina Mercedes S/N, 41012 Sevilla, Spain
{perezh,dorellana,mdelamor,lvalencia,ariscosn,marper}@us.es

Summary. In 2008, P-Lingua was born. The Research Group on Natural Computing worked on the development of simulation tools since the beginning of Membrane Computing. However, back in 2007, researchers from the group set out to the ambitious journey of creating a generic simulation framework for P systems. P-Lingua has evolved since then, offering more flexibility and a wider range of supported models. Many applications have also branched from this software project. In this paper, we briefly survey the evolution of P-Lingua to date, some of the associated applications, and prospective paths for upcoming challenges in the research area.

Key words: P systems, Simulation Software, P-Lingua

1 Introduction

Since the dawn of Membrane Computing, software tools to perform P system simulations have been developed. In fact, simulators have accompanied the evolution of P system models [34, 3, 30], given that they are the best way to aid in experimental validation and in formal verification. In the early years of the research area, simulators were conceived for specific models, leading to a plethora of interfaces and platforms, each one very different to each other. In this context, developing an uniform simulation framework was a need. This was the seed that led to P-Lingua project [36], which was the major work presented at Ignacio Pérez-Hurtado’s doctoral dissertation [30]. Further theses developed at the core of the Research Group on Natural Computing were carried out with specific contributions to P-Lingua [23, 8, 33, 15, 24].

P-Lingua aims at flexibility and extensibility to simulate P systems. The term itself is employed at different extents, so it is usually confused. P-Lingua can be used to denote:

• the software project: P-Lingua is an open-source software project published under GNU GPL license. All the developed code is freely available to the
community, and we hope that it helps other researchers to run their models and to develop new simulators. It includes the parser, the generation of files, the simulation framework.

- **the simulation framework**: more commonly known as pLinguaCore, it is Java framework that aims at simulating P systems in a flexible and extensible way. Several design pattern of object oriented programming were employed to help developing new simulations in a clear and clean manner.

- **the programming language**: instead of having an specific GUI to enter the information of the models, P-Lingua enables P-system designers to write their models into plain-text files in a very similar way (i.e. with similar syntax) than they write them down in paper or whiteboard.

- **the files with .pli extension**: files including models described in P-Lingua programming language have the .pli extension, and they are the input for the simulation software framework.

- **the command-line tool**: pLinguaCore can be employed from the Terminal (Linux) or Cmd (Windows) using a command-line tool with specific parameters to define the input .pli files, the output files (when used as a compiler) or the simulation options (when used as a simulator).

- **the translator/compiler to binary files**: efficient, parallel simulators came after P-Lingua, looking for small runtimes. They harness the P-Lingua parsing engine in order to process .pli files, and specific binary files are generated for them.

In this paper, we look back and survey the evolution of P-Lingua through the 11 years of developments and efforts in the research community. In Section 2, a gentle introduction of the P-Lingua programming language is made. In Section 3, a glance at the evolution of P-Lingua software is provided. Graphical User Interfaces (GUIs) based on P-Lingua are revisited in Section 4. The connection of P-Lingua to parallel simulators is summarized in Section 5. Conclusions and plans to future developments are displayed in Section 6.

## 2 P-Lingua Description Language

The simulation of P system computations over conventional silico machines has been done from the beginning of Membrane Computing.

The general structure of a simulator should contain three main modules:

- **The input**: A way to define the P system to be simulated as well as the initial configuration.

- **The simulation core**: The implementation of one or more simulation algorithms to reproduce the behavior of the P system to be simulated.

- **The output**: The representation of relevant information to the final user, it could be the representation of the final configuration, one computation, the whole computation tree, etc. In some cases, the information should be processed, developing statistics and providing graphics.
The simulation core usually depends on the software/hardware architecture to be used, highly parallel architectures such as CUDA use simulation algorithms very different than the ones used with approximations such as developing in Java or C/C++. The output depends on the final goal of the simulator and the final user background, for example, one application for simulating ecosystems should show information about populations, environment conditions such as weather, etc. Applications for membrane computing should show the generated computation, multisets of objects, executed rules, etc.

Regardless of the type of application, the definition of the P system is mandatory and it is a common module for all the simulators. One approximation is to hard-code the P system definition, but it could lead to errors and it is difficult to debug. Using graphical user interfaces introduces dependencies with the technology and operating systems, using GUIs for the P system definition also implies a maintenance effort and they can become obsolete over the years.

Thus, the solution used in P-Lingua is to use a programming language to define P systems. The files written in P-Lingua are close to the standard scientific notation, minimizing the effort of writing them. P-Lingua code is parametric using variables and iterators, modular using reusable source blocks and it can be translated to other formats by using parsers. We could say that the main goal of the P-Lingua language is to minimize the time from the blackboard to the computer.

As example, next is a P-Lingua code defining a family of recognizer P systems to solve the SAT problem in the framework of cell-like P systems with active membranes and elementary division rules based on the solution presented in [31]. It shows the main characteristics of the language. An exhaustive description of the language can be found in the literature [4, 5, 22, 30] and it is out of the scope of this review.

```plaintext
@model <membrane_division>
def Sat(m,n)
{
  /* Initial configuration */
  @mu = [[ ]'2]'1;

  /* Initial multisets */
  @ms(2) = d{1};

  /* Set of rules */
  [d{k}]'2 --> +[d{k}]-[d{k}] : 1 <= k <= n;

  { 
    +[x{i,1} --> r{i,1}]'2;
    -[nx{i,1} --> r{i,1}]'2;
    -[x{i,1} --> #]'2;
    +[nx{i,1} --> #]'2;
  } : 1 <= i <= m;
}
```
\[
\begin{align*}
\{ & \quad +[x_{i,j} \rightarrow x_{i,j-1}] \quad '2; \\
& \quad -[x_{i,j} \rightarrow x_{i,j-1}] \quad '2; \\
& \quad +[nx_{i,j} \rightarrow nx_{i,j-1}] \quad '2; \\
& \quad -[nx_{i,j} \rightarrow nx_{i,j-1}] \quad '2; \\
& \quad : 1 \leq i \leq m, 2 \leq j \leq n; \\
\}
\]

\[
\{ & \quad +[d_{k}] \quad '2 \rightarrow \Box d_{k}; \\
& \quad -[d_{k}] \quad '2 \rightarrow \Box d_{k}; \\
& \quad : 1 \leq k \leq n; \\
\}
\]

d_{k} \quad '2 \rightarrow \Box [d_{k+1}] : 1 \leq k \leq n-1;

[r_{i,k} \rightarrow r_{i,k+1}] \quad '2 : 1 \leq i \leq m, 1 \leq k \leq 2n-1;

[d_{k}] \rightarrow d_{k+1} \quad '1 : n \leq k \leq 3n-3;

[d_{3n-2}] \rightarrow d_{3n-1}, e \quad '1;

e \quad '2 \rightarrow \Box [c_{1}];

[d_{3n-1}] \rightarrow d_{3n} \quad '1;

[d_{k}] \rightarrow d_{k+1} \quad '1 : 3n \leq k \leq 3n+2m+2;

+[r_{1,2n}] \quad '2 \rightarrow \Box [r_{1,2n}];

-[r_{1,2n}] \rightarrow r_{i-1,2n}, r_{i,2n-1} \quad '2 : 1 \leq i \leq m;

r_{1,2n} \quad '2 \rightarrow \Box [r_{0,2n}];

-[c_{k}] \rightarrow c_{k+1} \quad '2 : 1 \leq k \leq m;

+[c_{m+1}] \quad '2 \rightarrow \Box [c_{m+1}];

[c_{m+1}] \rightarrow c_{m+2}, t \quad '1;

t \quad '1 \rightarrow \Box t;

+[c_{m+2}] \quad '1 \rightarrow \Box [Yes];

[d_{3n+2m+3}] \quad '1 \rightarrow \Box [No];
\}
/* End of Sat module */

/* Main module */
def main()
{

/* Call to Sat module for m=4 and n=6 */
call Sat(4,6);

/* Expansion of the input multiset */
@ms(2) += x_{1,1}, nx_{1,2}, nx_{2,2}, x_{2,3}, nx_{2,4}, x_{3,5}, nx_{4,6};

3 P-Lingua Simulation Software

PLinguaCore [36] is a GNU GPL library written in Java for parsing P-Lingua files, simulate computations and translate input P-Lingua files to other file formats. The library detects errors in the file and reports them such as a regular compiler tool. For each supported P system variant, several simulation algorithms are included. Eventually, the library translates files, which define a P system, between formats, for instance, from P-Lingua format to binary format.

Each version of the library is associated with an extension of the programming language and simulation engine to cover more types of P systems:

- **pLinguaCore 1.0**: The initial version. It was able to define active membrane P systems with division rules. pLinguaCore was in very early state. [4].
- **pLinguaCore 2.0**: Several cell-like P system models and built-in simulators for each supported model were added [5].
- **pLinguaCore 2.1**: Support for tissue-like P systems with division rules were added [22], plus fixed bugs.
- **pLinguaCore 3.0**: The simulation algorithm called DCBA for Population Dynamics P systems (PDP systems) was added [19], along with a new binary output file for PDP systems [17]. In this version, stochastic P systems are discontinued. Some bugs were fixed.
- **pLinguaCore 4.0**: Support for Spiking Neural P systems was added [13]. Moreover, Tissue-Like P systems with Cell Separation Rules was also added [29]. Some bugs were also fixed.

The development of pLinguaCore continued as part of MeCoSim tool, as depicted in next section.

4 GUIs for P-Lingua

As highlighted from the beginning of this survey paper, from its very beginning P-Lingua aimed to provide a general-purpose framework for the description and simulation of P systems. This would emerge as a greatly valuable tool for P systems experts to work with P systems by specifying them in a similar way that they would do with paper and pen in their theoretical studies, as shown in the previous sections.

However, in order to provide these users with more usable mechanisms to experiment with models based on P systems, they would require more visual elements to help them save time and clarify the evolution of the systems under study.

Additionally, membrane systems proved to be useful as a modelling framework for real problems in areas ranging from biology to economy, and a second group of
interest appeared into the scene: people not necessarily familiar with P systems, but using the models designed by experts to help understanding and managing the phenomena they were interested in. These kinds of *end users* would need higher level tools to handle the systems but not entering into technical details, and here some graphical interfaces would also play an essential role.

Among these visual interfaces to handle P systems based on the P-Lingua framework, it is worth mentioning two especially relevant ones: MeCoSim and MeCoGUI, described within the following subsections.

### 4.1 MeCoSim

From the very beginning in 2009-2010 [28], MeCoSim was designed with the two-fold intention expressed above, to help P systems designers with the modelling and verification tasks, and also allowing end users to handle solutions based on membrane systems [33, 30].

For the first goal, this software was always supported by the parsing, debugging and simulation engine given by P-Lingua, while the latter objective required further development to provide some mechanism where P systems designers could easily prepare end-user applications by a relatively simple configuration. Thus, MeCoSim was initially designed to enable the user defined customized interfaces, with inputs, outputs, charts, etc., adapted to each model or solution given by a family of P systems, expressed in MeCoSim language. This enables the users to enter their data for different initial conditions, instantiating the desired P systems within the family.

Apart from the initial goals, this software environment had from its conception two essential objectives: the flexibility required for the definition of any kind of custom applications and variety of P systems accepted, and the extensibility to increase its initial functionalities through the development using its plugins architecture and the integration with external packages. Thus, the seminal version of the software was enriched by the integration with different tools for the property extraction and verification of P systems [12, 9] and the development of plugins for graph-based problems, definition of propositional formulas, extensions of the language accepted by the generation of parameters, connection with external simulators based on Spin [10] and several other new features for designers and end users. Extensive documentation, case studies, explicative videos, etc. about everything related with MeCoSim can be found at [35].

### P-Lingua inside MeCoSim

Given the nature of this survey paper, devoted to P-Lingua, it is worth mentioning that MeCoSim is fully compatible with any version of P-Lingua, and uses pLinguaCore as its engine for parsing and simulation (additionally, it allows the simulation through other external simulators, but its most extended use makes use of pLinguaCore). Thus, MeCoSim will mainly expose the types of computing
models, with their parsers and simulators, provided by the version of pLingua we decide to deploy with MeCoSim. Therefore, depending on which version of pLinguaCore (see previous sections) we use, a different set of models and simulators will be available to use in MeCoSim.

However, the scope of MeCoSim and the collaboration with different external use brought new needs in terms of types of models to cover in terms and features of the P-Lingua language itself. The time constraints of P-Lingua project and MeCoSim-based applications used by different research groups were different, and was not possible to make both goals fit together with a sound coordination. Hence, a new version of pLinguaCore, let us say a fork from the project, started to grow inside MeCoSim project, and MeCoSim integrated this new version of pLinguaCore inside its distribution, progressively diverging from pLinguaCore 4. Despite this divergence, however, all models available in pLinguaCore 4 were also available inside pLinguaCore version used by MeCoSim. In this sense, we could informally consider this internal version as some unofficial pLinguaCore 4.5, including its predecessor plus some additional features in the language and some further models (with their parsers and simulators).

Among the main new computing models added within P-Lingua version inside MeCoSim we can find:

- Simple kernel P systems[9].
- Probabilistic Guarded (Scripted) P systems[7, 6].
- Evolutinal-communication P systems[25].
- Fuzzy Reasoning Spiking Neural P systems[14].
- Cell-like Spiking Neural P systems[32].
- P systems with minimal cooperation.
- Tissue P systems with promoters.

4.2 MeCoGUI

MeCoGUI is a minimal graphical user interface for P-Lingua which simplifies the process of parsing input parameters and generating results files. The current version encapsulates a version of P-Lingua which integrates Probabilistic Guarded Scripted P (PGSP) systems, a novel Membrane Computing framework tailored for modelling population dynamics [8, 7].

PGSP systems have been used to study the behaviour and conservation trends of *Pieris napi oleracea*, (P. n. oleracea, for short). This species, commonly known as mustard white butterfly, is native from eastern North America. The aim of the model is to predict population growth trends and evolutionary responses of several genotypes of this species, considering the concurrent invasions of the exotic plant garlic mustard (*Alliaria petiolata*), and another exotic, *Cardamine pratensis* [6].

4.3 pLinguaPlugin

pLinguaPlugin was an Eclipse plug-in which provided an end-user interface for using pLinguaCore. Hence, all services performed by pLinguaCore, such as com-
pilation of P system files, error detection, file translation and P system simulation were accessed in an intuitive and user-friendly way by means of pLinguaPlugin. It also provided ways to visually display configurations and modify specifications of P systems on real time. This component is free software published under GNU GPL license, and it is only compatible for pLinguaCore 2.0. After that, the support was discontinued. The plugin is still available at the P-Lingua website [36].

4.4 P-Lingua web analyzer

A web analyzer for P-Lingua was developed in PHP and deployed at the P-Lingua website [36]. An user can upload a .pli file, select to either simulate until a halting configuration or a certain number of steps, and by using pLinguaCore as a backend, simulate the model. When the simulation is finished, the output is shown on the website as a tree, where one can select certain configurations, and inside them, depict membranes, objects and rules executed. Moreover, both the simulation and the output data, dumped by pLinguaCore, is shown on the window. This tool was developed as a final project of computer engineering by J. González-Pareja in 2012.

5 Connection to Parallel Simulators on GPUs

pLinguaCore is the software framework devoted to perform P system simulations. As mentioned above, it is developed using object-oriented paradigm and considering design patterns. The objective of it is to offer readability, flexibility and extensibility of the framework for the research community. However, this comes at the cost of performance. Thus, a branch of P-Lingua project was the development of efficient, parallel simulators for P systems. This new project was later codenamed PMCGPU (Parallel simulators for Membrane Computing on the GPU) [37].

The platform employed to accelerate the simulation of P systems was the GPU (Graphics Processing Units), given that, since the introduction of CUDA, it is easily programmable and leverages a high degree of parallelism (from hundreds to thousands of processing cores) [11]. Multicore CPUs were also used to speedup simulations at a lower level through OpenMP. Several models have been simulated on GPUs so far, as surveyed in [16].

In the next subsections, we will show the two CUDA-based simulators that used input files generated by P-Lingua: PCUDA (for active membrane systems) [2] and ABCD-GPU (for PDP systems) [17]. The specific details of this input, binary formats can be consulted in [23].

5.1 Binary files for Active Membranes

The first GPU-based simulator was first presented in the 7th BWMC (2009) [21], focused on recognizer P systems with active membranes and elementary division
This was a generic simulator that required a description of the P system to be simulated as input, reproduced only one computation (confluent condition was assumed), and output the description of the configurations in a syntax close to pLinguaCore output module. In this sense, P-Lingua was employed as a compiler: parsing a P-Lingua file describing a P system with active membranes, and generating a compressed binary file with the unwrapped, processed information to be injected to PCUDA simulator. This binary file was carefully designed to compress the amount of bits dedicated to define (in this order): the alphabet, the membrane structure, the rules, and the initial configuration. The header of the file included information about the precision (amount of bits) employed for each part.

pLinguaCore was then extended (version 2.0) to support the generation of a specific binary format devoted to only one P system variant (active membranes). P-Lingua command line was extended distributed alongside the PCUDA simulator [37], so that users were able to generate binary files from .pli files by typing: java -jar pparser.jar inputFile.pli inputFile.bin. If the binary file is generated without errors, it can be used as input for the simulator, plus providing some extra information: number of membranes expected to be generated by the model (upper bound), number of objects (symbols) defined in the alphabet, number of threads to be launched on the GPU for each membrane.

5.2 Binary files for PDP systems

PDP systems were also considered for parallel simulators, given the real requirement of efficiency by the applications (named ABCD-GPU). The first version of ABCD-GPU was implemented in OpenMP to harness multicore CPUs [18], and was later extended for GPUs [20]. However, these initial versions aimed at testing the concept of executing DCBA algorithm [19] in parallel, and in order to stress the simulators, randomly generated PDP systems were used as input. They were created inside the simulator, so there was no need to input files. In [17], ABCD-GPU was extended with a input module for binary files. The binary format design follows the same concept than the one used for PCUDA. A header helped to establish the precision (bits) used for each part.

The full support of binary file generation for PDP systems came with pLinguaCore 3.0. The command line was again extended, and the user could parse .pli files and generate binaries by:

```
java -jar pLinguaCore.jar plingua inputFile.pli -bin inputFile.bin
```

If the binary is generated without errors, the file can be used to feed ABCD-GPU, plus indicating the following parameters: number of simulations to run, number of time steps to simulate, number of steps per cycle, and amount of cores to be used for the OpenMP version (when employed).
6 Conclusions and Future Work

P-Lingua is a software project devoted to provide tools to simulate P systems. Developing and maintaining this flexible project is of huge complexity. The wide variety of membrane systems is as large as the imagination of Membrane Computing researchers, and normally some models differ from others in many aspects. Gathering all together under the same software umbrella is perhaps not realizable, but using all the techniques and methods of software engineering might help to other developers. P-Lingua can be seen therefore not as a all-in-one software, but as a key stone to ease the way of developing new simulators.

Future version of P-Lingua are under development, and will be focused not only on flexibility, but also on efficiency. P-Lingua 5 [27] is being re-engineered from scratch using C++ object-oriented programming language, that will help to achieve lower runtimes when dealing with large P-Lingua files. This will open new scalable opportunities for next challenges in Membrane Computing.

P-Lingua is very alive, and we look forward to collaborations in order to extend and evolve it.

Acknowledgments

The authors acknowledge the support from the research project TIN2017-89842-P (MABICAP), cofinanced by “Ministerio de Economía, Industria y Competitividad” (MINECO) of Spain, through the “Agencia Estatal de Investigación” (AEI), and by “Fondo Europeo de Desarrollo Regional” (FEDER) of the European Union.

References


37. The PMCGPU project website. [http://sourceforge.net/p/pmcgpu](http://sourceforge.net/p/pmcgpu)