Biological Implementation of Algorithms

IPN-UNAM iGEM team

January 5, 2007

1 Problem Statement

The project is based on the systematic implementation of some algorithms using the biobricks provided by iGEM, as well as the development of new ones. The algorithms we will focus on are based on:

1. Models based on cellular automata

2. Reaction-diffusion mechanisms

3. Game-theoretic methods

We intend to combine both theoretical and experimental perspectives. On the theoretical side, we believe formal methods and models drawn from computer science such as Pi-Calculus [9] and state transition systems [10] can be used to express and test possible prototypes of biological systems before costly and potentially defective implementations. On the experimental side, standard techniques as those promoted by the iGEM project [3] can be applied to build quickly real biological models from the formal ones.

We provide in what follows a more detailed description of each one of these three algorithms.

1.1 Models based on cellular automata

There is an increasing interest in the so called unconventional computing via cellular automata [11]. We have developed several approaches to the design of logical gates and other specific aspects [6, 7, 8]. In particular, simulations related to rule 110 are very promising and it would be an important first step to be able to implement these simulations, so far made in a standard computer, in a cellular-based system. In principle, all the rules needed for the automata can be realized using the biobricks already available.

1.2 Reaction-diffusion mechanisms

Not completely unrelated with the previous issue is the implementation of some reaction-diffusion based algorithms (see [1, 5]). Most of them arise naturally in chemical systems and have been already implemented successfully (the BZ reaction being a good example). The "translation" of these implementations into a biological language would be not only of great practical importance, but could also shed light into fundamental theoretical issues related to the actual behavior of genetic regulatory networks. In [2] we develop a model for a genetic regulatory network for which we are implementing a hybrid model, meaning the coupling of a cellular automaton with a reaction-diffusion system. This and other proposals for simpler networks in root and leaves have been implemented as part of our research work and we believe it is possible to mimic them in bio-engineered networks to test the validity of our model and perform different experiments.

1.3 Game-theoretic methods

It seems very natural trying to program a population of bacteria (e.g. E. Coli) with some specific strategy and other population with a corresponding one. Our goal is to actually realize in a biological system that can be manipulated several evolutionary principles: Nash equilibria and evolutionary stable equilibria. We could also verify the theoretical proposal in classical games such as the hawk-dove, the prisoner's dilemma, etc. But we also expect to be able to obtain new insights about the emergence patterns in collective behavior (see for instance [4])

References

[1] Andrew Adamatzky, *Computing in Reaction-Diffusion and Excitable Media: Case Studies of Unconventional Procedures.* Molecular Computing. T. Sienko, A.

Adamatzky, N. Rambidi and M. Conrad, eds. MIT Press, 1999.

[2] Espinosa-Soto Carlos, Pablo Padilla-Longoria and Elena R. Alvarez-Buylla. 2004. *A gene regulatory network model for cell-date determination during Arabidopsis thaliana flower development that is robust and recovers experimental gene expression profiles.* The Plant Cell, 16: 2923-2939.

[3] Web site: http://parts2.mit.edu/wiki/index.php/Main_Page

[4] Martin A. Nowak, Evolutionary Dynamics,Belknap Press,2006.

[5] Nicholas G. Rambidi, *Chemical-Based Computing and Problems of High Computational Complexity: The Reaction-Diffusion Paradigm.* Molecular Computing. T. Sienko, A. Adamatzky, N. Rambidi and M. Conrad, eds. MIT Press, 1999.

[6] Genaro Juárez Martínez, Harold V. McIntosh and Juan C. Seck Tuoh Mora, *Production of gliders by collisions in Rule 110*, Lecture Notes in Computer Science 2801, 175-182, 2003.

[7] Genaro Juárez Martínez, Harold V. McIntosh and Juan C. Seck Tuoh Mora, *Gliders in Rule 110*, International Journal of Unconventional Computing, Vol. 2, Num. 1, pp. 1-49, January 2006.

[8] Genaro Juárez Martínez, Harold V. McIntosh, Juan C. Seck Tuoh Mora and Sergio V. Chapa Vergara, *Rule 110 objects and other constructions based-collisions*, International Journal of Unconventional Computing, to be published, 2006.

[9] R. Milner, J. Parrow and D. Walker, *A calculus of mobile processes, parts I and II*, Information and Computation, 100(1):1–77, 1992.

[10] Colin Stirling, *Modal and Temporal Properties of Processes*, Springer-Verlag 2001.

[11] Stephen Wolfram, *A New Kind of Science*, Wolfram Media, Inc., Champaign, Illinois, 2002.

2 Expected Outcome

- Working prototypes of the above-mentioned algorithms.
- Software related with this project. In particular, we have a preliminary version of a program related to cellular automata based on rule 110: OSXLCAU21. We intend to produce software applying the formal techniques and software tools mentioned in the first section.
- At least a couple of research papers from each of the algorithms for publication in specialized conference proceedings and/or journals.
- Curriculum material for the following first degree and MSc courses: Cell and Molecular Biology (I, II and III), Biotechnology I, Genetics I, Evolution II, all in Biology, and Advanced Seminar (Formal Languages in Biological Systems) in Computer Science.
- Some finished semester projects from the previous courses.
- At least 3 first degree dissertations in either Biology or Computer Science.

3 Schedule

March–May 2007. Overall design of algorithms completed.

May–November 2007. First working prototypes and related software.

November 2007. Presentation of results in the iGEM Jamboree.

December 2007. First papers submitted.

End of 2007–beginning 2008. Curriculum material on line.

February and May 2008. First degree dissertations ready.

4 Use of Funds

We present in the following a list of items to be paid with the funds provided. Costs appear after each entrance.

- Hardware
  - 2 servers. \$6,400
  - 3 “heavy” duty computers. \$6,600
  - 6 “standard” PC. \$6,000
- Software
  - 9 licenses for Vista. \$2,700
  - 9 licenses for MS Office \$4,500
  - 1 licenses for SQL server. \$6,000
  - 13 licenses for Microsoft .net
  - 1 license for MacVector (for sequential alignment and nucleotide design). \$3,500
- Other Equipment
  - Laminar flow bell. \$10,000
  - Photo-documentation system DigitDoc-it (including one printer). \$5000
- Reactive substances (see attached list). \$20,000
- Travel and accommodation
  - iGEM Jamboree (six attendants). \$12,000
  - 2007 Microsoft eScience Workshop (two attendants). \$5,000
  - Expenses for 2 visiting researchers. We intend to invite Luca Cardelli and Andrew Phillips from Microsoft Research (the latter has already attended our School on Formal Languages in Biological Systems). \$5,000
- Salaries (scholarships for students)

- 2 PhD students for a year. $20,000
- 7 first degree students for a year. $7,280

  Note: none of them has any other funding for their studies at the moment.

5 Dissemination and Evaluation

Peer-review evaluation is fundamental and for that reason we intend to submit as many papers as possible to conference and journal committees. Those papers, if published, will be the best way of disseminating our results.

The iGEM Jamboree is another source of feedback from colleagues working in the same subject and our team will attend the 2007 edition. The iGEM project maintains a wiki site where results from the participant teams are available.

Additionally, we intend to make available on line the curriculum material, dissertations and software both in our own websites and in the MSDN Academic Alliance Repository, if this material is considered suitable for that site. Software will be available in the terms stated in the BSD license or any similar license policy.

Finally, we will be looking forward to presenting our project at the 2007 Microsoft eScience Workshop.

6 Other Support

Both institutions, UNAM and IPN, will contribute to the project in different ways. In particular, we can count on the following:

- Funds from a closely related on-going project called Formal Languages in Biological Systems. Some of the expenses for speakers for a workshop on this subject will be paid for the project (up to $5,000). These funds are provided by the National University of Mexico (UNAM).
- The use of 2 PC's and 1 multi functional printer belonging to the same project.

- Another 4 PC's belonging to the Faculty of Science, UNAM.
- Toxic waste managing services from the Faculty of Science, UNAM.
- The use of the laboratory for molecular biology in the Faculty of Science which most of the equipment required for the project, except that requested in this proposal.

7 Qualifications of Principal Investigators

Juan S. Aranda Barradas. PhD in Biology (IPN). Lecturer at the Posgraduate and Research School in Biology, National Polytechnical Institute (IPN). Teaching and research in Biology.

Arturo Becerra Bracho. PhD in Biology (UNAM). Lecturer at the Department of Biology, Faculty of Science, National University of Mexico (UNAM). Teaching and research in Genetics and Evolution.

Francisco Hernández-Quiroz. PhD in Computer Science (Imperial College, London). Lecturer at the Department of Mathematics, Faculty of Science, National University of Mexico (UNAM). Teaching and research in Logic, Formal Languages and Programming Languages.

Genaro Juárez Martínez. PhD in Computer Science (IPN). Lecturer at the School of Computing, National Polytechnical Institute (IPN). Teaching and research in Computer Science, specifically in cellular automata.

Pablo Padilla Longoria. PhD in Mathematics (Courant Institute of Mathematical Sciences). Researcher at the Institute for Applied Mathematical Research, National University of Mexico (IIMAS-UNAM). Teaching and research in Mathematical Biology.

Rosaura Palma Orozco. MSc in Mathematics (Cinvestav). Lecturer at the School of Computing, National Polytechnical Institute (IPN). Teaching and research in Computer Science.