First Joint Meeting between the RSME and the AMS

Sevilla, June 18–21, 2003

Abstracts

Session 5

Biomolecular Mathematics

Organizers:

Fernando Guzmán (SUNY at Binghamton)
Thomas J. Head (SUNY at Binghamton)
Carlos Martín-Vide (Universitat Rovira i Virgili)
Mario Pérez (Universidad de Sevilla)

Index of Abstracts

(In this index, in case of multiple authors, only the speaker is shown)

Synonymous codon bias and gene expression

Alessandra Carbone (Institut des Hautes Études Scientifiques)

Proteins are formed out of 20 amino-acids which are coded in triplets of nucleotides, called codons. The four nucleotides (A,T,C,G) define 64 codons used in the cell. Codons are not uniformely employed in the cell but, on the contrary, certain codons are preferred and we speak about codon bias. There are several kinds of codon biases and some of them are linked to specific biological functions. I will introduce a simple algorithm based on sequence analysis to detect dominating codon bias in prokaryotic and eukaryotic organisms. The algorithm is based on a precise mathematical formulation of the problem. Predictions on gene expression levels follow for fast growing organisms.

Test tube systems

Rudolf Freund (Technische Universität Wien)

Test tube systems were introduced as biological computer systems based on DNA molecules and soon gave rise to various theoretical models based on rules reflecting operations on DNA molecules. Test tube systems based on the splicing operation or on the operations of cutting and pasting (or recombination, respectively) with simple filtering and redistribution strategies were found to be universal computational mechanisms. Using specific locally testable input and output filters, test tube systems based on the splicing operation or on the operations of cutting and recombination with only two tubes already gain universal computational power, whereas test tube systems with only one test tube can generate only regular languages. Moreover, co-operating distributed test tube systems with only two tubes and communicating those objects having undergone exactly three operations (splicing or cutting/recombination, respectively) have universal computational power, too.

Mathematics of gene assembly

Tero Harju (University of Turku)

The ciliates are complex single celled organisms that have a unique genetic feature of nuclear dualism: they possess two types of nucleus - the macronucleus, which provides for the RNA transcripts to operate the cell, and the micronucleus, which is a storage of the genes and which is involved only in the sexual process undergone by ciliates. One of the most complex DNA processing in nature known to us happens in ciliates while their micronuclear genome is rearranged to the macronuclear genome. This process of gene assembly is intriguing and captivating also from the computational point of view. We investigate here three molecular operations, *ld*, *hi*, and *dlad*, that have been postulated for this assembly process in the intramolecular fashion. These formal models of operations have three abstraction levels: MDS descriptors, legal strings and overlap graphs, and they all turn out to be equivalent as far as the operational model of gene assembly is concerned.

Recent results in natural computing

Thomas J. Head* (SUNY at Binghamton)

Mario J. Pérez Jiménez (Universidad de Sevilla)

Carlos Martín-Vide ((Universitat Rovira i Virgili)

Agent based computational models of cells and tissue

Mike Holcombe (University of Sheffield)

Computational models of biological systems are providing an important mechanism for trying to understand the way in which a complex biological system or subsystem behaves. The vast amount of data from genome and protonome projects needs to be exploited in ways that provide greater insight into the behaviour of systems at the cellular, tissue and organ levels, which are essentially self organising computational systems.

This paper reports work on using a mathematically based agent approach to modelling and simulating systems at different levels. The ground work is also being laid for the automated analysis of these models using techniques such as model checking.

Two case studies will be considered: the Tissure Physiome, where epithelium tissue (skin and urothelium) are being modelled and a model of Magnaporthe Griseum, rice blast fungus, where the whole organism is modelled and its life cycle related to the latest genetic data on this important crop disease.

Another class of semi-simple splicing languages

Rodica Ceterchi (University of Bucharest)

Carlos Martín Vide* (Universitat Rovira i Virgili)

K. G. Subramanian (Madras Christian College)

Following ideas from Goode, Pixton (2001) and Mateescu, Paun, Rozenberg, Salomaa (1998), we introduce in this paper a new class of splicing languages, which we call $(2, 4)$ semi-simple splicing languages. The semi-simple splicing languages of Goode, Pixton (2001) become the $(1, 3)$ class, and the two classes are incomparable. The $(2, 4)$ class is also composed of strictly locally testable languages.

Membrane computing-power and efficiency: Recent results

Gheorghe Paun (Universitat Rovira i Virgili)

Membrane computing is a rather young branch of natural computing which abstracts computing models from the structure and the functioning of the living cell. In short, in the compartments of membrane structure (a cell-like hierarchical arrangement of membranes) multisets of symbol-objects or sets of string-objects evolve according to sets of rules associated with the compartments. The model is distributed, parallel, and with a high degree of nondeterminism. Many types of membrane systems (also called P systems) were introduced in the literature, most of them being shown to be computationally universal. For systems able to create an exponential workspace in polynomialtime, e.g., by cell division, membrane creation, string replication, polynomial solutions to NP-complete problems can be devised (based on a time-space trade-off).

This domain of research is fastly developing (see `http://psystems.disco.unimib.it`) and a monograph (*Membrane Computing. An Introduction*) has been published in August 2002 by Springer-Verlag. The talk will present both the basic definitions of this domain and the main (classes of) results, focusing on notions and results which are not covered by the monograph.

Complexity classes in cellular computing with membranes

Mario J. Pérez Jiménez* (Universidad de Sevilla)

Álvaro Romero-Jiménez (Universidad de Sevilla)

Fernando Sancho-Caparrini (Universidad de Sevilla)

Given a class of recognizer membrane systems $\mathcal{F}$, the complexity class $\mathbf{PMC}_{\mathcal{F}}$ of all problems solvable in polynomial time by a family of P systems of type $\mathcal{F}$ is presented. This complexity class is stable by polynomial time reduction, and offer a new way to attack the $\mathbf{P} \neq \mathbf{NP}$ conjecture, now inside the framework of the cellular computing with membranes.

Using automated reasoning systems in natural computing

C. Graciani Díaz (Universidad de Sevilla)

Mario J. Pérez Jiménez* (Universidad de Sevilla)

The simple intuitive verification or even the use of formal notation does not ensure us that a designed program or a set of specifications are correct. The availability of general purpose theorem proving tools as PVS opens the way for the elaboration and development of techniques that will allow us to attack the processes of verification in the scope of Natural Computing. Also, those formalizations are done in a generic framework in which the concrete implementation of each operation is not important, but only their properties. This is a suitable way of working before implementing them in a real model and brings us the possibility of simulating real experiments or developing new ones.

On some abstract metrics for arbitrary contact structures

Mercè Llabrés Segura (Universitat de les Illes Balears)

Francesc A. Rosselló* (Universitat de les Illes Balears)

RNA molecules and proteins fold into tridimensional structures that play a key role in many biological processes. One of the most important problems in molecular biology is the comparison of these structures. In this work we consider their *contact structure*: their set of bonds between monomers. Two restrictions must be added to the definition of an *RNA contact structure*: two consecutive bases can never bond, and if two bases bond, then no one of them can bond with any other base.

Many notions of similarity between RNA contact structures have been considered in the literature. In a seminal paper on algebraic models of RNA contact structures, Reidys and Stadler introduced in 1996 several abstract metrics based on their algebraic models and independent of any notion of "graph editing"; they also asked for a further analysis of these metrics and for their generalization to the arbitrary contact structure setting.

In this paper we provide a solution to both problems. We compute explicitly the metrics they define and, next, we introduce new abstract metrics based on representations of graphs as polynomial ideals, that can be seen as generalizing those given by Reidys and Stadler and can be used on arbitrary contact structure, and we discuss their mathematical properties and their biophysical meaning.

Circular DNA-based algorithms to solve hard problems

Rani Siromoney (Madras Christian College)

NP-hard problems and NP-complete problems abound in Mathematics and Computer Science. Adleman used single-stranded DNA molecules to solve an instance of the HPP Problem. Head uses artificial circular plasmids to solve the satisfiability and other hard problems. We add to this collection.