Computation in Gene Expression

Hillol Kargupta[*]

Department of Computer Science and Electrical Engineering,
University of Maryland Baltimore County,
Baltimore, MD 21250, USA

Dirk Thierens[†]

Institute of Information and Computing Sciences,
Universiteit Utrecht,
Centrumgebouw Noord, office A356,
Padualaan 14, De Uithof, 3584CH Utrecht

The transformation of the information coded in the DNA to the mRNA, protein, and finally to the phenotype of a living organism is often called the *gene expression* process. It is a complex biochemical process. However, it can also be viewed as a computational process that involves a sequence of representation transformations. Representation transformations are often used in many fields for solving problems efficiently. Therefore, these transformations in gene expression allude to intriguing computational roles in genetic adaptation, search, and evolvability.

This special issue on the *Computation in Gene Expression* is an early effort to consolidate our understanding. It came out of the Genetic and Evolutionary Computation Conference (GECCO), 2000 workshop on this topic. This issue contains three papers. The first paper by Kargupta presents an analysis of genetic code-like representation transformations (GCTs) and their possible role in scalable genetic learning and search. It identifies some GCTs that convert the exponentially long Fourier representation of some nontrivial functions to an exponentially long representation with only a polynomial number of significant terms. This suggests the possibility of constructing efficient and approximate representations through GCTs in genetic learning and search. The second paper by Kennedy and Osborn offers a perspective of gene expression as a mechanism to parse a genomic language for constructing the phenotype. It develops a model of artificial cellular metabolism—*spider*—that expresses the operons encoded in the genome. The final paper by Silvescu and Honavar offers a perspective from a different angle. It takes a data driven approach to understand gene expression. It reports development of a boolean network-based model of a genetic regulatory

---

[*]Electronic mail address: hillol@cs.umbc.edu.
[†]Electronic mail address: dirk@cs.uu.nl.

network. It is interesting to note that in their boolean network the expression of any gene can depend on that of at most k genes. On the other hand, the paper by Kargupta shows that at least for some functions and GCTs, only the interactions between k genes for small values of k are significant. These three papers represent the growing body of literature on the computation in gene expression based on simulation, analytical, and data driven techniques.

We hope that this effort will contribute toward further understanding of the computation in gene expression.