

Preface

Knowledge gained from the sequencing of the human genome promises to change our lives. Powerful computing techniques have been used to acquire the knowledge gained so far, and still more powerful techniques will be required to fulfill the promises of genetically based drug design, medical diagnosis and treatment, and agricultural applications, among others. This issue of the *IBM Systems Journal* contains 19 papers devoted to deep computing for the life sciences.

We are indebted to several individuals for the acquisition, coordination, and organization of papers in this issue, as well as a companion double issue of the *IBM Journal of Research and Development*. S. L. Nunes of IBM Life Sciences Solutions provided early direction. Papers were solicited by J. M. Coffin, IBM Life Sciences Solutions, B. Robson, IBM Research Division, and W. C. Swope, IBM Research Division. We are especially grateful to Swope for his vision, guidance, and ceaseless efforts from inception to publication of these two issues.

The introductory paper, by Swope, is in two parts. The first part gives an overview of the biological processes and concepts that are central to the life sciences and explains how deep computing plays a key role. The second part introduces the papers in this issue of the *IBM Systems Journal* and also those in the companion issue of the *IBM Journal of Research and Development*, for readers who may be interested in further discussion. The papers here are primarily on the computational aspects; those in the companion issue are primarily on the scientific aspects of our joint topic. (See the introductory paper by Swope, Reference 4, for more information on the availability of this companion issue.)

“Deep computing” implies the use of powerful machines running sophisticated software using innova-

tive algorithms to solve very complex problems. The 18 papers that follow the introductory paper in this issue are organized by the type of computation required, with six papers in each category.

The first category, numerically intensive computing, contains papers on topics that range from the computational challenges of structural and functional genomics to the mapping, modeling, and visual exploration of structure-function relationships in the heart.

The second category is data mining and analysis. Here the computation involved includes pattern matching, decision support, and analysis distributed over the Internet. Papers describe methods for protein sequence classification and comparison, protein crystal growth, protein structure prediction, and DNA sequence analysis.

The third category is data management and integration. There are enormous quantities and many types of data involved in the study of the life sciences. Products and methods described in these papers include DiscoveryLink from IBM, K2/Kleisli and GUS from the University of Pennsylvania, GeneX from the National Center for Genome Research, and GeneMine from the University of California, Los Angeles.

The next issue of the *Journal* will focus on end-to-end computer security, from the use of “smart cards” and biometrics for user authentication, through security considerations for architecture, infrastructure, and an operating system, to secure coprocessors. An individual paper will be included on comparing production database workloads with performance benchmarks.

Marilyn L. Bates
Associate Editor

John J. Ritsko
Editor-in-Chief