

1. Record Nr.	UNINA9910299257803321
Autore	Erciyes Kayhan
Titolo	Distributed and Sequential Algorithms for Bioinformatics / / by Kayhan Erciyes
Pubbl/distr/stampa	Cham : , : Springer International Publishing : , : Imprint : Springer, , 2015
ISBN	3-319-24966-5
Edizione	[1st ed. 2015.]
Descrizione fisica	1 online resource (XVII, 367 p. 157 illus. in color.)
Collana	Computational Biology, , 1568-2684 ; ; 23
Disciplina	570.285
Soggetti	Bioinformatics Algorithms Systems biology Biological systems Computer science—Mathematics Computational Biology/Bioinformatics Algorithm Analysis and Problem Complexity Systems Biology Math Applications in Computer Science
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Bibliographic Level Mode of Issuance: Monograph
Nota di contenuto	Intro -- Preface -- Contents -- 1 Introduction -- 1.1 Introduction -- 1.2 Biological Sequences -- 1.3 Biological Networks -- 1.4 The Need for Distributed Algorithms -- 1.5 Outline of the Book -- Part I Background -- 2 Introduction to Molecular Biology -- 2.1 Introduction -- 2.2 The Cell -- 2.2.1 DNA -- 2.2.2 RNA -- 2.2.3 Genes -- 2.2.4 Proteins -- 2.3 Central Dogma of Life -- 2.3.1 Transcription -- 2.3.2 The Genetic Code -- 2.3.3 Translation -- 2.3.4 Mutations -- 2.4 Biotechnological Methods -- 2.4.1 Cloning -- 2.4.2 Polymerase Chain Reaction -- 2.4.3 DNA Sequencing -- 2.5 Databases -- 2.5.1 Nucleotide Databases -- 2.5.2 Protein Sequence Databases -- 2.6 Human Genome Project -- 2.7 Chapter Notes -- 3 Graphs, Algorithms, and Complexity -- 3.1 Introduction -- 3.2 Graphs -- 3.2.1 Types of Graphs -- 3.2.2 Graph Representations -- 3.2.3 Paths, Cycles, and Connectivity -- 3.2.4 Trees -- 3.2.5 Spectral Properties of Graphs --

3.3 Algorithms -- 3.3.1 Time and Space Complexities -- 3.3.2 Recurrences -- 3.3.3 Fundamental Approaches -- 3.3.4 Dynamic Programming -- 3.3.5 Graph Algorithms -- 3.3.6 Special Subgraphs -- 3.4 NP-Completeness -- 3.4.1 Reductions -- 3.4.2 Coping with NP-Completeness -- 3.5 Chapter Notes -- 4 Parallel and Distributed Computing -- 4.1 Introduction -- 4.2 Architectures for Parallel and Distributed Computing -- 4.2.1 Interconnection Networks -- 4.2.2 Multiprocessors and Multicomputers -- 4.2.3 Flynn's Taxonomy -- 4.3 Parallel Computing -- 4.3.1 Complexity of Parallel Algorithms -- 4.3.2 Parallel Random Access Memory Model -- 4.3.3 Parallel Algorithm Design Methods -- 4.3.4 Shared Memory Programming -- 4.3.5 Multi-threaded Programming -- 4.3.6 Parallel Processing in UNIX -- 4.4 Distributed Computing -- 4.4.1 Distributed Algorithm Design -- 4.4.2 Threads Re-visited -- 4.4.3 Message Passing Interface. 4.4.4 Distributed Processing in UNIX -- 4.5 Chapter Notes -- Part II Biological Sequences -- 5 String Algorithms -- 5.1 Introduction -- 5.2 Exact String Matching -- 5.2.1 Sequential Algorithms -- 5.2.2 Distributed String Matching -- 5.3 Approximate String Matching -- 5.4 Longest Subsequence Problems -- 5.4.1 Longest Common Subsequence -- 5.4.2 Longest Increasing Subsequence -- 5.5 Suffix Trees -- 5.5.1 Construction of Suffix Trees -- 5.5.2 Applications of Suffix Trees -- 5.5.3 Suffix Arrays -- 5.6 Chapter Notes -- 6 Sequence Alignment -- 6.1 Introduction -- 6.2 Problem Statement -- 6.2.1 The Objective Function -- 6.2.2 Scoring Matrices for Proteins -- 6.3 Pairwise Alignment -- 6.3.1 Global Alignment -- 6.3.2 Local Alignment -- 6.4 Multiple Sequence Alignment -- 6.4.1 Center Star Method -- 6.4.2 Progressive Alignment -- 6.5 Alignment with Suffix Trees -- 6.6 Database Search -- 6.6.1 FASTA -- 6.6.2 BLAST -- 6.7 Parallel and Distributed Sequence Alignment -- 6.7.1 Parallel and Distributed SW Algorithm -- 6.7.2 Distributed BLAST -- 6.7.3 Parallel/Distributed CLUSTALW -- 6.8 Chapter Notes -- 7 Clustering of Biological Sequences -- 7.1 Introduction -- 7.2 Analysis -- 7.2.1 Distance and Similarity Measures -- 7.2.2 Validation of Cluster Quality -- 7.3 Classical Methods -- 7.3.1 Hierarchical Algorithms -- 7.3.2 Partitional Algorithms -- 7.3.3 Other Methods -- 7.4 Clustering Algorithms Targeting Biological Sequences -- 7.4.1 Alignment-Based Clustering -- 7.4.2 Other Similarity-Based Methods -- 7.4.3 Graph-Based Clustering -- 7.5 Distributed Clustering -- 7.5.1 Hierarchical Clustering -- 7.5.2 k-means Clustering -- 7.5.3 Graph-Based Clustering -- 7.5.4 Review of Existing Algorithms -- 7.6 Chapter Notes -- 8 Sequence Repeats -- 8.1 Introduction -- 8.2 Tandem Repeats -- 8.2.1 Stoye and Gusfield Algorithm -- 8.2.2 Distributed Tandem Repeat Search. 8.3 Sequence Motifs -- 8.3.1 Probabilistic Approaches -- 8.3.2 Combinatorial Methods -- 8.3.3 Parallel and Distributed Motif Search -- 8.3.4 A Survey of Recent Distributed Algorithms -- 8.4 Chapter Notes -- 9 Genome Analysis -- 9.1 Introduction -- 9.2 Gene Finding -- 9.2.1 Fundamental Methods -- 9.2.2 Hidden Markov Models -- 9.2.3 Nature Inspired Methods -- 9.2.4 Distributed Gene Finding -- 9.3 Genome Rearrangement -- 9.3.1 Sorting by Reversals -- 9.3.2 Unsigned Reversals -- 9.3.3 Signed Reversals -- 9.3.4 Distributed Genome Rearrangement Algorithms -- 9.4 Haplotype Inference -- 9.4.1 Problem Statement -- 9.4.2 Clark's Algorithm -- 9.4.3 EM Algorithm -- 9.4.4 Distributed Haplotype Inference Algorithms -- 9.5 Chapter Notes -- Part III Biological Networks -- 10 Analysis of Biological Networks -- 10.1 Introduction -- 10.2 Networks in the Cell -- 10.2.1 Metabolic Networks -- 10.2.2 Gene Regulation Networks -- 10.2.3 Protein Interaction Networks -- 10.3 Networks Outside the Cell -- 10.3.1 Networks of the Brain -- 10.3.2 Phylogenetic Networks --

10.3.3 The Food Web -- 10.4 Properties of Biological Networks --
 10.4.1 Distance -- 10.4.2 Vertex Degrees -- 10.4.3 Clustering
 Coefficient -- 10.4.4 Matching Index -- 10.5 Centrality -- 10.5.1
 Degree Centrality -- 10.5.2 Closeness Centrality -- 10.5.3
 Betweenness Centrality -- 10.5.4 Eigenvalue Centrality -- 10.6
 Network Models -- 10.6.1 Random Networks -- 10.6.2 Small World
 Networks -- 10.6.3 Scale-Free Networks -- 10.6.4 Hierarchical
 Networks -- 10.7 Module Detection -- 10.8 Network Motifs -- 10.9
 Network Alignment -- 10.10 Chapter Notes -- 11 Cluster Discovery in
 Biological Networks -- 11.1 Introduction -- 11.2 Analysis -- 11.2.1
 Quality Metrics -- 11.2.2 Classification of Clustering Algorithms --
 11.3 Hierarchical Clustering -- 11.3.1 MST-Based Clustering -- 11.3.2
 Edge-Betweenness-Based Clustering.
 11.4 Density-Based Clustering -- 11.4.1 Clique Algorithms -- 11.4.2
 k-core Decomposition -- 11.4.3 Highly Connected Subgraphs
 Algorithm -- 11.4.4 Modularity-Based Clustering -- 11.5 Flow
 Simulation-Based Approaches -- 11.5.1 Markov Clustering Algorithm
 -- 11.5.2 Distributed Markov Clustering Algorithm Proposal -- 11.6
 Spectral Clustering -- 11.7 Chapter Notes -- 12 Network Motif Search
 -- 12.1 Introduction -- 12.2 Problem Statement -- 12.2.1 Methods of
 Motif Discovery -- 12.2.2 Relation to Graph Isomorphism -- 12.2.3
 Frequency Concepts -- 12.2.4 Random Graph Generation -- 12.2.5
 Statistical Significance -- 12.3 A Review of Sequential Motif Searching
 Algorithms -- 12.3.1 Network Centric Algorithms -- 12.3.2 Motif
 Centric Algorithms -- 12.4 Distributed Motif Discovery -- 12.4.1 A
 General Framework -- 12.4.2 Review of Distributed Motif Searching
 Algorithms -- 12.4.3 Wang et al.'s Algorithm -- 12.4.4 Schatz et al.'s
 Algorithm -- 12.4.5 Riberio et al.'s Algorithms -- 12.5 Chapter Notes
 -- 13 Network Alignment -- 13.1 Introduction -- 13.2 Problem
 Statement -- 13.2.1 Relation to Graph Isomorphism -- 13.2.2 Relation
 to Bipartite Graph Matching -- 13.2.3 Evaluation of Alignment Quality
 -- 13.2.4 Network Alignment Methods -- 13.3 Review of Sequential
 Network Alignment Algorithms -- 13.3.1 PathBlast -- 13.3.2 IsoRank
 -- 13.3.3 MaWish -- 13.3.4 GRAAL -- 13.3.5 Recent Algorithms --
 13.4 Distributed Network Alignment -- 13.4.1 A Distributed Greedy
 Approximation Algorithm Proposal -- 13.4.2 Distributed Hoepman's
 Algorithm -- 13.4.3 Distributed Auction Algorithms -- 13.5 Chapter
 Notes -- 14 Phylogenetics -- 14.1 Introduction -- 14.2 Terminology --
 14.3 Phylogenetic Trees -- 14.3.1 Distance-Based Algorithms --
 14.3.2 Maximum Parsimony -- 14.3.3 Maximum Likelihood -- 14.4
 Phylogenetic Networks -- 14.4.1 Reconstruction Methods -- 14.5
 Chapter Notes -- 15 Epilogue.
 15.1 Introduction -- 15.2 Current Challenges -- 15.2.1 Big Data
 Analysis -- 15.2.2 Disease Analysis -- 15.2.3 Bioinformatics Education
 -- 15.3 Specific Challenges -- 15.3.1 Sequence Analysis -- 15.3.2
 Network Analysis -- 15.4 Future Directions -- 15.4.1 Big Data Gets
 Bigger -- 15.4.2 New Paradigms on Disease Analysis -- 15.4.3
 Personalized Medicine -- Index.

Sommario/riassunto

This unique textbook/reference presents unified coverage of
 bioinformatics topics relating to both biological sequences and
 biological networks, providing an in-depth analysis of cutting-edge
 distributed algorithms, as well as of relevant sequential algorithms. In
 addition to introducing the latest algorithms in this area, more than
 fifteen new distributed algorithms are also proposed. Topics and
 features: Reviews a range of open challenges in biological sequences
 and networks, beginning with an informal description of the problem
 before defining it formally Describes in detail both sequential and
 parallel/distributed algorithms for each problem, briefly discussing

software packages if there are any available Suggests approaches for distributed algorithms as possible extensions to sequential algorithms, when the distributed algorithms for the topic are scarce Proposes a number of new distributed algorithms in each chapter, to serve as potential starting points for further research Concludes each chapter with self-test exercises, a summary of the key points, a comparison of the algorithms described, and a literature review This clearly-written and easy to follow work is ideal as a textbook for graduate and senior undergraduate students of computer science and biology, and as a self-study guide for any interested reader with a basic background in discrete mathematics and algorithms. Researchers in bioinformatics will also find the book to be a useful reference on this subject. Dr. K. Erciyes is Rector of Izmir University, Turkey, where he also serves as a professor in the Computer Engineering Department. His other publications include the Springer title Distributed Graph Algorithms for Computer Networks.
