

# ALGEBRAIC AND DISCRETE MATHEMATICAL METHODS FOR MODERN BIOLOGY

**RAINA ROBEVA, EDITOR**

Inspired by the national initiative toward a new biology, this work offers a collection of modules introducing methods from modern discrete mathematics into the undergraduate math and biology curricula. Each module begins with a question from contemporary biology, followed by the description of mathematical methods and theory appropriate for the search of answers. Projects and exercises embedded in the text utilize freely accessible or widely available software for visualization, simulation, and analysis used in modern biology research. This book is appropriate for mathematics courses such as finite mathematics, discrete structures, linear algebra, abstract/modern algebra, graph theory, probability, bioinformatics, statistics, biostatistics, and modeling, as well as for biology courses such as genetics, cell and molecular biology, biochemistry, ecology, and evolution. The companion website includes solutions to all exercises and additional materials including tutorials, projects, data sets, and computer code.

## Key Features

- Examines significant questions in modern biology and their mathematical treatments
- Presents important concepts and methods from discrete mathematics in the context of essential biology
- Features material appropriate for both mathematics and biology courses
- Presents chapters in modular format, so coverage does not need to follow the Table of Contents
- Introduces projects appropriate for undergraduate research
- Requires no calculus as a prerequisite



**Raina Robeva** was born in Sofia, Bulgaria. She has a PhD in Mathematics from the University of Virginia and has led multiple NSF-funded curriculum development projects at the interface of mathematics and biology. She is the lead author of the textbook *An Invitation to Biomathematics* (2008) and the lead editor of the volume *Mathematical Concepts and Methods in Modern Biology: Using Modern Discrete Models* (2013), both published by Academic Press. Robeva is the founding Chief Editor of the research journal *Frontiers in Systems Biology*. She is a professor of Mathematical Sciences at Sweet Briar College and lives in Charlottesville, Virginia.



**ACADEMIC PRESS**

An imprint of Elsevier  
store.elsevier.com

ISBN 978-0-12-801213-0



9 780128 012130



# Contents

Contributors	ix	2.6 Conclusions	48
Preface	xi	References	49
<b>1. Graph Theory for Systems Biology: Interval Graphs, Motifs, and Pattern Recognition</b>		<b>3. Adaptation and Fitness Graphs</b>	
<i>John R. Jungck and Rama Viswanathan</i>		<i>Kristina Crona and Emilie Wiesner</i>	
1.1 Introduction	1	3.1 Introduction	51
1.2 Revisualizing, Recognizing, and Reasoning About Relationships	3	3.2 Fitness Landscapes and Fitness Graphs	52
1.2.1 Basic Concepts from Graph Theory	3	3.2.1 Basic Terminology and Notation	52
1.2.2 Interval Graphs in Biology	6	3.2.2 Fitness, Fitness Landscapes, and Fitness Graphs	53
1.3 Example I—Differentiation: Gene Expression	15	3.2.3 Epistasis	55
1.4 Example II—Disease Etiology	20	3.3 Fitness Graphs and Recombination	58
1.5 Conclusion	25	3.4 Fitness Graphs and Drug Cycling	60
Acknowledgments	26	References	63
References	26	<b>4. Signaling Networks: Asynchronous Boolean Models</b>	
<b>2. Food Webs and Graphs</b>		<i>Réka Albert and Raina Robeva</i>	
<i>Margaret (Midge) Cozzens</i>		4.1 Introduction to Signaling Networks	65
2.1 Introduction	29	4.2 A Brief Summary of Graph-Theoretic Analysis of Signaling Networks	66
2.2 Modeling Predator-Prey Relationships with Food Webs	29	4.3 Dynamic Modeling of Signaling Networks	69
2.3 Trophic Levels and Trophic Status	30	4.4 The Representation of Node Regulation in Boolean Models	70
2.3.1 Background and Definitions	31	4.5 The Dynamics of Boolean Models	72
2.3.2 Adding Complexity: Weighted Food Webs and Flow-Based Trophic Levels	35	4.6 Attractor Analysis for Stochastic Asynchronous Update	75
2.3.3 Flow-Based Trophic Level	36	4.7 Boolean Models Capture Characteristic Dynamic Behavior	77
2.4 Competition Graphs and Habitat Dimension	37	4.8 How to Deal with Incomplete Information when Constructing the Model	80
2.4.1 Competition Graphs (also Called Niche Overlap Graphs and Predator Graphs)	37	4.8.1 Dealing with Gaps in Network Construction	81
2.4.2 Interval Graphs and Boxicity	37	4.8.2 Dealing with Gaps in Transition Functions	82
2.4.3 Habitat Dimension	40	4.8.3 Dealing with Gaps in Initial Condition	84
2.5 Connectance, Competition Number, and Projection Graphs	41	4.8.4 Dealing with Gaps in Timing Information	85
2.5.1 Connectance	42	4.9 Generate Novel Predictions with the Model	85
2.5.2 Competition Number	43		
2.5.3 Projection Graphs	44		



4.10 Boolean Rule-Based Structural Analysis of Cellular Networks	86	6.8 Conclusion	137
4.11 Conclusions	90	References	138
References	90		
<b>5. Dynamics of Complex Boolean Networks: Canalization, Stability, and Criticality</b>		<b>7. BioModel Engineering with Petri Nets</b>	
<i>Qijun He, Matthew Macauley and Robin Davies</i>		<i>Mary Ann Blätke, Monika Heiner and Wolfgang Marwan</i>	
5.1 Introduction	93	7.1 Introduction	141
5.2 Boolean Network Models	95	7.2 Running Case Study	144
5.2.1 Gene Regulatory Networks	95	7.3 Petri Nets ( $\mathcal{PN}$ )	146
5.2.2 Network Topology	96	7.3.1 Modeling	146
5.2.3 Network Topology and Random Networks	99	7.3.2 Analysis	153
5.2.4 Boolean Functions	100	7.3.3 Further Reading	159
5.2.5 Boolean Networks	102	7.3.4 Exercises	160
5.3 Canalization	104	7.4 Stochastic Petri Nets ( $\mathcal{SPN}$ )	162
5.3.1 Canalizing Boolean Functions	104	7.4.1 Modeling	162
5.3.2 Nested Canalizing Functions	105	7.4.2 Analysis	165
5.3.3 Canalizing Depth	109	7.4.3 Further Reading	169
5.3.4 Dominant Variables of NCFs	110	7.4.4 Exercises	170
5.4 Dynamics Over Complex Networks	112	7.5 Continuous Petri Nets ( $\mathcal{CPN}$ )	172
5.4.1 Boolean Calculus	113	7.5.1 Modeling	172
5.4.2 Derrida Plots and the Three Dynamical Regimes	115	7.5.2 Analysis	173
5.4.3 Ensembles of RBNs	116	7.5.3 Further Reading	175
Acknowledgments	118	7.5.4 Exercises	176
References	118	7.6 Hybrid Petri Nets ( $\mathcal{HPN}$ )	177
		7.6.1 Modeling	178
<b>6. Steady State Analysis of Boolean Models: A Dimension Reduction Approach</b>		7.6.2 Analysis	180
<i>Alan Veliz-Cuba and David Murrugarra</i>		7.6.3 Further Reading	181
6.1 Introduction	121	7.6.4 Exercises	182
6.2 An Example: Toy Model of the <i>lac</i> Operon	122	7.7 Colored Petri Nets	183
6.3 General Reduction	125	7.7.1 Further Reading	186
6.3.1 Definition	125	7.7.2 Exercises	186
6.3.2 Examples	125	7.8 Conclusions	187
6.4 Implementing the Reduction Algorithm Using Boolean Algebra	128	Acknowledgments	189
6.5 Implementing the Reduction Algorithm Using Polynomial Algebra	129	7.9 Supplementary Materials	189
6.5.1 Background	129	References	189
6.5.2 Using Polynomial Algebra Software to Reduce Boolean Networks	130		
6.6 Applications	131	<b>8. Transmission of Infectious Diseases: Data, Models, and Simulations</b>	
6.6.1 The <i>lac</i> Operon	131	<i>Winfried Just, Hannah Callender, M. Drew LaMar and Natalia Toporikova</i>	
6.6.2 Th-Cell Differentiation	133	8.1 Introduction: Why Do We Want to Model Infectious Diseases?	193
6.7 AND Boolean Models	134	8.2 Mathematical Models of Disease Transmission	198
6.7.1 Background	135	8.2.1 Transmission Probabilities	199
		8.2.2 The Time Line of Within-Host Dynamics	201
		8.2.3 Movement Between Compartments	203
		8.2.4 Basic Model Types: <i>SEIR</i> , <i>SIR</i> , <i>SI</i> , and <i>SIS</i>	206



8.2.5	How to Model Time and Run Simulations	208	10.3.1	Nonindependence of Multiple Traits	250
8.3	How Does the Computer Run Simulations?	210	10.3.2	The Genetic Variance-Covariance Matrix	252
8.3.1	Meet the Simulator	210	10.3.3	Simultaneous Selection on Multiple Traits	253
8.3.2	How to Load the Die	212	10.3.4	Predicting the Outcome of Selection on Covarying Traits	255
References		214	10.3.5	Evolution of the <b>G</b> Matrix Itself	257
9.	Disease Transmission Dynamics on Networks: Network Structure Versus Disease Dynamics		References		258
	<i>Winfried Just, Hannah Callender and M. Drew Lamar</i>		11.	Metabolic Analysis: Algebraic and Geometric Methods	
9.1	Introduction	217		<i>Terrell L. Hodge, Blair R. Szymczyna and Todd J. Barkman</i>	
9.2	Models Based on the Uniform Mixing Assumption	218	11.1	Introduction	261
9.2.1	Compartment-Based Models	218	11.2	Encoding the Reactions: Linear Algebraic Modeling	262
9.2.2	The Basic Reproductive Ratio $R_0$	220	11.3	Adding Reaction Kinetics: Algebraic Formulation of Mass-Action Kinetics	271
9.3	Network-Based Models	224	11.4	Directions for Further Reading and Research: Metabolic Pathways	273
9.3.1	Networks and Graphs	225	11.5	NMR and Linear Algebraic Methods	274
9.3.2	Disease Transmission on Networks	229	11.6	NMR Spectroscopy and Applications to the Study of Metabolism	274
9.3.3	Examples of Contact Networks	230	11.6.1	Principles of NMR Spectroscopy	275
9.3.4	Additional Graph-Theoretic Notions	231	11.6.2	The NMR Spectrum	277
9.3.5	Erdős-Rényi Random Graphs	233	11.6.3	NMR Investigations of Metabolism	281
9.4	Suggestions for Further Study	234	11.7	NMR for Metabolic Analysis and Mathematical Methods: Directions of Further Research	289
Acknowledgments		235	11.8	Supplementary Materials	290
References		235	References		290
10.	Predicting Correlated Responses in Quantitative Traits Under Selection: A Linear Algebra Approach		12.	Reconstructing the Phylogeny: Computational Methods	
	<i>Janet Steven and Bessie Kirkwood</i>			<i>Grady Weyenberg and Ruriko Yoshida</i>	
10.1	Introduction	237	12.1	Introduction	293
10.2	Quantifying Selection on Quantitative Traits	238	12.1.1	Sequences and Alignments	297
10.2.1	Describing Traits Mathematically	238	12.2	Quantifying Evolutionary Change	299
10.2.2	Quantifying Reproduction and Survival	241	12.2.1	Probabilistic Models of Molecular Evolution	299
10.2.3	Describing the Relationship Between Fitness and a Trait	242	12.2.2	Common Model Extensions	306
10.2.4	Determining the Genetic Component of Quantitative Traits	245	12.3	Reconstructing the Tree	306
10.2.5	Estimating Heritability in a Trait	246	12.3.1	Distance-Based Methods	306
10.2.6	The Breeder's Equation	247	12.3.2	Maximum Parsimony	309
10.2.7	The Price Equation	249	12.3.3	Methods Based on Probability Models	310
10.3	Covariance Among Traits Under Selection	249			



12.4 Model Selection	312
12.5 Statistical Methods to Test Congruency Between Trees	313
References	316

### 13. RNA Secondary Structures: Combinatorial Models and Folding Algorithms

*Qijun He, Matthew Macauley and  
Robin Davies*

13.1 Introduction	321
13.2 Combinatorial Models of Noncrossing RNA Structures	324
13.2.1 Partial Matchings and Physical Constraints	324
13.2.2 Loop Decomposition	327
13.3 Energy-Based Folding Algorithms for Secondary Structure Prediction	329
13.3.1 Maximizing Bond Strengths via Dynamic Programming	329
13.3.2 Minimum Free Energy Folding	333
13.4 Stochastic Folding Algorithms via Language Theory	335
13.4.1 Languages and Grammars	335

13.4.2 The Knudsen-Hein Grammar for RNA Secondary Structures	337
13.4.3 Secondary Structure Prediction Using SCFGs	340
13.4.4 Summary	341
13.5 Pseudoknots	341
Acknowledgments	344
References	344

### 14. RNA Secondary Structures: An Approach Through Pseudoknots and Fatgraphs

*Christian M. Reidys*

14.1 Introduction	347
14.2 Fatgraphs and Shapes	349
14.3 Genus Recursion	354
14.4 Shapes of Fixed Topological Genus	357
Acknowledgments	361
References	361

Index	363
-------	-----