

Introduction to Modeling for Biosciences

David J. Barnes • Dominique Chu

Introduction to Modeling for Biosciences

 Springer

David J. Barnes
Computing Laboratory
University of Kent
Canterbury, Kent CT2 7NF
UK
d.j.barnes@kent.ac.uk

Dr. Dominique Chu
Computing Laboratory
University of Kent
Canterbury, Kent CT2 7NF
UK
D.F.Chu@kent.ac.uk

ISBN 978-1-84996-325-1
DOI 10.1007/978-1-84996-326-8
Springer London Dordrecht Heidelberg New York

e-ISBN 978-1-84996-326-8

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Control Number: 2010931520

© Springer-Verlag London Limited 2010

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act 1988, this publication may only be reproduced, stored or transmitted, in any form or by any means, with the prior permission in writing of the publishers, or in the case of reprographic reproduction in accordance with the terms of licenses issued by the Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to the publishers.

The use of registered names, trademarks, etc., in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant laws and regulations and therefore free for general use.

The publisher makes no representation, express or implied, with regard to the accuracy of the information contained in this book and cannot accept any legal responsibility or liability for any errors or omissions that may be made.

Cover art by snailsnail
Cover design: KünkelLopka GmbH, Heidelberg

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

尊法世憲

*David dedicates this book to
Helen,
Ben, (Hannah and John), John and Sarah*

Preface

In this book we seek to provide a detailed introduction to a range of modeling techniques that are appropriate for modeling in biosciences. The book is primarily intended for bioscientists, but will be equally useful for anybody wishing to start modeling in biosciences and related fields. The topics we discuss include agent-based models, stochastic modeling techniques, differential equations and Gillespie's stochastic simulation algorithm. Throughout, we pay particular attention to the needs of the novice modeler. We recognise that modeling in science in general (and in biology, in particular) requires both *skills* (i.e., programming, developing algorithms, and solving equations) and *techniques* (i.e., the ability to recognise what is important and needs to be represented in the model, and what can and should be left out). In our experience with novice modelers we have noticed that: (i) both skill and technique are equally important; and (ii) both are normally lacking to some degree.

The philosophy of this book, therefore, is to discuss both aspects—the technical side, and the side that concerns being able to identify the right degree of abstraction. As far as the latter area is concerned, we do not believe that there is a set of rules that, if followed, will necessarily lead to a successful modeling result. Therefore, we have not provided a list of such rules. Instead, we adopt a practical approach which involves walking the reader through realistic and concrete modeling projects. In doing so, we highlight and comment on the process of abstracting the real system into a model. The motivation for this approach is that it is akin to apprenticeship, allowing the reader both to observe practical expertise and to generate personal understanding and intuition, that will ultimately help them to formulate their own models.

Included in the book are practical introductions to a number of useful tools, such as the Maxima computer algebra system, the PRISM model checker, and the Repast Symphony agent modeling environment. Some of the chapters also include exercises to help the reader sharpen their understanding of the topics. The book is supported by a web site, <http://www.cs.kent.ac.uk/imb/>, that includes source code of many of the example models we discuss.

Canterbury, UK

David J. Barnes
Dominique Chu

Contents

- 1 Foundations of Modeling 1**
 - 1.1 Simulation vs. Analytic Results 3
 - 1.2 Stochastic vs. Deterministic Models 5
 - 1.3 Fundamentals of Modeling 6
 - 1.4 Validity and Purpose of Models 11

- 2 Agent-Based Modeling 15**
 - 2.1 Mathematical and Computational Modeling 15
 - 2.1.1 Limits to Modeling 17
 - 2.2 Agent-Based Models 21
 - 2.2.1 The Structure of ABMs 22
 - 2.2.2 Algorithms 25
 - 2.2.3 Time-Driven Algorithms 26
 - 2.2.4 Event-Driven Models 28
 - 2.3 Game of Life 30
 - 2.4 Malaria 34
 - 2.4.1 A Digression 37
 - 2.4.2 Stochastic Systems 39
 - 2.4.3 Immobile Agents 43
 - 2.5 General Consideration when Analyzing a Model 46
 - 2.5.1 How to Test ABMs? 47
 - 2.6 Case Study: The Evolution of Fimbriation 48
 - 2.6.1 Group Selection 49
 - 2.6.2 The Model 51

- 3 ABMs Using Repast and Java 79**
 - 3.1 The Basics of Agent-Based Modeling 80
 - 3.2 An Outline of Repast Concepts 83
 - 3.2.1 Contexts and Projections 84
 - 3.2.2 Model Parameterization 86
 - 3.3 The Game of Life in Repast S 87
 - 3.3.1 The model.score File 88

3.3.2	The Agent Class	89
3.3.3	The Model Initializer	103
3.3.4	Summary of Model Creation	104
3.3.5	Running the Model	105
3.3.6	Creating a Display	106
3.3.7	Creating an Agent Style Class	107
3.3.8	Inspecting Agents at Runtime	109
3.3.9	Review	109
3.4	Malaria Model in Repast Using Java	110
3.4.1	The Malaria Model	110
3.4.2	The model.score File	111
3.4.3	Commonalities in the Agent Types	112
3.4.4	Building the Root Context	112
3.4.5	Accessing Runtime Parameter Values	113
3.4.6	Creating a Projection	114
3.4.7	Implementing the Common Elements of the Agents	115
3.4.8	Completing the Mosquito Agent	118
3.4.9	Scheduling the Actions	119
3.4.10	Visualizing the Model	120
3.4.11	Charts	121
3.4.12	Outputting Data	124
3.4.13	A Statistics-Gathering Agent	124
3.4.14	Summary of Concepts Relating to the Malaria Model	127
3.4.15	Running Repast Models Outside Eclipse	128
3.4.16	Going Further with Repast S	130
4	Differential Equations	131
4.1	Differentiation	131
4.1.1	A Mathematical Example	136
4.1.2	Digression	139
4.2	Integration	141
4.3	Differential Equations	144
4.3.1	Limits to Growth	147
4.3.2	Steady State	150
4.3.3	Bacterial Growth Revisited	152
4.4	Case Study: Malaria	154
4.4.1	A Brief Note on Stability	161
4.5	Chemical Reactions	166
4.5.1	Michaelis-Menten and Hill Kinetics	168
4.5.2	Modeling Gene Expression	173
4.6	Case Study: Cherry and Adler's Bistable Switch	177
4.7	Summary	182
5	Mathematical Tools	183
5.1	A Word of Warning: Pitfalls of CAS	183

- 5.2 Existing Tools and Types of Systems 185
- 5.3 Maxima: Preliminaries 187
- 5.4 Maxima: Simple Sample Sessions 189
 - 5.4.1 The Basics 189
 - 5.4.2 Saving and Recalling Sessions 194
- 5.5 Maxima: Beyond Preliminaries 195
 - 5.5.1 Solving Equations 196
 - 5.5.2 Matrices and Eigenvalues 198
 - 5.5.3 Graphics and Plotting 200
 - 5.5.4 Integrating and Differentiating 205
- 5.6 Maxima: Case Studies 209
 - 5.6.1 Gene Expression 209
 - 5.6.2 Malaria 210
 - 5.6.3 Cherry and Adler’s Bistable Switch 212
- 5.7 Summary 214

- 6 Other Stochastic Methods and Prism 215**
 - 6.1 The Master Equation 217
 - 6.2 Partition Functions 225
 - 6.2.1 Preferences 227
 - 6.2.2 Binding to DNA 231
 - 6.2.3 Codon Bias in Proteins 235
 - 6.3 Markov Chains 236
 - 6.3.1 Absorbing Markov Chains 240
 - 6.3.2 Continuous Time Markov Chains 242
 - 6.3.3 An Example from Gene Activation 244
 - 6.4 Analyzing Markov Chains: Sample Paths 246
 - 6.5 Analyzing Markov Chains: Using PRISM 248
 - 6.5.1 The PRISM Modeling Language 249
 - 6.5.2 Running PRISM 251
 - 6.5.3 Rewards 257
 - 6.5.4 Simulation in PRISM 261
 - 6.5.5 The PRISM GUI 263
 - 6.6 Examples 264
 - 6.6.1 Fim Switching 265
 - 6.6.2 Stochastic Versions of a Differential Equation 268
 - 6.6.3 Tricks for PRISM Models 270

- 7 Simulating Biochemical Systems 273**
 - 7.1 The Gillespie Algorithms 273
 - 7.1.1 Gillespie’s Direct Method 274
 - 7.1.2 Gillespie’s First Reaction Method 275
 - 7.1.3 Java Implementation of the Direct Method 276
 - 7.1.4 A Single Reaction 278
 - 7.1.5 Multiple Reactions 279

7.1.6	The Lotka-Volterra Equation	281
7.2	The Gibson-Bruck Algorithm	284
7.2.1	The Dependency Graph	285
7.2.2	The Indexed Priority Queue	285
7.2.3	Updating the τ Values	286
7.2.4	Analysis	288
7.3	A Constant Time Method	289
7.3.1	Selection Procedure	290
7.3.2	Reaction Selection	292
7.4	Practical Implementation Considerations	293
7.4.1	Data Structures—The Dependency Tree	294
7.4.2	Programming Techniques—Tree Updating	295
7.4.3	Runtime Environment	296
7.5	The Tau-Leap Method	297
7.6	Dizzy	297
7.7	Delayed Stochastic Models	301
7.8	The Stochastic Genetic Networks Simulator	303
7.9	Summary	305
A	Reference Material	307
A.1	Repast Batch Running	307
A.2	Some Common Rules of Differentiation and Integration	307
A.2.1	Common Differentials	307
A.2.2	Common Integrals	308
A.3	Maxima Notation	309
A.4	PRISM Notation Summary	310
A.5	Some Mathematical Concepts	310
A.5.1	Vectors and Matrices	310
A.5.2	Probability	313
A.5.3	Probability Distributions	314
A.5.4	Taylor Expansion	315
	References	317
	Index	319