

# SpringerBriefs in Electrical and Computer Engineering

Control, Automation and Robotics

*Series Editors*

Tamer Başar  
Antonio Bicchi  
Miroslav Krstic

For further volumes:  
<http://www.springer.com/series/10198>

Mathukumalli Vidyasagar

# Computational Cancer Biology

An Interaction Network Approach

Mathukumalli Vidyasagar  
Bioengineering Department  
The University of Texas at Dallas  
Richardson  
USA

ISSN 2192-6786 ISSN 2192-6794 (electronic)  
ISBN 978-1-4471-4750-3 ISBN 978-1-4471-4751-0 (eBook)  
DOI 10.1007/978-1-4471-4751-0  
Springer London Heidelberg New York Dordrecht

Library of Congress Control Number: 2012950853

© The Author(s) 2012

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

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

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media ([www.springer.com](http://www.springer.com))

*To Mike White*

# Preface

During the first four decades after completing my Ph.D., I was a ‘card-carrying control theorist’. Sometime toward the end of this period I became aware that there was a major revolution taking place in biology, whereby it became pretty easy for anyone with a sufficiently large budget to generate massive amounts of raw data. The challenge, I was repeatedly told, was to process these data to identify patterns and draw valid conclusions. Accordingly I tried to read several books and articles on bioinformatics and computational biology, but never felt sufficiently enlightened as to what the underlying issues were, or enthused to pursue the subject. After I moved to the University of Texas at Dallas, I met Professor Michael A. White, Scientific Director of the Harold Simmons Cancer Center at the famed UT Southwestern Medical Center, also in Dallas. I can say this was a transformative event in my life. Mike is a rarity among biologists I have met, someone who is able to convey a broad picture of biological issues, and also to have an open mind toward those from alien cultures such as myself. Through intense interactions with Mike over the past two plus years, I have been able to formulate several statistical and algorithmic problems that are both interesting and challenging to those with a systems and control background, as well as useful to cancer biologists. While I have a long way to go, at least I have commenced on what promises to be a richly rewarding journey.

Another stroke of luck was Edwin Chong, in his capacity as General Chair of the 2011 joint Conference on Decision and Control and the European Control Conference, inviting me to give a plenary lecture. Under normal conditions, the plenary lecturer just shows up and gives his talk, hands over a copy of the slides, and that’s that. But since this was a joint CDC-ECC, I had to prepare a proper journal article. This forced me to channel my meandering thoughts into something more coherent, and resulted in the tutorial paper [1]. So when Springer-Verlag asked me to write a brief monograph, it was natural to expand that paper into the present work, by elaborating on some of the ideas and including subsequent research. Due to the page limitations, even this work describes only a part of the research that my students and I are currently carrying out.

During the two years plus that Mike and I have been talking, I have picked up enough of the rudiments of cancer biology that I can now, with some confidence, act as an ambassador of the cancer biology community to the control theory community. I have not yet had the courage to try it in the other direction. Nonetheless, for his definitive role in converting me into at least a passable imitation of a computational cancer biologist, I take great pleasure in dedicating this book to Mike.

I acknowledge with gratitude financial support from the National Science Foundation Award #1001643, the Cecil & Ida Green Endowment at UT Dallas, and from the Harold Simmons Comprehensive Cancer Center at UT Southwestern Medical Center.

Hyderabad and Dallas, September 2012

Mathukumalli Vidyasagar

## Reference

1. Vidyasagar, M.: Probabilistic methods in cancer biology. *Eur. J. Control* **17**(5–6), 483–511 (2011)

# Contents

<b>1</b>	<b>The Role of System Theory in Biology</b>	<b>1</b>
1.1	Introduction	1
1.2	Some Facts and Figures About Cancer	3
1.3	Advances in Data Generation	4
1.3.1	Genome Sequencing	5
1.3.2	Microarray Experiments	6
1.3.3	Chromatin Immunoprecipitation and ChIP-seq	8
1.3.4	Small Interfering RNAs	8
1.3.5	Micro-RNAs	9
1.4	Role for the Systems and Control Community	9
1.5	Organization of the Monograph	11
	References	12
<b>2</b>	<b>Analyzing Statistical Significance</b>	<b>13</b>
2.1	Basic Statistical Tests	14
2.1.1	The $t$ -Test	14
2.1.2	The Kolmogorov–Smirnov (K–S) Tests	16
2.2	Significance Analysis for Microarrays	22
2.3	Gene Set Enhancement Analysis	24
2.4	Pattern Analysis for Microarrays	26
	References	28
<b>3</b>	<b>Inferring Gene Interaction Networks</b>	<b>31</b>
3.1	Background	31
3.2	Problem Formulations	36
3.3	Methods Based on Mutual Information	39
3.4	Methods Based on Bayesian Networks	43
3.5	A Unified Interpretation	46

3.6	A New Algorithm Based on the Phi-Mixing Coefficient. . . . .	50
3.6.1	The Phi-Mixing Coefficient: Definition and Computation. . . . .	50
3.6.2	Data Processing Inequality for the Phi-Mixing Coefficient . . . . .	55
3.6.3	A New Algorithm for Inferring GINs. . . . .	57
3.6.4	A Case Study: A Lung Cancer GIN . . . . .	61
3.7	Evaluation and Validation of Competing Approaches . . . . .	66
	References . . . . .	67
<b>4</b>	<b>Some Research Directions . . . . .</b>	<b>69</b>
4.1	Harmonizing Prior Knowledge with Phixer Output GIN. . . . .	70
4.2	Identification of Genomic Machines . . . . .	71
4.3	Separating Drivers from Passengers. . . . .	75
	References . . . . .	80



## Editors' Bios

**Tamer Başar** is with the University of Illinois at Urbana-Champaign, where he holds the academic positions of Swanlund Endowed Chair, Center for Advanced Study Professor of Electrical and Computer Engineering, Research Professor at the Coordinated Science Laboratory, and Research Professor at the Information Trust Institute. He received the B.S.E.E. degree from Robert College, Istanbul, and the M.S., M.Phil, and Ph.D. degrees from Yale University. He has published extensively in systems, control, communications, and dynamic games, and has current research interests that address fundamental issues in these areas along with applications such as formation in adversarial environments, network security, resilience in cyber-physical systems, and pricing in networks.

In addition to his editorial involvement with these *Briefs*, Başar is also the Editor-in-Chief of *Automatica*, Editor of two Birkhäuser Series on *Systems & Control* and *Static & Dynamic Game Theory*, the Managing Editor of the *Annals of the International Society of Dynamic Games* (ISDG), and member of editorial and advisory boards of several international journals in control, wireless networks, and applied mathematics. He has received several awards and recognitions over the years, among which are the Medal of Science of Turkey (1993); Bode Lecture Prize (2004) of IEEE CSS; Quazza Medal (2005) of IFAC; Bellman Control Heritage Award (2006) of AACC; and Isaacs Award (2010) of ISDG. He is a member of the US National Academy of Engineering, Fellow of IEEE and IFAC, Council Member of IFAC (2011–2014), a past president of CSS, the founding president of ISDG, and president of AACC (2010–2011).

**Antonio Bicchi** is Professor of Automatic Control and Robotics at the University of Pisa. He graduated at the University of Bologna in 1988, and was a postdoc scholar at M.I.T. A.I. Lab between 1988 and 1990.

His main research interests are in:

- dynamics, kinematics, and control of complex mechanical systems, including robots, autonomous vehicles, and automotive systems;

- haptics and dextrous manipulation; and
- theory and control of nonlinear systems, in particular hybrid (logic/dynamic, symbol/signal) systems.

He has published more than 300 papers on international journals, books, and refereed conferences.

Professor Bicchi currently serves as the Director of the Interdepartmental Research Center “E. Piaggio” of the University of Pisa, and President of the Italian Association of Researchers in Automatic Control. He has served as Editor-in-Chief of the Conference Editorial Board for the IEEE Robotics and Automation Society (RAS), and as Vice President of IEEE RAS, Distinguished Lecturer, and Editor for several scientific journals including the *International Journal of Robotics Research*, the *IEEE Transactions on Robotics and Automation*, and *IEEE RAS Magazine*. He has organized and co-chaired the first World Haptics Conference (2005), and Hybrid Systems: Computation and Control (2007). He is the recipient of several best paper awards at various conferences, and of an Advanced Grant from the European Research Council. Antonio Bicchi has been an IEEE Fellow since 2005.

**Miroslav Krstic** holds the Daniel L. Alspach chair and is the founding director of the Cymer Center for Control Systems and Dynamics at University of California, San Diego. He is a recipient of the PECASE, NSF Career, and ONR Young Investigator Awards, as well as the Axelby and Schuck Paper Prizes. Professor Krstic was the first recipient of the UCSD Research Award in the area of engineering and has held the Russell Severance Springer Distinguished Visiting Professorship at UC Berkeley and the Harold W. Sorenson Distinguished Professorship at UCSD. He is a Fellow of IEEE and IFAC. Professor Krstic serves as Senior Editor for *Automatica* and *IEEE Transactions on Automatic Control* and as Editor for the Springer series *Communications and Control Engineering*. He has served as Vice President for Technical Activities of the IEEE Control Systems Society. Krstic has co-authored eight books on adaptive, nonlinear, and stochastic control, extremum seeking, control of PDE systems including turbulent flows, and control of delay systems.