

Die Reihe

Biometrie und Medizinische Informatik Greifswalder Seminarberichte

wird herausgegeben von:

Universitätsprofessor Dr. rer. nat. habil. Karl-Ernst Biebler

und Dr. rer. nat. Bernd Jäger

Institut für Biometrie und Medizinische Informatik der

Ernst-Moritz-Arndt-Universität Greifswald

Rathenaustraße 48

17487 Greifswald

Anliegen der *Greifswalder Seminarberichte Biometrie und Medizinische Informatik* ist es, die Zusammenarbeit im Berührungsfeld von Medizin, Biowissenschaften, Mathematik und Informatik zu fördern.

In diesem Sinne steht die Mitarbeit an den Publikationen allen denen offen, die sich einer solchen Vermittlung zwischen den Wissenschaften verpflichtet fühlen, und die dazu einen Beitrag leisten wollen.

Über den Autor:

Prof. Dr. rer. nat. habil. Dietmar Cieslik ist Mitarbeiter des Instituts für Mathematik und Informatik der Ernst-Moritz-Arndt-Universität Greifswald.

DISCRETE STRUCTURES IN BIOMATHEMATICS

An introduction for application

Dietmar CIESLIK

Ernst-Moritz-Arndt University

Greifswald, Germany

Biometrie und medizinische Informatik
Greifswalder Seminarberichte

Dietmar Cieslik

Discrete structures in biomathematics

An introduction for application

Shaker Verlag
Aachen 2006

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at <http://dnb.d-nb.de>.

Copyright Shaker Verlag 2006

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publishers.

Printed in Germany.

ISBN-10: 3-8322-5691-1

ISBN-13: 978-3-8322-5691-3

ISSN 1439-5320

Shaker Verlag GmbH • P.O. BOX 101818 • D-52018 Aachen

Phone: 0049/2407/9596-0 • Telefax: 0049/2407/9596-9

Internet: www.shaker.de • e-mail: info@shaker.de

CONTENTS

PREFACE	ix
1 BASICS	1
1.1 Sets	1
1.2 The number of objects in a set	4
1.3 Ordered pairs, relations, correspondences	5
1.4 The multiplication principle	12
1.5 The addition principle	13
1.6 Factorials	15
1.7 Metric spaces	16
1.8 Infinite sets	18
1.9 Orders of growing and asymptotics	27
2 SELECTING OBJECTS	31
2.1 The number of subsets	31
2.2 Binomial coefficients	33
2.3 The Hardy-Weinberg equilibrium	41
2.4 Selections with repetitions	46
3 PARTITIONS	49
3.1 Partition of a set	49
3.2 Partitions of a given size	51
3.3 The Stirling numbers of the first kind	51
3.4 The Stirling numbers of the second kind	53
3.5 The interrelation between the Stirling numbers and the Bell numbers	55

4	RECURRENCE RELATIONS	57
4.1	Linear recurrence relations	58
4.2	The dynamics of population growth	66
4.3	Numerical recurrences	69
5	COMBINATORIAL PROBABILITY	73
5.1	Events and probabilities	73
5.2	The algebra of probabilities	74
5.3	Conditional probability and independent events	76
5.4	The birthday paradox	77
5.5	Random variables	79
5.6	Random Genetic Drift	88
5.7	Watson's paradox	89
5.8	The theory of information	91
5.9	The origin of life and Eigen's paradox	99
5.10	Sequence spaces	101
6	MARKOV PROCESSES	105
6.1	Transitions	105
6.2	Two-states processes	106
6.3	Evolutionary models	108
7	SIMILARITY OF WORDS	111
7.1	Distances between words	112
7.2	Similarity of words	113
7.3	The interrelation between distance and similarity	118
7.4	Specific cases	121
7.5	The algorithm	122
7.6	Optimal multiple alignments	123
8	GRAPHS AND TREES	125
8.1	Graphs	125
8.2	The metric closure of a network	127
8.3	Trees	129
8.4	Minimum spanning trees	130
8.5	Labelled and semi-labelled trees	131

8.6	Unlabelled trees	136
8.7	Binary trees	138
8.8	Rooted trees	139
9	TRAVELLING ROUND A GRAPH	143
9.1	Eulerian cycles	143
9.2	Hamiltonian graphs and k -orders	144
9.3	The shortest superstring problem	147
10	CLASSIFICATIONS	151
10.1	Classification and evolution	153
10.2	Classifications and trees	154
10.3	Pair grouping	158
11	THE PHYLOGENY	161
11.1	Phylogenetic trees	161
11.2	The perfect phylogeny problem	165
11.3	Distance methods	167
11.4	The maximum parsimony method	169
11.5	Consensus trees	171
	REFERENCES	173
	INDEX	179