## Contents

<table>
<thead>
<tr>
<th>Part I: Foundations of Biomedical Informatics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Biomedical Data</td>
</tr>
<tr>
<td>1.1 The Nature and Representation of Biomedical Data</td>
</tr>
<tr>
<td>1.1.1 What Can Be Represented in a Computer?</td>
</tr>
<tr>
<td>1.1.2 DNA and the Genetic Code</td>
</tr>
<tr>
<td>1.1.3 Anatomy</td>
</tr>
<tr>
<td>1.1.4 Medical Laboratory Data</td>
</tr>
<tr>
<td>1.1.5 Medical Images</td>
</tr>
<tr>
<td>1.1.6 Metadata</td>
</tr>
<tr>
<td>1.2 Objects, Metadata, and Serialization</td>
</tr>
<tr>
<td>1.2.1 A Simple Solution Using Tags (Keywords)</td>
</tr>
<tr>
<td>1.2.2 An Object-oriented Design</td>
</tr>
<tr>
<td>1.2.3 A Better Solution Using Meta-objects</td>
</tr>
<tr>
<td>1.2.4 Medical Images: Incorporating Binary Data</td>
</tr>
<tr>
<td>1.3 XML</td>
</tr>
<tr>
<td>1.4 Database Systems and Ideas</td>
</tr>
<tr>
<td>1.4.1 The Relational Model</td>
</tr>
<tr>
<td>1.4.2 The Entity-Attribute-Value Model</td>
</tr>
<tr>
<td>1.5 Data Quality</td>
</tr>
<tr>
<td>1.6 Data, Information, and Knowledge</td>
</tr>
<tr>
<td>1.7 Summary</td>
</tr>
</tbody>
</table>

| 2 Symbolic Biomedical Knowledge             |
| 2.1 Biomedical Theories and Computer Programs |
| 2.1.1 A World Class Reasoning Example      |
| 2.1.2 Biological Examples                 |
| 2.1.3 Symbolic Theories                   |

<table>
<thead>
<tr>
<th>Foreword</th>
</tr>
</thead>
<tbody>
<tr>
<td>xiii</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Preface</th>
</tr>
</thead>
<tbody>
<tr>
<td>xvii</td>
</tr>
</tbody>
</table>
2.2 Logic and Inference Systems
2.2.1 Predicate Calculus 106
2.2.2 Unsound Inference: Abduction and Induction 124
2.2.3 First Order Logic 125
2.2.4 Rule-based Programming for FOL 133
2.2.5 Limitations of FOL 139

2.3 Frames, Semantic Nets, and Ontologies
2.3.1 A Simple Frame System 143
2.3.2 Extensions to Frame Systems 159
2.3.3 Frame System Implementations, Interfaces, and APIs 162
2.3.4 Frames and Object-oriented Programming Languages 162

2.4 Description Logics
2.4.1 Description Logic Languages 166
2.4.2 Examples from GALEN 167
2.4.3 The Semantic Web 170
2.4.4 Back to the Future 173

2.5 Search
2.5.1 A General Framework for Search 174
2.5.2 Uninformed Search Strategies 178
2.5.3 Informed Search Strategies 180
2.5.4 Path Search 182

2.6 Summary 183

3 Probabilistic Biomedical Knowledge
3.1 Probability and Statistics
3.1.1 Probability 186
3.1.2 Statistics 187
3.1.3 The Laws of Probability 188
3.1.4 Conditional Probability 190
3.1.5 Independence 191
3.1.6 Random Variables and Estimation 192

3.2 Application and Generalization of Bayes' Rule
3.2.1 Simple Bayesian Inference 201
3.2.2 Non-Boolean Variables 202
3.2.3 Bayes Nets 203

3.3 Utility and Decision Modeling
3.3.1 A Decision Analysis Vignette 208
3.3.2 Graph Structure and Probability Assignment 209
3.3.3 Determining Utilities 210
3.3.4 Computation of Expected Values 211

3.4 Information Theory
3.4.1 Encoding of Messages 213
3.4.2 Entropy and Information 216
3.4.3 Efficient Encoding 217
3.4.4 Error Detection and Correction 220
3.4.5 Information Theory in Biology and Medicine 220

3.5 Summary 221
Contents

4 Biomedical Information Access 223
  4.1 Information Retrieval Systems 223
  4.2 IR System Design 227
    4.2.1 Indexing 227
    4.2.2 Processing Queries 232
    4.2.3 Searching and Matching 233
    4.2.4 Ranking the Results 233
    4.2.5 Performance Evaluation 234
  4.3 Intelligent Query Processing 235
  4.4 Network Access to Document Repositories 235
  4.5 Natural Language Processing 246
  4.6 Beyond Text Documents 248
    4.6.1 Biomedical Images 249
    4.6.2 Electronic Medical Records 250
  4.7 Summary 250

Part II
Biomedical Ideas and Computational Realizations

5 Computing With Genes, Proteins, and Cells 253
  5.1 Getting Data from Files 255
  5.2 Computing With DNA Sequences 259
  5.3 Mapping Genes to Proteins 264
  5.4 Computing With Protein Sequences 268
    5.4.1 Representing Amino Acid and Proteins 268
    5.4.2 Simple Properties of Proteins from Sequence Data 271
    5.4.3 Amino Acid Profiles 272
  5.5 The Gene Ontology 277
  5.6 Biochemical Pathways 280
    5.6.1 Searching for Pathways 281
    5.6.2 Biochemical Reachability Logic 287
  5.7 Simulation With State Machines 291
    5.7.1 State Machines 291
    5.7.2 A Simple State Machine 293
    5.7.3 Simulating the Cell Cycle in Yeast 297
  5.8 Summary 302

6 Modeling Biological Structure 303
  6.1 The UW Foundational Model of Anatomy 304
    6.1.1 The Components of the FMA 304
    6.1.2 Representing Anatomical Relations in the FMA 305
    6.1.3 Metaclases 305
    6.1.4 The part-of Relationship 308
  6.2 A Simple Network Interface – the FMS 308
## Contents

6.3 Semantic Consistency Checking and the FMA  
6.3.1 The Lymphatic System: An Example  
6.3.2 Current State of the Lymphatics in the FMA  
6.4 Summary  

7 Drug Interactions  
7.1 Drug Information Catalogs  
7.2 Reasoning About Pharmacokinetics  
7.2.1 Procedural Knowledge About Enzymatic Metabolism  
7.2.2 A Rule-based Drug Interaction System  
7.2.3 Truth Maintenance Systems  
7.2.4 A Drug Interactions Theory with Justifications  
7.2.5 Evidence-based Reasoning  
7.3 Reasoning About Pharmacodynamics  
7.4 Summary  

8 Medical Data Communication  
8.1 EMR Systems and PACS  
8.1.1 EMR Systems  
8.1.2 Medical Images and PACS  
8.2 Networks and Protocols  
8.2.1 Network Architecture  
8.2.2 Network Client and Server Design  
8.3 HL7  
8.3.1 The HL7 Message Exchange Protocol  
8.3.2 HL7 Message Structure  
8.3.3 Scanning HL7 Messages  
8.3.4 HL7 Message Field Content  
8.4 DICOM: Medical Image Information Agents  
8.4.1 A Short History of DICOM  
8.4.2 About the DICOM Standard  
8.5 How to Implement DICOM  
8.5.1 The DICOM State Machine  
8.5.2 Parsing and Generation of PDUs  
8.5.3 Parsing and Generation of Commands and Data  
8.5.4 How It All Turned Out  
8.6 Summary  

9 Cancer Radiotherapy Planning  
9.1 Radiation Therapy  
9.2 Radiotherapy Planning Software  
9.3 Locating the Target  
9.3.1 Computing the CTV  
9.3.2 Computing the PTV  
9.4 Influence Diagrams in RTP  
9.5 Summary
### Contents

10 Safety and Security  
10.1 A Bit of Personal History and Perspective 418  
10.2 Some Real-Life Stories 419  
10.2.1 Radiation Therapy Planning Systems 420  
10.2.2 Who is Master, the Human or the Computer? 422  
10.2.3 The Therac-25: A Real-Time Disaster 422  
10.2.4 Prism 423  
10.2.5 The UW Neutron Therapy Facility 423  
10.2.6 DICOM and Internet Security 424  
10.2.7 Intravenous Infusion Pumps on the Internet 424  
10.3 Safety by Design: Radiotherapy Machines 425  
10.3.1 A Simple Linac Control System 426  
10.3.2 Adding a Dosimetry System 428  
10.3.3 Formal and Automated Analysis 429  
10.4 IT Security in Health Care Organizations 429  
10.5 Epilogue 431  

#### Appendix

A Lisp Notes, Software, and Other Resources 433  
A.1 Lisp Notes 433  
A.1.1 General Notes on Lisp 433  
A.1.2 CLOS and MOP 434  
A.1.3 Optimization 436  
A.1.4 Books on Lisp 437  
A.2 Software 438  
A.2.1 Data and Knowledge Resources 438  
A.2.2 Common Lisp Systems, Libraries, and Resources 439  
A.2.3 Languages and Tools for Knowledge Representation 439  
A.2.4 Code from the Book 440  
A.3 Other Resources 441  

Bibliography 443  

Index 459