# LIST OF FIGURES

1.1 Geometric models of protein surfaces. 3

1.2 Geometry of a simplified two dimensional model molecule. 5

1.3 The family of alpha shapes or dual simplicial complexes for a two-dimensional toy molecule. 7

1.4 An illustration of *a family of alpha shapes of HIV-1 protease* and *flips*. 8

1.5 An example of analytical area calculation. 10

1.6 Discrete flow of empty space illustrated for two dimensional disks. 12

1.7 The computed surface pockets of binding sites on Ras21 protein and FtsZ protein. 12

1.8 An illustration of *locally Delaunay edge* and *flips*. 13

1.9 Scaling of voids and pockets of a set of 636 proteins representing most of the known protein folds. 15

1.10 Protein function prediction as illustrated by the example of alpha amylases. 19

2.1 The Miyazawa-Jernigan model of chemical reaction. 36

2.2 Schematic drawing of the Delaunay complex and the alpha shape of a two-dimensional molecule. 45

2.3 Schematic illustration of non-interacting pairs of residues. 46

2.4 Geometric views of the inequality requirement for protein scoring function. 49

2.5 Recognition of binding surface patch of protein targets using geometric potential function. 55

3.1 The Ising model of 30×30 size, with a total of $30 \times 30 = 900$ sites. 80

3.2 Illustration of rejection sampling. 81

3.3 Generating self-avoiding chain by sequential importance sampling. 88

3.4 An example of a phylogenetic tree. 91

4.1 The stochastic network of a toggle switch. 120

4.2 The steady state probability landscape of a toggle switch. 120

4.3 Different selection of cell fate of *E. coli* infected by phage lambda and a model of the epigenetic circuit for lysogeny maintenance. 122

4.4 The probability landscape of the epigenetic circuits of lysogeny maintenance in phage lambda. 123

4.5 Instability, shallow threshold, and switching inefficiency of the network against fluctuation in UV irradiation in mutant phage lambda. 124

5.1 Illustration of the labeled directed graph representation for a molecular interaction network. The arc $\mathsf{B} \dashv \mathsf{A}$ indicates a negative influence of B on A, *i.e.*, an increase in the amount of protein B causes a decrease in the amount of protein A. The pathway $\mathsf{B} \to \mathsf{C} \to \mathsf{A} \dashv \mathsf{D}$ induces a positive influence of B on D since the product of labels of its arcs is $1 \times (-1) \times (-1) = 1$. 135

5.2 A Boolean circuit composed of logical AND, OR and NOT gates that encodes relationships between three proteins and two genes. For example, either Protein B must be absent or Protein C must be present (or both) to activate Gene Y. 135

5.3 **(a)** A Boolean network with three binary states $s_1, s_2, s_3$. **(b)** The associated directed graph. A fixed point of the network is given by $\vec{s} = \big(s_1, s_2, s_3\big) = (0, 1, 0)$. 138

5.4 An algorithmic framework for synthesizing signal transduction networks [5]. The optimization steps involving Tr and Pnc are explained in Section 5.3.3. 141

5.5 Dynamic programming algorithm to find all reachabilities. 143

5.6 Pictorial illustration of the iterative calculations of the dynamic programming algorithm in Fig. 5.5. 143

5.7 The transitive reduction (Tr) problem. 145

5.8 An example of obtaining a reduced network via transitive reduction. The obtained network is not minimal (see Exercise 5.4). 145

5.9 A greedy algorithm to solve Tr. 146

5.10 An example of a family of graphs for which the greedy algorithm has an approximation ratio of 2. The greedy algorithm may remove the arcs $v_i \to v_{i+1}$ for $i = 1, 2, \ldots, n-1$ providing a solution with $2n$ arcs, but an optimal solution with $n+1$ arcs is possible by selecting the arcs $v_0 \to v_1$, $v_i \to v_{i+1}$ for $i = 1, 2, \ldots, n-1$, and $v_n \to v_0$. 147

5.11 The pseudo-node collapse (Pnc) problem [5]. 148

5.12 A system of seven elements. 150

5.13 Equivalence of dynamical properties may depend on node functions. 152

5.14 **(a)** The Markov-chain algorithm for generating random networks by arc swapping. **(b)** A pictorial illustration of arc swapping. 155

5.15 A schematic diagram for the overview of the Mra approach. 156

5.16 Linear algebraic formulation of the experimental design question for the Mra approach. 158

5.17 A combinatorially equivalent reformulation of (5.6). 161

5.18 Two well-known algorithms to solve $\mathbf{SC}_1$ [84]. 161

5.19 Improved randomized approximation algorithm for $\mathbf{SC}_\gamma$ [14]. 162

5.20 **(a)** Measurements of expression levels of 5 genes $\mathscr{G}_1$, $\mathscr{G}_2$, $\mathscr{G}_3$, $\mathscr{G}_4$ and $\mathscr{G}_5$ at two successive time steps; variable $x_i$ correspond to gene $\mathscr{G}_i$. **(b)** A causal relationship and Boolean formula that explains the causal relationship of other variables to $x_5$ based only on the data shown in **(a)**. **(c)** Another causal relationship and Boolean formula for $\mathscr{G}_5$ that is consistent with the data in **(a)**. 163

5.21 **(a)** Data matrix $X = \left(x_{i,j}\right)$ (quantized to four values) for measurement of expression levels of $m = 5$ genes at $n + 1 = 4$ time points. **(b)** The universe and sets corresponding to gene $\mathscr{G}_2$ in the hitting set formulation of Fig. 5.22**(a)**. 164

5.22 **(a)** A hitting set formulation of the combinatorial approach for gene $\mathscr{G}_i$. **(b)** A greedy algorithm for **HS** that iteratively selects a new element of the universe that hits a maximum number of sets not hit yet. 165

5.23 Two dimensional Roc space obtained by plotting FPR versus TPR values. 168

5.24 Three $n$-node graphs (shown for $n = 8$) discussed in Exercise 5.6. 175

5.25 Contraction of a cycle of length 4. 176

6.1 A Boolean network of two species interaction. 181

6.2 A discrete-time sigmoidal neural network and its graphical representation. 182

6.3 **(a)** The threshold gate function. **(b)** The sigmoidal gate function. 182

6.4 **(a)** A continuous-state discrete-time ANN with a continuous gate function $g$. **(b)** The difference equation model corresponding to the ANN in (a). $R_i$ is the maximum rate of synthesis of gene $v_i$, $\lambda_i$ is the degradation rate of the product from gene $v_i$, and the threshold $\theta_i$ summarizes the effect of general transcription factors on gene $v_i$. **(c)** The specific activation function $g$ used by Kyoda *et al.* [62]. **(d)** The topological version of the model in **(b)** indicating excitory and inhibitory effects. 186

6.5 Rectangular partition of state space induced by system (6.7). 190

6.6 **(a)** An example of a NFA. The input $0101 \in \Sigma^*$ is accepted by the NFA since there is a directed path from the initial state $q_0$ to a final state $q_2$ labeled 0101. **(b)** An example of a piecewise-linear hybrid automata (PLHA) with two continuous state variables and no inputs. A hypothetical trajectory of the dynamics is shown by thick black lines with arrows. 191

6.7 Hexagonal lattice of cells for Delta-Notch protein signalling. A cell with its six neighbors is shown amplified. 192

6.8 A system composed of five interconnected subsystems. 193

6.9 Two signed graphs. The graph in **(a)** is sign-consistent, but the graph in **(b)**, which differs in just one edge from **(a)**, is not sign-consistent since it has two paths in its undirected version with different parity between nodes $x_1$ $x_4$, namely a direct path of odd parity and a path of even parity transversing node $x_5$. Self-loops, which in biochemical systems often represent degradation terms, are ignored in the definition. 197

6.10 Definition of the sign consistency problem. 198

6.11 Deletion of the arc $(x_2, x_4)$ makes the given signed graph consistent. The node labels are shown besides the nodes. 198

7.1 A schematic diagram of the early development of a *Drosophila* embryo. Each hexagon represents a cell, and neighboring cells interact to form a collective behavior. In this figure, an initial striped pattern of the genes en and wg induces the production of the gene hh, but only in those cells that are producing en. 210

7.2 [8] The *Drosophila* segment polarity regulatory network for one cell with the interpretation of the regulatory role of PTC on the reaction CI → CN as PTC → CN and PTC ⊣ CI. 211

7.3 A 1-dimensional 3-cell *Drosophila* segment polarity regulatory network with cyclic boundary conditions. 212

7.4 A pictorial representation of the network manually synthesized by Li *et al.* [15]. Nodes in the network correspond to proteins, genes and other small molecules (*e.g.*, RAC1 is a small GTPase protein). 213