

Contents

<i>Preface</i>	<i>page</i>	<i>xi</i>
1 Introduction		1
1.1 Original problems		2
1.2 Additional problems		5
1.3 Further reading		23
2 Pairwise alignment		24
2.1 Original problems		24
2.2 Additional problems and theory		43
2.2.1 Derivation of the amino acid substitution matrices (PAM series)		46
2.2.2 Distributions of similarity scores		57
2.2.3 Distribution of the length of the longest common word among several unrelated sequences		62
2.3 Further reading		65
3 Markov chains and hidden Markov models		67
3.1 Original problems		68
3.2 Additional problems and theory		77
3.2.1 Probabilistic models for sequences of symbols: selection of the model and parameter estimation		86
3.2.2 Bayesian approach to sequence composition analysis: the segmentation model by Liu and Lawrence		95
3.3 Further reading		102
4 Pairwise alignment using HMMs		104
4.1 Original problems		105
4.2 Additional problems		113
4.3 Further reading		125