


References


Index

\( \Omega(f(n)) \) 30
\( \Theta(f(n)) \) 30
\( \lambda \) see string, empty
2-factor 209

accepted mutation 95
adenine 10
adjacent 25
algorithm 28
   consistent 32
alignment 81ff.
   compatible 109ff.
   global 84ff.
   local 84, 90
   multiple 101ff.
   of two strings 82ff.
   computation 88
   optimal local 90
   score 83
   semiglobal 91
alignment tree 110
alphabet 23
amino acid 7
   hydrophilic 353
   hydrophobic 353
   polar 353
approximation algorithm 32f.
approximation ratio 32
arc 338
arc set 338
asymptotic approximation 362
autocorrelation 224
   polynomial 225
automaton 39
backtracking 135
bad character rule 44
   preprocessing 45
base pair
   reachable 322
   valid 320
Bernoulli string 224
binary tree 28
   directed 28
   undirected 28
binding site 213
bit score 100
BLAST 99ff.
BLOCKS database 97
BLOSUM matrix 97
Boyer–Moore algorithm 49
branch and bound 368
breakpoint 241
bulge 322
C1F see consecutive ones form
C1P see consecutive ones property
Celera Genomics 120
CG-islands 228, 233
CHAIN 341
chain termination method 18
character see symbol
character matrix 269
   binary 269
child 27
chimeric clone 16, 158, 174
Chomsky normal form 332
chord 26
Index

chromosome 12
clique problem 342
close 144
close library 144
cloning 15
CNF see Chomsky normal form
column distance graph 161
COMPACTAddTree problem 266
companion see companion column
companion column 293
companion graph 298
companion pair
  equated 293
  negated 293
companion row 293
complete pair matrix 290
component graph 303
compression 177
computing problem 28
concatenation 24
connected component 299, 346
consecutive ones form 146
consecutive ones property 146
testing for 156
consensus 102, 173
  ε-consensus 199
distance to 102
consensus string approximation 215
consensus string problem 214
contact 354
edge 353ff.
contig 198
cost function 29
cost measure 54
  logarithmic 54
  uniform 54, 69
coverage 172
  incomplete 174
Crossing 341
cycle 26
  simple 26
cycle cover 190
cycle cover superstring algorithm 191
CYK algorithm 335, 373
  stochastic 335
cytosine 10
database search 97ff.
DDP 127
debovo protein structure prediction 352ff.
Dec(0,1)ShortestSuperseq problem 105
DecDDP 128
decision problem 28
  corresponding 32
DecLAPCS 343
DecMaxCut problem 366
DecMULTSPALIGN problem 105
DecSCS 182
deletion 19, 83
denaturation 15
deoxyribonucleic acid see DNA
derivation 329
  rule see production
tree 333
diagonal run 98
digest 124
diploid 287
disjoint DDP 131
distance 179
distance graph
  for SCS 179
  for the COMPACTAddTree problem 266
distance set 127
divide and conquer 219
DNA 10
DNA array see DNA chip
DNA chip 19
DNA contamination problem 63
DNA fingerprinting 217
DNA sequencing 119ff.
domain 351
double digest approach 124ff.
double digest problem see DDP
dynamic programming 84, 104, 231,
  314, 324
  edge 25
  directed 27
  edge labeling 28
  edit distance 84
  edit graph 88
  embedding 353
  emission probability 228
  Eulerian cycle 26, 202
  computing an 205
Eulerian path 26, 202
EXPERFPHYL problem 270
false negative 20, 145
false positive 20, 145
FASTA 98f.
feasible solution 29
fingerprint 143
fingerprinting 143
fission 249f.
Fitch algorithm 277f.
folding point 359
forcing pattern 294
forcing component 299
forcing graph 298
forcing pair 294
fragment assembly problem 172
fragment conflict 306
fragment conflict graph 306
frequency analysis 224
full digest 125
function
  polynomially bounded 30
function computing problem 29
fusion 249f.
gap 82
gap scoring 94
gap symbol 82
gel electrophoresis 16
gene 12
genome 13
genome rearrangements 237ff.
genotype 287f.
genotype matrix 289
  regular 293
genotypical characters 268
global alignment problem 84
GMnP 159
good suffix rule 44
  preprocessing 49
grammar
  context-free 329
  stochastic context-free 329, 331
graph 25ff.
  bipartite 26
  complete 25
  connected 26
  directed 27
edge-weighted 28
Eulerian 203
greedy method 182
greedy superstring algorithm 183
grid lattice 353
guanine 11
hairpin loop 322
Hamiltonian cycle 26
Hamiltonian cycle problem 28, 31
Hamming distance 162, 214, 277
haplotype 287f.
haplotype matrix 289
haplotyping 287ff.
helix 350
heterozygous 288
hidden Markov model 228ff.
  high scoring pair 100
hit 100
HMM see hidden Markov model
HMM decoding problem 231
homologous genes 239
homology 13
homomorphism 25
homozygous 288
host 15
hot spot 98
HP model 352ff.
HP(d) problem 355
HSP see high scoring pair
Human Genome Project 120
hybridization 11, 15
hybridization data 143
hybridization matrix 144
incident 25
indicator function 296
  feasible 296
insert 15
insertion 19, 83
inside-outside algorithm 373
interchromosomal transformations 238
interior loop 323
interval graph 169
intrachromosomal transformations 238
inversion see reversal
Landau symbols 30
LAPCS 340
lattice 353
layout 173
lce 217
LCS 338
leaf 27
left-diverse 67
left-symbol 67
letter see symbol
Levenshtein distance see edit distance
lexicographical order 69
link 199
weakest 199
local alignment problem 90
local search 281
longest arc-preserving common subsequence problem see LAPCS
longest common extension 217
longest common extension problem 217
longest common subsequence problem see LCS
longest common substring problem 61
loop 323, 350
lowest common ancestor 218
magic word 214
majority voting 103, 174
mapping
by hybridization 143ff.
optical 168
restriction site- 141, 143
with non-unique probes 165f.
with unique probes 146ff.
marker 120, 123
match 83
mate pair 120
maximum compression common
superstring problem see MCCS
MAXQUARTETTCONSIST problem 282
MbH see mapping, by hybridization
MCCS 177
median string 214
merge 25
metric 251, 258
MFR problem 309
minimum fragment removal problem see MFR problem
minimum SNP removal problem see MSR problem
minimum vertex bipartizer problem 310
MinOSR problem 248
MinPARTOP problem 280
MinSR problem 240
approximation 245
MinSYNTENY problem 251
approximation 254
mismatch 83
motif 351
MSR problem 309
MULTCONSENSUSALIGN problem 103
multicontig model 198
multicontig problem 199
multiple loop 323
MULTSPALIGN problem 104
neighbors
connected 354
topological 354
NESTED 341
non-terminal 329
nondeterminism 31
NP 31
NP-complete 31
NP-hard 31ff.
nucleic acid 9ff.
nucleotide 10
Nussinov algorithm 324f.
O(f(n)) 30
optimization goal 83
optimization problem 29
ordering constraints 365
orientation
unknown 174
ov 25
overlap 25, 173
approximate 92f.
computation 93
computation 29, 64
empty 25
generalized 41
overlap graph 179
overlap-layout-consensus scheme 172
P 31
PAM distance 95
PAM matrix 95
PAM unit 95
parameterized complexity 373
parent 27
parsimony principle 275ff.
parsimony problem 276
parsimony score 277
partial digest 131
partial digest approach 131ff.
  ideal data 131
partial digest problem see PDP
  atomic distances 134
  backtracking algorithm 136
  level 134
path 26
  directed 27
  in a Hidden Markov Model 228
  simple 26
pathlabel 54
pattern 37
PDP 133
perfect phylogeny haplotyping problem
  see PPH problem
PerfPhyl problem 270
permutation 239
  directed 239, 247
  extended representation 241
  signed 239
  undirected 239
phenotypical characters 268
phylogenetic tree 257ff.
  perfect 269, 289
  construction 272
  existence 275
  undirected perfect 290
  unrooted 276
phylogeny see phylogenetic tree
physical map 120, 123
physical mapping 123ff.
PLAIN 341
point set 133
polynomial-time approximation scheme
  see PTAS
polynomial-time reduction 32
polypeptide chain 9
Pos 54
position set 126
PPH problem 290
PQ-tree 147
  consistent permutation 149
empty 149
empty vertex 151
equivalent 149
front 148
full vertex 151
legal operation 148
partial vertex 151
pattern 150
pertinent subtree 151
pertinent vertex 151
reduction 149
substitute 150
universal 149
Pref 25
pref 25
prefix 24
preprocessing 39
primary structure
  of proteins 350
  of RNA 320
primer 16
probe 20, 144
production 329
protein 7ff.
  core model 364
  de novo structure prediction 352ff.
  folding 349
  HP model 352ff.
  primary structure 9, 350
  quaternary structure 351
  secondary structure 350
  structure model 364
  structure prediction 349ff.
  tertiary structure 350
protein threading 363ff.
  branch and bound 370
protein threading problem 366
pseudoknot 323
PTAS 217
quartet 281
  consistent 281
  optimal 281
quartet method 281ff.
quartet puzzling 282f.
quaternary structure
  of proteins 351
read 18
reciprocal translocation 238
reconstruction model 197
reconstruction problem 197
relation computing problem 29
repeat 66, 174
approximate 66
exact 66
maximal exact 66
compact representation 68
repeat search problem 66
resolving pair 288
restriction enzymes 15, 123
restriction site 15, 123
restriction site mapping 123ff.
reversal 238
for directed permutation 248
for undirected permutation 239
reverse complement 11
ribonucleic acid see RNA
RNA 10
primary structure 320
secondary structure 320
substructure 322
tertiary structure 321
RNA substructure
size of 323
root 27
SbH see sequencing, by hybridization
SbH reconstruction problem 202
scoring function 83
good 112
scoring matrix 94f.
SCS 176
secondary structure
of proteins 350
of RNA 320ff.
self-avoiding walk 353
sequence 23
sequencing see DNA sequencing
by hybridization 171, 201ff.
shotgun 171ff.
sequencing errors 18
set partition problem 128
sheet 350
shortest common superstring problem 176
decision version 182
with compression measure 178
with length measure 178
shortest covering string problem 165
shotgun method 120
shotgun sequencing see sequencing,
shotgun
signals 213ff.
similarity 83
computation 87
similarity matrix 84
simulated annealing 281
single nucleotide polymorphism see SNP
skew algorithm 72f.
SNP 288ff.
SNP conflict 311
SNP matrix 306
gapless 311
reduced 311
weight 313
SNP site 288
sorting
counting sort 71
radix sort 72
SP-score 103
spacing constraints 365
spanning tree algorithm 33
spectrum 201
spectrum graph 202
stack 135
stacked pair 322
star alignment 110f.
state 40
accepting 40
initial 40
stem 322
string 23ff.
arc-annotated 337f.
compatible 201
empty 23
length 23
reverse complementary 11
simply compatible 201
string depth 54
string homomorphism see homomorphism
string matching 37ff.
approximate 91
Boyer–Moore algorithm 44ff.
naive approach 38
Index

with automata 39ff.
with suffix trees 57
string matching automaton 40f.
construction 42
string matching problem 38
naive approach 38
strip 242
ascending 242
descending 242
STS probes 146
subgraph 26
induced 26
subsequence 24
arc-preserving common 339
consistent mapping 339
longest arc-preserving common 340
longest common 338
substitution 19, 83
substring 24
frequent 223
infrequent 223
longest common 61
substring edit distance 197
substring free 177
substring problem 58
Suff 25
suff 25
suffix 24
suffix array 68ff.
inverse 73
suffix similarity 46
suffix tree 50ff., 218
compact 53
construction 55
generalized 59
simple 50
construction 51
sum of pairs see SP-score
supersequence 105
shortest common 105
superstring
induced 181
shortest common 176
trivial 177
symbol 23
syntenic distance 249f.
syntenic operations 250
synteny graph 253
synteny problem 251
T-contig 199
tandem repeat 217
approximate 217
computation 222
tandem repeat problem 219
tandem repeats 217ff.
taxa 257
taxon see taxa
terminal 329
tertiary structure
of proteins 350
of RNA 321
text 37
threading 365
three point condition 259
threshold problem 32
thymine 11
Time 30
transition function 40
transition probability 228
translocation 248ff.
transposition 238
traveling salesman problem see TSP
tree 27
additive 265ff.
compact additive 266
computation 267
directed 27
ordered 27
rooted 27
spanning 27
ultrametric 259
distance 259
tree compatible 271
triangle inequality 33, 112, 164, 259
trie see suffix tree, simple
tRNA 321
TSP 29, 164
metric 33
turnpike reconstruction problem 168
ultrametric 259
UNLIMITED 341
UPGMA algorithm 261, 263
uracil 11

variance 224
vertex 25
balanced 204
degree 25
indegree 27
inner 27
outdegree 27
semi-balanced 204
vertex labeling 28
vertical run 313
Viterbi algorithm 232
Watson–Crick complement 11
whole genome shotgun approach 120
Zuker algorithm 326f.
Natural Computing Series


A.A. Freitas: Data Mining and Knowledge Discovery with Evolutionary Algorithms. XIV, 264 pages, 74 figs., 10 tables. 2002


L. Sekanina: Evolvable Components. From Theory to Hardware Implementations. XVI, 194 pages. 2004

G. Ciobanu, G. Rozenberg (Eds.): Modelling in Molecular Biology. X, 310 pages. 2004


M. Amos: Theoretical and Experimental DNA Computation. XIV, 170 pages, 78 figs. 2005


