Index

accumulation point, 176
address in a tree, 52
algorithm
$A_{k-REV}$, 227
$A_{k-TSS}$, 220
ALERGIA, 340
ALERGIA-COMPATIBLE, 340
ALERGIA-TEST, 339
BACKWARD, 92
BACKWARD, 92
BAUM-WELCH, 362
BUILD-PTT, 377
CONSISTENT, 319
CONSTRUCTING a PFA from a FFA, 332
DSAI-COMPATIBLE, 346
EDSM, 292, 293
ELIMINATING $\lambda$-transitions, 97
EM, 361
EM, 361
Encoding a fraction as a string, 205
enumeration, 152
Enumeration identification, 203
FORWARD, 91
generating a string with a PFA, 90
GOLD, 243
halving, 166
INITIALISE-K-REV-CF, 310
INSIDE, 101
INSIDE – OUTSIDE, 367
K-REV-CF, 310
$k – TSS$ machine to DFA, 219
LARS, 319
LSTAR, 275
LSTAR-CLOSE, 276
LSTAR-CONSISTENT, 276
LSTAR-INITIALISE, 275
LSTAR-USEEQ, 276
MDI, 352
MDI-COMPATIBLE, 352
MDL, 289
MERGE, 226
NEWRULE, 318
ONWARD-PTT, 378
OSTIA, 376, 384
OSTIA-FOLD, 382
OSTIA-MERGE, 381
OSTIA-OUTPUTS, 380
OSTIA-PUSHBACK, 381
OUTSIDE, 102
PATTERN_LEARNER, 232
PROMOTE, 243
Reduction between presentations, 158
RPNI, 255
STERN-BROCOT identification, 204
STOCHASTIC-FOLD, 335
STOCHASTIC-MERGE, 334
Transforming the $\lambda$-PFA into a $\lambda$-PFA with just one
initial state, 96
VITERBI, 93
alphabet, 48
auxiliary, 46
final, 46
argmin, 46
automaton
deterministic finite, 74
derived, 127
deterministic with look-ahead, 224
frequency finite, 331
$k$-reversible, 224
maximal canonical, 125
$\mu$-distinguishable, 338
non-deterministic finite, 71
parity, 132
probabilistic finite, 87
pushdown, 82
quotient, 124
reversal, 223
universal, 126

ball
centre, 65
good, 65
of strings, 65
radius, 65

\(\text{CFL}(\Sigma)\), 78
characteristic sample, 154
characteristic set, 178
co-emission probability, 109
cocone, 65
complement of a language, 50
cone, 65
consistent NFA, 72
convergence point, 150
CS, 154
derivation
\(k\)-deep, 307
DFA, 74
DFFA, 331
distance, 47
cost matrix, 57
edit, 56
Hamming, 55
\(L_1(x, y)\), 47
\(L_2(x, y)\), 47
\(L_\infty(x, y)\), 47
metric, 47
prefix, 56
distribution, 86
regular, 95
DFFA, 89
DSAI, 346
e-good hypothesis, 200
Euclidean metric, 47
EX, 163
expansiveness, 80
expectation maximisation, 361

FFA, 331
finite elasticity, 176
finite thickness, 177
frequency, 331
frequency prefix tree acceptor, 333
\(\mathcal{G}\), 144
genetic algorithm, 283
grammar, 77

dcontext-free, 78
deterministic linear, 81
even linear, 81
even linear, 302
\(k\)-reversible context-free, 309
linear, 81, 302
pure, 314
very simple, 303
grammar class
algorithmically minimisable, 178
\(k\rightarrow \mathcal{REVG}(\Sigma)\), 225
graph colouring, 295

hierarchical order, 49
hyper-plane, 233

identification in the limit
behaviourally correct, 150, 175
from queries, 187
from text, 175
IPE-polynomial time, 155
MC-polynomial time, 155
polynomially with queries, 164
strong polynomial with probability one, 206
weak polynomial with probability one, 205
with probability one, 202
\(\text{ifdef}\), 45
implicit prediction error, 155
incremental learning, 266
infinite elasticity, 176, 181
INSIDE probability, 367
IPE, 155

\(k\)-ancestor, 308
\(K\)-contexts, 307
\(k\)-contexts, 308
\(k\)-reversible, 309
kernel
all \(k\)-subsequences, 62
all-subsequences, 62
gap-weighted subsequences, 63
string, 60

\(L\), 50, 144
\(L_1(x, y)\), 47
\(L_2(x, y)\), 47
\(L_\infty(x, y)\), 47
\(L\), 144
\(L_{F_A}\), 71
\(L_{F_R}\), 71
label in a tree, 52
\(\lambda\), 48
language, 49
complement, 50
infinitary, 53
$k$-testable, 218
prefix-closed, 50
probabilistic, 86
regular, 74
suffix-closed, 50
symmetric difference, 50
language class, 50
language class $\mathcal{CFG}(\Sigma)$, 78
$\mathcal{DB}(\Sigma)$, 160
$\mathcal{DETP}$, 95
$\mathcal{DFA}(\Sigma)$, 74
$\mathcal{ELG}(\Sigma)$, 161
$\mathcal{lcs}(L)$, 50
learner, 149
conservative, 155
consistent, 155
strongly consistent, 238
weakly consistent, 238
learning algorithm, 149
length-lex order, 49
length-lexicographic order, 49
lexicographic order, 49
limit point, 176
$\mathcal{MAT}$, 269
maximum likelihood, 361
ML, 361
MCA, 125
metric, 47
mind changes, 155
multiset, 46
$\mathcal{N}$, 46
$[n]$, 46
naming function, 71, 74
NFA, 71
non-erasing pattern, 231
normal form
Chomsky, 79
proper, 79
quadratic, 79
reduced, 78
$\mathcal{NP}$, 47
$\mathcal{O}$, 46
onward transducer, 377
order
hierarchical, 49
length-lex, 49
prefix, 49
OUTSIDE probability, 367
overall polynomial time, 153
$\mathcal{P}$, 47
PAC-learning, 200
polynomial, 200
partition
block, 46
class, 46
finer, 46
lattice, 124
pattern, 230
$\mathcal{PCFG}$, 100
perplexity, 111
$\mathcal{PFA}$, 87
polynomial number of mind changes, 155
polynomial characteristic sample, 154
polynomial update time, 154
polynomially bounded learner, 187
$\mathcal{POMPD}$, 356
$\mathcal{PPTA}$, 104
$\mathcal{PRED}_k$, 223
$\leq_{\text{pref}}$, 49
prefix distance, 56
prefix order, 49
prefix tree acceptor, 77, 238
$\mathcal{Pres}$, 147
$\phi$, 147
presentation, 147
consistent, 148
informant, 148
text, 148
$\mathcal{PPTA}$, 77, 238
$\mathcal{P}$, 47
problem
$\mathcal{SAT}$, 133
consistency, 119
PTA, 77
query
correction, 163, 193
EQ, 163
extended membership, 163
extended prefix language, 163
MQ, 163
sampling, 163
specific sampling, 163
subset, 163
WEQ, 163

\( \mathbb{R} \), 46
rational transducer, 373
reasonable encoding, 152
\( \text{REGP}(\Sigma) \), 95
regular expression, 76
regular reduction, 303
relative frequency, 331
row
compatible, 246
obviously different, 246
sample, 76
informed, 238
sequential transducer, 374
set
argmax, 46
cardinality, 46
equivalent classes, 46
partition, 46
prefix, 50
prefix-closed, 243
suffix, 50
suffix-closed, 243
\( \Sigma^* \), 48
state
BLUE, 240
RED, 240
WHITE, 240
accepting, 70
candidate, 240
final, 70
initial, 71
neutral, 70
rejecting, 70
sink, 75
string, 48
distinguishing, 338
k-predecessor
k-predecessor, 223
k-successor, 223

longest common prefix, 50
longest common suffix, 50
shortest prefix, 259
subsequence, 48
substring, 48
superstring, 48
structural completeness, 127
subsequential transducer, 375
SSQ, 163
SUCC, 223
support, 46
symmetric difference between two languages, 50
symmetrical structural completeness, 128
table
closed, 246, 273
consistent, 247, 272, 273
equivalent prefixes, 272
equivalent rows, 272
hole in a, 245, 270
observation, 244
redundant, 244
Tabu search, 286
tell tale set, 177
TEXT presentation, 174
transduction, 372
tree, 51
domain, 51
frontier of, 52
leaf of, 52
node of, 52
root of, 52
subtree, 51, 52
UA, 126
variation metric, 47
VC-dimension, 116, 117
violation (of k-reversibility), 225