

Index

- D-MAP, 270, 274, 275
- D-MAR, 270, 275
- D-MPE, 270, 274
- D-PR, 270, 274
- E-MAJSAT, 271, 274
- MAXSAT, 275
- MAJSAT, 271, 274
- SAT, 271, 274
- W-MAXSAT, 257, 281
- WMC, 277, 278, 280

- a-cutset, 183, 224
- absolute error of estimate, 386
- ADDs, 327
 - and compiling networks, 334
 - and variable elimination, 332
 - operations, 328
- AIC score, 460
- Akaike Information Criterion (AIC), 460
- algebraic decision diagrams, *see also* ADDs, 327
- alleles, 110
- almost simplicial node, 207
- ancestor, 54
- aperiodic Markov chain, 402, 410
- arithmetic circuit, 291
 - compiling from Bayesian networks, 300
 - propagation, 293
 - semantics of partial derivatives, 292
 - with local structure, 319
- asymptotically Normal, 384, 392, 394, 443
- automated reasoning, 1

- balanced dtree, 228
- Bayes
 - conditioning, 32
 - factor, 42
 - rule, 38
 - theorem, 38
- Bayesian approach for learning Bayesian networks, 477
- Bayesian Dirichlet (BD) score, 500
- Bayesian Information Criterion (BIC), 461, 498
- Bayesian network, 57
 - arithmetic circuit representation, 287
 - chain rule, 58, 60
 - CNF encodings, 319
 - CPT, 57
 - dimension, 460
 - dynamic, *see also* DBN, 98
 - family, 57
 - functional, 316
 - granularity, 90
 - instantiation, 57
 - monitors, 78
 - multiply-connected, 108, 143
 - parameter, 57
 - parametrization, 57
 - polynomial representation, 290
 - semantics of partial derivatives, 292
 - polytree, 108, 142
 - probability of evidence, 76
 - pruning, 144
 - simulating, 379
 - singly-connected, 108, 142
 - structure, 57
- Bayesian parameter estimates, 483, 492
- Bayesian scoring measure, 499
- BD score, 500
- BDe score, 501
- BDeu score, 502
- belief propagation, 340
 - algorithm, 163, 175
 - edge-deletion semantics, 354
 - generalized, 349
 - iterative, 343
 - loopy, 344
 - message schedule, 345
 - messages, 342
- belief revision, 4
- BER, 104
- Beta distribution, 491
- Bethe free energy, 363, 364
- BIC score, 461, 498
- binary jointree, 169, 222
- bit error rate, 104
- BP, *see* belief propagation, 340
- branch decomposition, 232, 234
- bucket elimination, 135, 147
- bypass variable, 91, 95, 116

- cache allocation, 192
- candidate method, 486
- case, 442
- case analysis, 37, 178, 326
- causal support, 342
- causality, 12, 53, 55, 56, 64, 70, 85, 88, 92, 102, 112, 115, 116
- CD distance, 418
- CDF, 46, 101
- central limit theorem, 384

- chain rule
 - for Bayesian networks, 58, 60, 131, 274, 403, 415, 447
 - for probability, 37, 39, 60
- channel coding, 102
- Chebyshev's inequality, 382
- chordal graph, 230
- chordality axiom, 68
- circuit complexity, 291, 313
- classifier
 - naive Bayes, 467, 469
 - tree-augmented naive Bayes (TAN), 467
- clause, 21
 - IP, 319
 - PI, 319
 - unit, 21
- clique, 135, 203, 204, 208
 - induced by elimination order, 231
 - maximal, 203, 231
- clone, 255
- cluster, 158, 164, 187, 216, 224
 - sequence, 204
- CNF, 21
 - converting a sentence into CNF, 24
 - encoding of Bayesian network, 277, 279, 280, 319
 - weighted, 280
- co-varying parameters, 422, 479
- code word, 102
- collect phase of jointree algorithm, 162, 167
- compatible instantiations, 21
- compiling Bayesian networks, 287
 - using CNF encodings, 304
 - using jointrees, 302
 - using variable elimination, 301, 334
 - with local structure, 322, 334
- complete graph, 203
- complexity classes, 271
- composition axiom, 59
- conditional
 - independence, 35
 - KL-divergence, 469
 - log-likelihood function, 469
 - mutual information, 37
 - probability table, *see also* CPT, 56
- conditioning, 178
 - a probability distribution on evidence, 31, 32
 - a propositional sentence on a literal, 23
 - recursive, 181
 - using a loop-cutset, 180
 - with local structure, 323
- confidence interval, 382
- conjunctive normal form, *see also* CNF, 21
- consistency, 17
- consistent
 - estimate, 384
 - instantiations, 21
 - sentence, 17
- constrained
 - elimination order, 258
 - optimization, 452
 - treewidth, 259
- constraint-based approach for learning Bayesian networks, 440, 466
- context, 186, 224
- context-specific
 - caching, 324
 - decomposition, 324
 - independence, 313, 316, 324
- continuous
 - distribution, 46, 101, 382, 489–491
 - variables, 12, 46, 101, 105, 491
- contracting edge, 208
- contraction axiom, 61
- convolutional codes, 105
- CPT, 56
 - ADD representation, 327
 - decision tree representation, 117
 - decomposition, 315
 - deterministic, 78, 101, 103, 118
 - large size, 114
 - non-tabular representation, 117, 327
 - rule-based representation, 118
- cumulative distribution function, *see also* CDF, 46
- cutset, 183, 224
- cutset conditioning algorithm, 180
- D-MAP, 69
- d-separation, 63, 65, 86, 94, 223, 364, 454, 506
 - completeness, 68
 - complexity, 66
 - soundness, 67
- DAG, 53
- data set, 442
- DBN, 98, 105, 106, 396, 400
- decision theory, 5, 12
- decoder, 103, 104
- decomposability, 22, 304, 306, 320
- decomposable graph, 230
- decomposition
 - axiom, 59
 - graph, *see also* dgraph, 189
 - tree, *see also* dtree, 183, 224
- deduction theorem, 25
- degeneracy of graph, 208
- degree
 - of graph, 208
 - of node, 203
- degrees of belief, 4, 27
- deleting edges, 255
- Delta method, 413, 414
- dependence MAP, *see also* D-MAP, 69
- depth-first search
 - for cache allocation, 193
 - for computing MAP, 260
 - for computing MPE, 250
 - for finding optimal elimination order, 209
 - for learning network structure, 464
- descendants, 54
- detailed balance, 402
- determinism, 22, 304, 306, 316, 320, 325
- deterministic
 - CPT, 78, 93, 101, 103, 118
 - node, 78
 - variable, 78
- dgraph, 189

- diagnosis, 85, 87, 92
- diagnostic support, 342
- direct sampling, 383, 385
- directed acyclic graph, *see also* DAG, 53
- Dirichlet distribution, 490
- discriminative model, 441, 467
- disjunctive normal form, *see also* DNF, 21
- distance measure
 - CD distance, 418
- distribute phase of jointree algorithm, 162, 167
- distribution
 - Beta, 491
 - continuous, 46, 101, 382, 489–491
 - Dirichlet, 490
 - empirical, 442
 - exponential, 101
 - Gaussian, *see also* Normal distribution, 46
 - importance, 393
 - initial, 402
 - jointree factorization, 223
 - lifetime, 101
 - marginal, 78
 - Normal, 46, 104, 382, 384, 392, 394, 443
 - proposal, 393
 - sampling, 383
 - standard Normal, 46, 47, 382
 - stationary, 402
 - strictly positive, 62
 - transition, 402
- DNF, 21
- dtree, 183, 214
 - balanced, 228
 - for DAG, 224
 - for graph, 234
 - from elimination order, 225
 - properties, 225
 - width, 187, 224
- dynamic Bayesian network, *see also* DBN, 98
- edge
 - contraction, 208
 - deletion, 255
 - fill-in, 138, 204
- effective treewidth, 144
- elimination
 - and triangulation, 230
 - heuristics, 205
 - of a factor, 153
 - of a node, 204
 - prefix, 206
 - tree, 155
- elimination order, 204
 - and triangulated graphs, 230
 - constrained, 258
 - from dtree, 216
 - from jointree, 214
 - optimal, 205, 209
 - perfect, 205, 230
 - width, 134
- EM, 446
 - for MAP estimates, 488, 496
 - for ML estimates, 450
- emission probabilities, 58
- empirical distribution, 442
- entropy, 30, 33, 346, 349, 444, 458, 517
 - relative, 518
- equivalence, 17
- equivalent sample size, 489
- error
 - absolute, 386
 - relative, 387
 - standard, 385
- estimate
 - consistent, 384
 - unbiased, 384
- estimating
 - a conditional probability, 392
 - an expectation, 383
- event, 16
- evidence
 - and local structure, 317, 322
 - entering into a jointree, 167
 - hard, 39
 - indicator, 148, 163, 166, 342
 - probability of, 139, 141
 - soft, 39, 80, 104, 150, 354, 434, 435
 - variable, 77, 84, 143
 - virtual, 46, 49, 81
- exhaustiveness, 17
- expectation, 381
- expectation maximization, *see also* EM, 446
- expected
 - count, 448, 487
 - log-likelihood function, 451
 - value, 381
- exponential distribution, 101
- extended factor, 247, 248
- factor, 129
 - division, 171, 176
 - elimination, 153
 - extended, 247
 - extended operations, 248
 - from tabular to ADD representations, 331
 - maximization operation, 246
 - multiplying, 131
 - reducing, 139
 - represented by ADDs, 327
 - summing out, 129
 - trivial, 129
- false negative, 38, 44, 80, 87
- false positive, 38, 44, 80, 87
- feedback loops, 105
- fill-in edges, 138, 204
- filled-in graph, 204
- first-order languages, 12
- Forward algorithm, 249
- free energy, 363
- function blocks, 93
- functional
 - dependencies, 316
 - networks, 316
- functionally determined variable, 118
- fuzzy logic, 5

Gamma function, 490
 Gaussian
 distribution, *see also* Normal distribution, 46
 noise, 104
 GBP, *see* generalized belief propagation, 349
 generalized belief propagation, 349
 generative model, 441
 genes, 110
 genetic linkage, 109
 genotype, 110
 Gibbs sampling, 403
 global parameter independence, 482
 gradient ascent algorithm
 for estimating network parameters, 451
 graph
 admitting a perfect elimination order, 230
 chordal, 230
 complete, 203
 contraction, 208
 contraction degeneracy, 209
 decomposable, 230
 degeneracy, 208
 degree, 208
 filled-in, 204
 interaction, 135, 202
 maximum minimum degree, 208
 minor, 209
 MMD, 208
 moral, 203
 sequence, 204
 triangulated, 230
 graphical models, xiii
 graphoid axioms, 58, 62
 greedy search
 for cache allocation, 195
 for learning network structure, 463
 haplotype, 110
 hard evidence, 39
 heuristic
 elimination, 205
 min-degree, 137, 205
 min-fill, 138, 205
 Hidden Markov Model, *see also* HMM, 55
 hidden variable, 445
 HMM, 55, 58, 249, 396
 emission probabilities, 58
 Forward algorithm, 249
 most probable state path, 249
 sensor probabilities, 58
 sequence probability, 249
 transition probabilities, 58
 Veterbi algorithm, 249
 Hoeffding's inequality, 386
 Hugin architecture, 166, 170, 308
 hypergraph partitioning, 227
 I-MAP, 68, 69
 IBP, *see also* belief propagation, 343

IJGP, *see also* iterative joingraph propagation, 352
 implication, 17
 importance
 distribution, 393
 sampling, 393
 independence, 34
 context-specific, 313, 316, 324
 contraction axiom, 61
 d-separation, 63, 65, 86, 94, 223, 364, 454, 506
 decomposition axiom, 59
 graphoid axioms, 62
 intersection axiom, 62
 of network parameters, 481
 positive graphoid axioms, 62
 symmetry axiom, 59
 triviality axioms, 62
 using DAGs, 53
 variable-based, 36, 59
 weak union axiom, 61
 independence MAP, *see also* I-MAP, 68
 indicator
 clause, 319
 variable, 319
 inequality
 Chebyshev's, 382
 Hoeffding's, 386
 information
 bits, 102
 gain, 518
 instantiation, 20
 interaction graph, 135, 202
 intermediary variable, 84
 intermittent faults, 97
 intersection axiom, 62
 inward phase of jointree algorithm, 162, 167
 IP clause, 278, 319
 irreducible Markov chain, 402
 iterative belief propagation, *see also* belief propagation, 343
 iterative joingraph propagation, 352
 Jeffrey's rule, 40
 joingraph, 350
 joint marginal, 138, 140
 jointree, 164, 213, 223
 algorithm, 164
 collect phase, 162, 167
 distribute phase, 162, 167
 Hugin architecture, 166, 170, 308
 inward phase, 162, 167
 outward phase, 162, 167
 pull phase, 162, 167
 push phase, 162, 167
 Shenoy-Shafer architecture, 166, 167, 217, 223, 308
 balanced, 229
 binary, 169, 222
 for DAG, 216
 for graph, 217, 234
 from dtree, 222
 from elimination order, 220
 minimal, 165, 166, 217

- property, 164, 216
- transforming, 217, 262
- width, 164, 216
- junction tree, *see also* jointree, 216
- K2 score, 463, 500
- K3 score, 463
- KL divergence, 346, 443, 470, 471, 473, 511, 518
- knowledge base, 14
- knowledge-based systems, 1
- Kullback-Leibler divergence, *see* KL divergence, 518
- Laplace approximation, 497
- large-sample approximation, 496
- latent variable, 445
- law of large numbers, 384
- law of total probability, 37
- learning Bayesian networks
 - the Bayesian approach, 477
 - the constraint-based approach, 440, 466
 - the likelihood approach, 439
- lifetime distribution, 101
- likelihood
 - approach for learning Bayesian networks, 439
 - equivalence, 501
 - function, 443
 - of a structure, 444
 - of parameters, 443
 - principle, 440
 - weighting, 395
- likelihood ratio, 51
- literal, 21
- local parameter independence, 482
- local search
 - for computing MAP, 263
 - for estimating network parameters, 446, 451, 486, 495
 - for learning network structure, 462
- local structure, 291, 300, 313
 - and arithmetic circuits, 319, 334
 - and CNF encodings, 319
 - and conditioning, 323
 - and determinism, 316, 325
 - and evidence, 317, 322
 - and variable elimination, 326
 - exposing, 318
- log-likelihood
 - of parameters, 444
 - of structure, 444
- logical
 - connectives, 14
 - constraints, 100, 316
 - properties, 16
 - relationships, 17
- loop-cutset, 180
- loopy belief propagation, *see also* belief propagation, 344
- lower bound
 - on treewidth, 208
- MAP, 82, 96, 100
 - computing using depth-first search, 260
 - computing using local search, 263
 - computing using variable elimination, 258
 - instantiation, 258
 - probability, 258
 - upper bound, 261
 - variables, 83
- MAP parameter estimates, 478, 485–487, 495
- MAR, 454
- marginal
 - distribution, 78
 - joint, 138, 140
 - likelihood, 485, 486
 - large-sample approximation, 496
 - stochastic sampling approximation, 496
 - posterior, 78, 138, 141
 - prior, 78, 133
- marginalizing a variable, 130
- Markov
 - blanket, 71
 - boundary, 71
- Markov chain, 402
 - aperiodic, 402, 410
 - converge, 411
 - detailed balance, 402
 - homogeneous, 402
 - irreducible, 402
 - recurrent states, 402
 - reversible, 402
 - simulation, 401
 - states, 402
- Markov Chain Monte Carlo, *see also* MCMC, 401
- Markov networks, xiii, 12
- Markovian assumptions, 55
- maximal clique, 203
- maximizer circuit, 296
- maximizing out, 246, 248
- maximum a posteriori, *see also* MAP, 243
- maximum likelihood (ML) parameter estimates, 443, 450
- maximum minimum degree, 208
- MCMC, 401
- MDL score, 461
- mean, 381
- message schedule, 345
- micro models, 115
- min-degree heuristic, 137, 205
- min-fill heuristic, 138, 205
- minimal I-MAP, 69
- minimal jointree, 165, 166, 217
- Minimum Description Length, *see also* MDL, 461
- minimum-cardinality models, 26
- missing data at random, 454
- missing-data mechanism, 453
- MLF, 290, 305, 339
- MMD, 208
- model, 16
 - complexity, 460
 - counting, 277, 319
 - dimension, 460
 - from design, 92, 98
 - from expert, 85, 87
 - persistence, 98
 - selection, 499

- model-based reasoning, 2
- monitors, 78
- monotonicity, 18
- monotonicity problem, 2
- Monte Carlo simulation, 383
- moral graph, 203
- moralization, 203
- most probable explanation, *see also* MPE, 244
- MPE, 81
 - computing using circuit propagation, 296
 - computing using depth-first search, 250
 - computing using variable elimination, 245
 - instantiation, 245
 - probability, 244
 - reducing to W-MAXSAT, 257, 280
 - sensitivity analysis, 424
 - upper bound, 255
- multi-linear function, *see also* MLF, 290
- multivalued variables, 19
- multiply-connected network, 108, 143
- multiplying factors, 131
- mutual exclusiveness, 17
- mutual information, 36, 360, 363, 368, 456, 457, 517

- naive Bayes, 86, 91, 149, 265, 468, 510
 - classifier, 467, 469
- negation normal form, *see also* NNF, 22
- network, *see* Bayesian network, 57
- NNF circuit, 22, 304–306, 320
 - decomposable, 22, 309
 - deterministic, 22, 309
 - smooth, 22, 309
- node
 - almost simplicial, 207
 - simplicial, 207
- noise model, 104
- noisy-or model, 115, 150, 318, 339
- non-descendants, 55
- non-informative prior, 494
- non-monotonic logic, xiii, 3
- Normal distribution, 46, 104, 382, 384, 392, 394, 443

- Occam's razor, 460
- odds, 41, 51, 120, 419
- optimal elimination order, 205, 209
 - best-first search, 210
 - depth-first search, 209
- outward phase of jointree algorithm, 162, 167
- overfitting, 459

- P-MAP, 69
- parameter
 - clause, 319
 - co-varying, 422, 479
 - equality, 321
 - estimates
 - Bayesian, 483, 492
 - MAP, 478, 485–487
 - maximum likelihood (ML), 443, 450
 - independence, 481
 - modularity, 501
 - set, 479, 489
 - variable, 319
- parametric structure, *see also* local structure, 313
- parents, 54
- partial derivatives
 - of maximizer polynomials, 299
 - of network polynomials, 292
- particle filtering, 396
- PDF, 46, 101
- pedigree, 110
- perfect elimination order, 205, 230
 - and triangulated graphs, 230
- perfect MAP, *see also* P-MAP, 69
- persistence model, 98
- phenotype, 110
- PI clause, 278, 319
- polytree
 - algorithm, 7, 163, 175, 180, 340
 - network, 108, 142
- positive graphoid axioms, 62
- possible world, 15
- posterior marginal, 78, 138, 141
- potential, *see also* factor, 129
- preprocessing rules, 206
- prior
 - equivalence, 501
 - knowledge, 480
 - marginal, 78, 133
- probability, 27
 - as an expectation, 383
 - density function, *see also* PDF, 46
 - distribution, 27
 - jointree factorization, 223
 - of evidence, 76, 139, 141
- projecting a factor, 130
- proposal distribution, 393
- propositional sentence
 - consistent, 17
 - satisfiable, 17
 - semantics, 15
 - syntax, 13
 - valid, 17
- pruning
 - edges, 144
 - network, 144
 - nodes, 143
- pull phase of jointree algorithm, 162, 167
- push phase of jointree algorithm, 162, 167

- qualification problem, 3
- query variable, 84, 143

- random
 - restarts, 264, 462
 - sample, 378
 - variable, 381
- Rao-Blackwell sampling, 391
- recombination
 - event, 111
 - fraction or frequency, 111
- recursive conditioning algorithm, 181

- reducing factor, 139
- redundant bits, 102
- refutation theorem, 25
- rejection sampling, 392, 404, 409
- relational languages, 12
- relative
 - entropy, 518
 - error, 387
- reliability
 - analysis, 98
 - block diagram, 98
- reversible Markov chain, 402
- running intersection property, 218, 219

- SamIam, 76–78, 81, 89, 90
- sample
 - mean, 383
 - variance, 385
- sampling
 - direct, 383, 385
 - distribution, 383
 - Gibbs, 403
 - importance, 393
 - likelihood weighting, 395
 - particle filtering, 396
 - Rao-Blackwell, 391
 - rejection, 392, 404, 409
 - sequential importance, 396
- satisfiability, 17
- satisfiable sentence, 17
- score equivalence, 501
- scoring measure
 - AIC, 460
 - Bayesian, 499
 - BD, 500
 - BDe, 501
 - BDeu, 502
 - BIC, 461
 - K2, 463, 500
 - K3, 463
 - MDL, 461
- selective model averaging, 499
- sensitivity analysis, 89, 100, 119, 417
 - network independent, 418
 - network specific, 423
 - of most probable explanations, 424
- sensitivity function, 423
- sensor probabilities, 58
- separator, 157, 164, 216
- sequential importance sampling, 396
- Shenoy-Shafer architecture, 166, 167, 217, 223, 308
- simplicial node, 207
- simulating
 - a Bayesian network, 379
 - a Gibbs chain, 404
 - a Markov chain, 401
- simulation
 - Markov chain, 401
 - Monte Carlo, 383
- single-fault assumption, 86
- singly-connected network, 108, 142

- smoothness, 22, 25, 304, 306, 320
- soft evidence, 39, 46, 80, 104, 150, 354, 434, 435
 - and continuous variables, 46
 - as a noisy sensor, 44
- soft-max parameter space, 452
- splitting variables, 255
- standard
 - deviation, 382
 - error, 385
- standard Normal distribution, 46, 47, 382
- state of belief, 27
- stationary distribution, 402
- stochastic sampling, 378
- strictly positive distribution, 62
- structure
 - local, 291, 300, 313
 - parametric, 313
- structure-based algorithms, 143, 313
- sufficient statistic, 442
- summing out, 129
- symmetry axiom, 59
- systematic search
 - for cache allocation, 193
 - for computing MAP, 260
 - for computing MPE, 250
 - for computing optimal elimination orders, 209
 - for learning network structure, 464

- TAN classifier, 467
- term, 21
- transition probabilities, 58
- tree
 - contraction, 228
 - decomposition, 231, 234
- treewidth, 141, 142, 165
 - complexity, 231
 - constrained, 259
 - dtree definition, 227
 - effective, 144
 - elimination order definition, 205
 - jointree definition, 222
 - lower bound, 208, 209
 - of DAG, 203
 - of graph, 205
- triangulated graph, 230
- trivial
 - factor, 129
 - instantiation, 21
- triviality axiom, 62
- truth assignment, 15
- turbo codes, 107

- unbiased estimate, 384
- unit
 - clause, 21
 - resolution, 23, 326
- upper bound
 - for MAP, 261
 - for MPE, 255

- valid sentence, 17
- validity, 17

- variable
 - assignment, 15
 - binary, 14
 - Boolean, 14
 - bypass, 91, 95, 116
 - clone, 255
 - continuous, 12, 46, 101, 105, 491
 - deterministic, 78
 - evidence, 84, 143
 - functionally determined, 118
 - hidden, 445
 - instantiation, 20
 - intermediary, 84
 - latent, 445
 - maximizing out, 246, 248
 - multi-valued, 19
 - propositional, 14
 - query, 84, 143
 - splitting, 255
 - variable elimination
 - for compiling networks, 334
 - for inference, 126
 - using ADDs, 332
 - with local structure, 326
 - variable-based independence, 36
 - variance, 381
 - Viterbi algorithm, 249
 - virtual evidence, 46, 49, 81
 - weak union axiom, 61
 - weighted CNF, 280
 - weighted model counting, 277
 - with local structure, 319
 - WER, 104
 - width
 - of a-cutset, 184
 - of context, 186
 - of cutset, 184
 - of dtree, 187, 224
 - of elimination order, 134, 205
 - of elimination prefix, 206
 - of elimination tree, 158
 - of jointree, 164, 216
 - word error rate, 104
 - world, 15
 - zero counts, 443