

Patrick Naïm, Pierre-Henri Willemin,
Philippe Leray, Olivier Pourret, Anna Becker

Réseaux bayésiens

- Introduction intuitive aux réseaux bayésiens
- Fondements théoriques et algorithmes
- Méthodologie de mise en œuvre
 - Domaines d'application et études de cas détaillées
- Outils logiciels : Bayes Net Toolbox, BayesiaLab, Hugin, Netica et Elvira

Réseaux bayésiens

3^e édition

**Patrick Naïm, Pierre-Henri Wuillemin,
Philippe Leray, Olivier Pourret, Anna Becker**

Avec la contribution de Bruce G. Marcot,
Carmen Lacave et Francisco J. Díez

EYROLLES

Table des matières

Première partie : introduction aux réseaux bayésiens	1
1 Approche intuitive	3
1.1 Une représentation graphique de la causalité	3
1.1.1 Circulation de l'information dans un graphe causal	4
▶ Un exemple	5
▶ Le cas général	6
1.1.2 D-séparation (blocage)	7
1.1.3 Conclusion	10
1.2 Une représentation probabiliste associée	11
1.2.1 Transposition	11
▶ Définition des probabilités	12
1.2.2 Premier exemple : validité de la formalisation probabiliste	12
▶ Modélisation	12
▶ Discussion	14
1.2.3 Deuxième exemple : dépendances et indépendances	17
▶ Modélisation	17
▶ Utilisation du modèle	18
▶ Circuits d'information et indépendances	20
1.2.4 Les réseaux bayésiens	21
▶ Définition	21
▶ Propriétés	22
▶ Utilisation et difficultés	22

2	Introduction aux algorithmes	25
2.1	Inférence	25
2.1.1	Approche intuitive	26
	▶ Chaînes	27
	▶ Arbres	28
	▶ Polyarbres	28
2.1.2	Cas général	28
	▶ Conditionnement	30
	▶ Arbre de jonction	31
	▶ Méthodes approchées	32
2.2	Apprentissage	32
2.2.1	Apprentissage de paramètres	34
	▶ Cas général	35
2.2.2	Apprentissage de structure	36
	▶ Critère	36
	▶ Recherche : structures contraintes	37
	▶ Recherche : le cas général	38
2.3	Modèles continus	38
2.4	Liens avec d'autres méthodes	39
3	Exercices	41
3.1	Pour commencer	42
3.1.1	Daltonisme	42
3.1.2	Langues orientales	42
3.1.3	Détection d'une maladie animale	42
3.1.4	Provenance d'un composant	43
3.2	Grands classiques	43
3.2.1	Jet de deux dés	43
3.2.2	Trois coffres	43
3.2.3	Trois prisonniers	43
3.2.4	Meurtres en Floride	44

3.3	Cas tests	44
3.3.1	Diagnostics médicaux contradictoires	44
3.3.2	Contrôles antidopage	45
3.3.3	Fiabilité d'un système	45
3.3.4	Détection d'incendie	46
3.3.5	Au tribunal	46
3.3.6	Gestion d'un parc de véhicules	46
3.4	Plus difficiles	47
3.4.1	Âges	47
3.4.2	Décision de justice	47
3.4.3	Modèle génétique	47
3.4.4	Contrôle d'un procédé	48
3.4.5	Jeu télévisé	48
3.4.6	Mesure de température	49
3.4.7	Durée de vie d'une ampoule électrique	49
3.4.8	Mesure d'une superficie	50
3.4.9	Réseau électrique	50
3.5	Aide à la décision	50
3.5.1	Dilemme... cornélien	50
3.5.2	Parapluie	51
3.5.3	Tournoi de tennis	51
3.6	Exercices théoriques	52
3.6.1	Pouvoir de modélisation des réseaux bayésiens	52
3.6.2	Apprentissage de probabilités	53
3.6.3	Indépendances 2 à 2	53
3.7	Commentaires et solutions des exercices	53

Deuxième partie : cadre théorique et algorithmes 71

4 Modèles graphiques et indépendances 73

4.1 Graphoïdes 74

4.1.1	Modèles d'indépendance	74
4.1.2	Semi-graphoïde et graphoïde	74
4.2	Modèle d'indépendance et loi de probabilité	76
4.3	Modèles d'indépendance et séparation dans les graphes . .	78
4.4	Modèles non orientés : réseaux de Markov	80
4.4.1	Définition	80
4.4.2	Séparation et indépendances : propriétés de Markov	81
4.4.3	Réseaux de Markov et factorisation	82
4.4.4	Limites	83
4.5	Modèles orientés : réseaux bayésiens	84
4.5.1	Définitions	84
4.5.2	Propriétés de Markov dans les graphes orientés . .	86
4.5.3	Réseaux bayésiens et factorisation	87
4.5.4	Limites	89
4.6	Pourquoi des arcs plutôt que des arêtes ?	90
4.6.1	Factorisation	90
4.6.2	Sémantique et causalité	91
4.6.3	Pragmatisme	92
5	Propagations	93
5.1	Propagation par messages locaux dans un arbre	94
5.1.1	Décomposition de l'information	95
	► Calcul de $\lambda(x)$	97
	► Calcul des $\pi(x)$	98
	► Synthèse et écriture itérative	99
	► Cas des racines, des feuilles et des nœuds informés	102
5.1.2	Algorithme de propagation : <i>polytree propagation</i> . .	103
5.2	Conditionnement global	104
5.2.1	Principe de la coupe	104
5.2.2	Propagation conditionnée	105
5.3	Arbre de jonction	106

5.3.1	Moralisation et Triangulation	106
5.3.2	Propagation dans l'arbre de jonction	110
5.4	Méthodes approchées	111
5.4.1	Méthodes exactes sur des topologies approchées . .	111
5.4.2	Méthodes stochastiques	112
	▶ Connaissance parfaite de la loi à simuler	113
	▶ Connaissance imparfaite de la loi à simuler	113
6	Apprentissage	117
6.1	Apprentissage des paramètres	118
6.1.1	À partir de données complètes	118
	▶ Apprentissage statistique	118
	▶ Apprentissage bayésien	120
6.1.2	À partir de données incomplètes	121
	▶ Nature des données manquantes	121
	▶ Traitement des données MCAR	122
	▶ Traitement des données MAR	122
	▶ Apprentissage statistique et algorithme EM	122
	▶ Apprentissage bayésien et algorithme EM	125
6.1.3	Incorporation de connaissances	126
	▶ Comment demander à un expert d'estimer une probabilité ?	126
	▶ Quelles probabilités estimer ?	127
	▶ Comment fusionner les avis de plusieurs experts ?	130
6.2	Apprentissage de la structure	131
6.2.1	Introduction	131
6.2.2	Hypothèses	132
6.2.3	Notion d'équivalence de Markov	133
6.2.4	Recherche d'indépendances conditionnelles	136
	▶ Tests d'indépendance conditionnelle	137
	▶ Algorithmes PC et IC	139

	► Quelques améliorations	144
6.2.5	Algorithmes basés sur un score	144
	► Les scores possibles	145
	► Déterminer un <i>a priori</i> sur les structures	148
	► Pourquoi chercher la meilleure structure ?	149
	► Recherche dans l'espace des réseaux bayésiens	149
	► Algorithmes basés sur un score et données incomplètes	158
	► Recherche dans l'espace des classes d'équivalence de Markov	161
6.2.6	Méthodes hybrides	170
6.2.7	Incorporation de connaissances	171
	► Structures de réseaux bayésiens pour la classification	172
	► Structures de réseaux bayésiens avec variables latentes	175
	► Autres structures particulières	176
6.2.8	Découverte de variables latentes	177
	► Recherche d'indépendances conditionnelles	177
	► Algorithmes basés sur un score	179
6.2.9	Cas particulier des réseaux bayésiens causaux	179
	► Définition	180
	► Apprentissage sans variables latentes	180
	► Apprentissage avec variables latentes	182

Troisième partie : méthodologie de mise en œuvre et études de cas 185

7	Mise en œuvre des réseaux bayésiens	187
7.1	Pourquoi utiliser des réseaux bayésiens ?	187
7.1.1	Acquisition des connaissances	188
	► Un recueil d'expertise facilité	188

▶	Un ensemble complet de méthodes d'apprentissage	189
▶	Un apprentissage incrémental	189
7.1.2	Représentation des connaissances	191
▶	Un formalisme unificateur	191
▶	Une représentation des connaissances lisible	192
7.1.3	Utilisation de connaissances	192
▶	Une gamme de requêtes très complète	192
▶	Optimisation d'une fonction d'utilité	193
7.1.4	Limites des réseaux bayésiens	195
▶	Un recul encore insuffisant pour l'apprentissage	195
▶	Utilisation des probabilités	195
▶	Lisibilité des graphes	196
▶	Les variables continues	196
▶	La complexité des algorithmes	196
7.1.5	Comparaison avec d'autres techniques	197
7.2	Où utiliser des réseaux bayésiens ?	197
7.2.1	Caractéristiques générales	198
▶	Une connaissance explicite ou implicite du domaine	198
▶	Une utilisation complexe ou évolutive	199
7.2.2	Classification des applications par types	200
▶	Modèles symboliques	200
▶	Modèles numériques	201
7.2.3	Classification des applications par domaines	201
▶	Santé	201
▶	Industrie	202
▶	Défense	203
▶	Banque/finance	203
▶	Marketing	206
▶	Informatique	207

►	Gestion des connaissances	208
7.3	Comment utiliser des réseaux bayésiens ?	208
7.3.1	Identification des variables et de leurs espaces d'états	209
7.3.2	Définition de la structure du réseau bayésien	210
7.3.3	Loi de probabilité conjointe des variables	211
8	Exemples d'applications	213
8.1	Détection de fraude (ATT)	213
8.2	Aide à la décision en temps réel (NASA)	216
8.3	Autres applications (par domaines)	219
8.3.1	Industrie	219
8.3.2	Santé	224
8.3.3	Informatique et télécommunications	225
8.3.4	Défense	227
9	Étude de cas n°1 : gestion globale des risques d'une entreprise	231
9.1	La méthode GLORIA	232
9.2	Horizon de temps et objectifs de l'entreprise	233
9.3	Construction du réseau bayésien	234
9.3.1	Identification des variables	234
9.3.2	Identification des relations entre variables	236
9.4	Lois de probabilité des variables	237
9.4.1	Variables sommets	237
9.4.2	Variables intermédiaires	237
9.4.3	Exemple	238
9.5	Résultats de la méthode GLORIA	238
9.5.1	Probabilité de non-atteinte des objectifs	238
9.5.2	Simulation	238
9.5.3	Diagramme probabilité/gravité	240
9.5.4	Criticité des risques	242
10	Étude de cas n°2 : modélisation et quantification des risques opé-	

rationnels	245
10.1 Gestion des risques, incertitude et connaissance	246
10.2 Présentation de la démarche	247
10.3 Modélisation des scénarios de risque	249
10.3.1 Présentation de la méthode	249
► Objectifs	249
► Difficultés	250
► Connaissance ou données	251
► Un processus de gestion des connaissances	252
10.3.2 Le modèle Exposition - Survenance - Gravité (XSG)	253
► La vulnérabilité	254
► Exposition - Survenance - Gravité	254
► Utilisation des réseaux bayésiens	255
► Avantages de l'utilisation des réseaux bayésiens	256
10.3.3 Définition des scénarios	257
► Identification des vulnérabilités	257
► Sélection des vulnérabilités	258
► Étude détaillée des vulnérabilités	258
10.3.4 Quantification des scénarios	258
► Définir l'exposition, la survenance, et la gravité	259
► Modéliser l'exposition	259
► Modéliser la survenance	260
► Modéliser la gravité	262
10.3.5 Résumé	266
10.4 Conclusion	267
11 Étude de cas n°3 : étude d'un système électrique	269
11.1 Modélisation d'un réseau électrique	270
11.1.1 Variables aléatoires	270
11.1.2 Dépendances entre variables	271
11.1.3 Choix d'un modèle mathématique	273

11.2	Étude du réseau électrique en région PACA	274
11.2.1	Contexte	274
11.2.2	Construction du modèle	275
11.2.3	Résultats de l'étude	276
12	Étude de cas n°4 : questionnaire adaptatif pour la vente de crédit en ligne	279
12.1	Un réseau bayésien comme modèle de score	280
12.1.1	Données et prétraitement	281
12.1.2	Modélisation	282
12.1.3	Le modèle obtenu	284
12.2	Utilisation du réseau bayésien	285
12.3	Résultats et conclusion	289
13	Étude de cas n°5 : gestion de ressources naturelles et analyses de risques	293
13.1	Revue des méthodes	294
13.1.1	Pourquoi les réseaux bayésiens ?	295
13.1.2	Méthodes de création de réseaux bayésiens	295
	▶ Utilisation de diagrammes d'influence	295
	▶ Probabilités associées aux variables	296
	▶ Construction de réseau bayésien à partir d'expertise ou de données	296
	▶ Utilisation de variables <i>proxy</i>	298
13.2	Exemples de réseaux bayésiens	299
13.2.1	Modèles de prévision pour la faune et la flore	299
	▶ Modélisation des musaraignes pygmées dans le bassin intérieur de la Colombie Britannique (États-Unis)	299
	▶ Modélisation de la grouse cendrée dans le bassin intérieur de la Columbia (États-Unis)	301
	▶ Faune et flore du Nord-Ouest Pacifique des États-Unis	302

► Faune et Flore de l'Ouest du Canada	304
13.2.2 Utilisation de réseaux bayésiens pour la rétrovision	304
13.2.3 Les réseaux bayésiens comme modèles de décision	307
13.3 Utilisation des réseaux bayésiens pour étudier la faune et la flore et gérer les ressources naturelles	311
13.3.1 Couplages avec d'autres modèles	311
13.3.2 Gestion adaptative	311
13.3.3 Prise de décision en univers incertain et considéra- tion des types d'erreurs	313
13.3.4 Mise à jour et affinage des modèles	313
13.4 Conclusion et perspectives	314
14 Étude de cas n°6 : diagnostic médical	317
14.1 Sources d'incertitudes en médecine	318
14.2 Construction de réseaux bayésiens médicaux	320
14.2.1 Construction de réseaux bayésiens à partir de bases de données médicales	321
14.2.2 Construction à l'aide d'experts humains	322
► Construction du graphe causal	322
► Application de modèles canoniques	324
► Acquisition d'informations quantitatives	325
14.3 Un exemple de modèle : PROSTANET	326
14.3.1 Structure du graphe	327
14.3.2 Recueil de probabilités	328
14.3.3 Déboguage	331
14.3.4 Évaluation	332
14.3.5 Historique des versions	332
14.4 Conclusion	332
Annexes	335
A Théorie des graphes	337

A.1	Définitions générales	337
A.2	Notions orientées	340
A.3	Notions non orientées	341
A.4	Typologie et propriétés des graphes	343
B	Probabilités	347
B.1	Probabilités	347
B.1.1	Définitions principales	348
B.1.2	Probabilités sur plusieurs variables	349
	► Probabilités jointes	350
	► Probabilités marginales	351
	► Probabilités conditionnelles	352
B.2	Indépendance conditionnelle	354
B.2.1	Définitions	354
B.2.2	Propriétés	357
C	Outils	359
C.1	Bayes Net Toolbox (BNT)	359
C.1.1	Présentation	359
C.1.2	Modélisation	360
C.1.3	Apprentissage	360
C.1.4	Inférence	361
C.2	BayesiaLab	361
C.2.1	Présentation	361
C.2.2	Modélisation	361
C.2.3	Apprentissage	363
C.2.4	Exploitation	365
C.2.5	Analyse	367
C.2.6	Prise en compte de la dimension temporelle	369
C.2.7	Aide à la décision	369
C.2.8	Compléments	370

C.2.9	Conclusion	370
C.3	Hugin	371
C.3.1	Présentation	371
C.3.2	Construction des modèles	371
C.3.3	Inférence	374
C.3.4	Apprentissage	376
C.3.5	Compléments	377
C.3.6	Conclusion	377
C.4	Netica	378
C.4.1	Présentation	378
C.4.2	Construction des modèles	378
C.4.3	Inférence	379
C.4.4	Apprentissage	380
C.4.5	Autres fonctionnalités	380
C.4.6	Conclusion	382
C.5	Elvira	382
C.5.1	Introduction	382
C.5.2	Le format Elvira	384
C.5.3	Interface graphique	384
C.5.4	Principales fonctionnalités	385
Bibliographie		389
Liste des figures		411
Liste des tables		417
Index		421