

TABLES DE MATIERES

Introduction générale.....	1
CHAPITRE 1 : Contexte du diagnostic à base de modèles	4
1. Introduction.....	5
2. Définition.....	5
3. Méthodes de diagnostic.....	6
3.1. Diagnostic à base de connaissances.....	6
3.1.1. Diagnostic par systèmes experts.....	6
3.1.2. Systèmes experts flous	8
3.2. Diagnostic à base de modèles	10
3.2.1. Diagnostic par Réseaux Bayésiens	11
4. Conclusion	12
CHAPITRE 2 : Logique Floue et Réseau Bayésien.....	13
1. Introduction.....	14
2. Logique Floue.....	14
2.1. Introduction	14
2.2. Définition.....	14
2.3. Historique	15
2.4. Application	15
2.5. Théorie des sous-ensembles flous.....	15
2.6. Opérateurs en logique floue.....	17
2.6.1. L'opérateur NON (complément)	17
2.6.2. L'opérateur ET (intersection)	17
2.6.3. L'opérateur OU (union).....	18
2.7. La commande floue	18
2.7.1. Fuzzification.....	19
2.7.2. Inférence.....	20
2.7.3. Défuzzification	22
2.8. Conclusion	22
3. Réseaux Bayésiens	23
3.1. Introduction	23
3.2. Définition.....	23
3.3. Historique	24
3.4. Le théorème de Bayes.....	24
3.5. Construction du réseau bayésien.....	24

3.6. Modélisation par réseau bayésien	25
3.6.1. Liens d'influences et graphes causaux	25
3.6.2. D-séparation et indépendance conditionnelle	26
3.7. Inférence dans les réseaux bayésiens	27
3.8. Apprentissage dans les réseaux bayésiens	27
3.9. Applications utilisant les réseaux bayésien	28
3.10. Exemple simple faisant intervenir des faits incertains	29
3.11. Comparatif de la technique réseau bayésien vis-à-vis d'autres techniques	32
3.12. Conclusion	34
4. Conclusion	34
CHAPITRE 3 : Présentation des Réseaux Bayésiens Flous : Fondement et outil	35
1. Introduction	36
2. Définition	37
3. La théorie de Réseau Bayésien Flou	37
3.1. Première approche de Réseau Bayésien Flou	38
3.1.1. Probabilité Flou	38
3.1.2. Equation Bayésienne Floue	38
3.2. Deuxième approche de Réseau Bayésien Flou	40
3.2.1. Structure du Réseau Bayésien Flou	40
3.2.2. Propagation de croyance dans le Réseau Bayésien Flou	40
3.2.2.1. Quelques notions de base	40
3.2.2.2. Quelques suppositions	40
3.2.2.3. Propagation de croyance dans le cas d'un seul parent	41
3.2.2.4. Propagation de croyance dans le cas des parents multiples	42
4. Avantage des Réseaux Bayésiens Flous	43
5. Recherches effectuées dans les réseaux bayésiens flous	44
6. Notre approche pour l'inférence dans les Réseaux bayésiens flous	44
6.1. Présentation de notre algorithme	44
6.1.1. Arbre de jonction	44
6.1.2. Soft evidence	45
6.1.3. Algorithme proposé	47
7. Conclusion	48
CHAPITRE 4 : Réalisation et test	49
1. Introduction	50
2. Editeur Réseau Bayésien Flou	50
2.1. Editeur Réseau Bayésien classique	50
2.1.1. Aspect graphique	50

2.1.2. Algorithme d'inférence	51
2.2. Représentation des données floues	53
3. Exemples d'inférence flou	54
3.1. Exemple simple	54
3.2. Exemple de diagnostic du cancer de la prostate.....	55
3.2.1. Cancer de la prostate	55
3.2.2. Problématique et solution	57
3.2.3. Réseau Bayésien Flou pour le diagnostic du cancer de la prostate.....	58
4. Conclusion	62
Conclusion générale	63
Bibliographie.....	66

LISTE DES FIGURES

Figure 1 : Principe d'un système expert Flou [Ketata et al., 2005]	9
Figure 2 : Représentation des fonctions $\mu_A(x)$ et non ($\mu_A(x)$).....	17
Figure 3 : Illustration de la fonction intersection.....	18
Figure 4 : Illustration de la fonction union	18
Figure 5 : Schéma d'un contrôleur flou	19
Figure 6 : Illustration des sous-ensembles des trois variables Distance, Vitesse et Freinage	20
Figure 7 : Illustration de l'inférence en appliquant la méthode Max-Min	21
Figure 8 : Graphe causal entre A et B.....	25
Figure 9 : Graphe du réseau bayésien.....	30
Figure 10 : Exemple de RBF.....	41
Figure 11 : Exemple pour Mise à jour de Croyance avec soft evidence.....	47
Figure 12 : Présentation générale de l'éditeur	50
Figure 13 : Graphe initial.....	51
Figure 14 : Réseau après choix de l'algorithme.....	52
Figure 15 : Résultat final	53
Figure 16 : Représentation de la variable floue Age	54
Figure 17 : Résultat final	55
Figure 18 : Sous-ensembles flous de la variable Taux de l'APS	58
Figure 19 : RBF pour le diagnostic du cancer de la prostate	59
Figure 20 : Sous-ensembles flous de la variable Taux de l'APS.....	61
Figure 21: Degré d'appartenance aux sous-ensembles flous de la variable Taux de l'APS.....	62

LISTE DES TABLEAUX

Tableau 1. <i>Fonctions d'appartenances entre la théorie classique et la théorie floue</i>	16
Tableau 2. <i>Table de probabilité (Retard Fantasio)</i>	31
Tableau 3. <i>Table de probabilité (Retard Gaston)</i>	31
Tableau 4. <i>Comparatif des réseaux bayésiens à d'autres techniques [Naïm et al., 2004]</i>	32
Tableau 5. <i>Comparaison entre la théorie de probabilité et la logique floue</i>	36
Tableau 6. <i>Résultat du recensement de la population</i>	38
Tableau 7. <i>Données du bureau fiscal</i>	39
Tableau 8. <i>Distribution de probabilités à priori de A</i>	41
Tableau 9. <i>Distribution de probabilités conditionnelles de B</i>	41
Tableau 10. <i>Distribution de probabilités conditionnelles de C</i>	41
Tableau 11. <i>Distribution de probabilités conditionnelles de E</i>	42

Introduction générale

" Calculer la probabilité d'un événement n'a aucun sens une fois que l'on sait qu'il s'est produit. L'apparition de la vie, celle des dinosaures, celle des Hommes, a résulté d'un grand nombre de bifurcations dans le cours des processus se déroulant sur notre planète; chacune de ces bifurcations s'est produite alors que de nombreuses autres étaient possibles; chacune avait une probabilité faible, mais il fallait bien qu'une de ces possibilités se produise."

*La science à l'usage des non-scientifiques,
2003 Albert Jacquard, né en 1925.*

Nous assistons aujourd'hui à des révolutions dans les moyens de fabrication, de production, de traitement de l'information et de communication. Ces révolutions ont contribué à la production d'outils et systèmes de plus en plus complexes par leurs tailles et leurs modes de fonctionnement.

Ceci impose le recours à des méthodes de diagnostic de plus en plus sophistiquées tel que les méthodes de diagnostic à base de connaissance ou les méthodes plus récentes de diagnostic à base de modèle.

Ces derniers ont surmonté certaines limites inhérentes aux approches de diagnostic à base de connaissance en utilisant un modèle de système afin de prédire sa panne ou chercher son origine.

Parmi les modèles utilisés pour décrire la connaissance sur un système, on distingue les Réseaux Bayésiens (RBs) qui permettent la manipulation de connaissance incertaine.

Cependant, l'une des limites majeures que rencontrent les utilisateurs des réseaux bayésiens est d'exprimer les connaissances non précises surtout dans les diagnostics où beaucoup d'informations sont floues. En faite, nous ne pouvons pas exprimer toutes sortes de connaissances par un langage précis.

La solution est donc de combiner les réseaux bayésiens avec la logique floue, d'où la naissance des Réseaux Bayésiens Flous (RBFs).

L'objectif de ce projet de mastère est d'appréhender la problématique des Réseaux Bayésiens Flous, ensuite de modéliser quelques systèmes par des réseaux bayésiens flous et de développer des algorithmes d'inférence à des fins de diagnostic exploitant la richesse de ces modèles.

Dans ce projet nous cherchons à résoudre quelques problèmes constatés lors du diagnostic à base de Réseau Bayésien classique. Ces problèmes concernent essentiellement la modélisation qui s'avère difficile en présence de connaissances ambiguës. Les algorithmes d'inférence et de diagnostic sont non génériques et non adaptés au contexte flou.

Ce rapport est organisé selon quatre chapitres :

Le premier chapitre propose quelques techniques de diagnostic qu'on peut classer essentiellement, en deux catégories : diagnostic à base de connaissances et diagnostic à base

de modèles. Notre objectif dans ce chapitre est de présenter la méthode de diagnostic de modèle et comment utiliser un réseau bayésien comme modèle.

Le deuxième chapitre détaille les notions de base de la logique floue et du réseau bayésien, qui sont parmi les techniques les plus intéressantes pour l'aide au diagnostic. Ce modèle combine les avantages des réseaux bayésiens en termes de flexibilité et de facilité d'interprétation et les avantages de la logique floue en termes de flexibilité.

Le troisième chapitre met en évidence la théorie de réseau bayésien flou. Il valorise ses avantages, invoque les applications qui l'utilisent, et détaille les approches qui existent pour inférer ce genre de modèle.

Enfin, le quatrième chapitre aborde la réalisation et le test de notre application. Celle-ci se compose d'un éditeur graphique pour la modélisation de RBF. Ainsi, qu'un algorithme d'inférence permettant la mise à jour de données classiques ou floues.

CHAPITRE 1

Contexte du diagnostic à base de modèles

*"Il entre dans toutes les actions
humaines plus de hasard que de
décision." André Gide (1869-1951).*

1 Introduction

Le diagnostic est une activité présente dans de nombreux milieux : médical, administratif, industriel. Le terme diagnostic est issu du domaine médical où il signifie "Action de déterminer une maladie d'après ses symptômes". Le diagnostic technique peut être défini comme la recherche de l'origine des pannes ou défauts afin de réparer un système.

Généralement, le processus du diagnostic repose sur le principe de l'observation et de l'analyse des symptômes d'une défaillance, dans le but de trouver le remède garantissant le rétablissement et réduire la probabilité de défaillance. Les décisions prises suite à un diagnostic sont généralement le fruit d'une expérience, d'un savoir-faire ou d'un historique.

Dans ce premier chapitre, nous en proposons quelques techniques de diagnostic qu'on peut classer, essentiellement, en deux catégories : diagnostic à base de connaissances et diagnostic à base de modèles.

2 Définition

Le diagnostic est défini [1] comme étant le raisonnement menant à l'identification de la cause ou l'origine d'une défaillance, d'un problème ou d'une maladie, à partir des caractères ou symptômes relevés par des observations, des contrôles ou des tests.

Cependant, le diagnostic est un terme commun à plusieurs disciplines, sa définition est liée à la nature du système considéré et du résultat attendu. Le mot diagnostic n'a pas exactement la même définition suivant les disciplines et les auteurs.

Par exemple, Reiter a défini un diagnostic comme suit : « *Supposons connue une description d'un système ainsi qu'une observation du comportement anormal. Le problème du diagnostic consiste à déterminer les composants dont le dysfonctionnement expliquerait les différences observées* » [Reiter, 1986]. Il utilise des littéraux pour représenter le système et les liens entre ses composants, un diagnostic est défini comme une conjonction de littéraux qui sont consistants avec la description et le comportement observé et qui inclut un littéral pour chacun des représentants des composants.

En outre, Van der Gaag [Gaag et al., 1994] a défini la notion de diagnostic comme l'acheminement qui permet d'identifier l'ensemble le plus probable de désordres qui expliquent les manifestations observées en cas de problème.

3 Méthodes de diagnostic

Les outils de diagnostic mis à la disposition des spécialistes en maintenance et en supervision sont nombreux, mais leur efficacité est très inégale. Ils appartiennent généralement soit aux méthodes à base de connaissances soit aux méthodes plus récentes de diagnostic à base de modèles. Nous allons présenter, dans la suite, ces principales méthodes de diagnostic issues de travaux dans le domaine de l'intelligence artificielle.

3.1 Diagnostic à base de connaissances

Dans de nombreuses méthodes de diagnostic, la connaissance exacte du fonctionnement du système n'est pas indispensable, l'expérience et les tests sur site constituent la base pour les approches qui exploitent des connaissances de nature heuristique. Parmi les méthodes de diagnostic à base de connaissances nous distinguons principalement le diagnostic par systèmes experts.

3.1.1 Diagnostic par systèmes experts

Un système expert (SE) est un système d'aide à la décision basé sur un moteur d'inférence¹ et sur une base de connaissances [2]. Il est la transcription logicielle de la réflexion d'un expert dans un domaine donné. Il est capable de déduction logique et de produire une solution qui semble la plus juste.

Toutefois, il reste un outil d'aide à la décision et est loin de pouvoir remplacer l'intelligence d'un expert, d'ailleurs il n'est concevable que pour les domaines dans lesquels il existe des experts humains.

Les SEs ne sont en aucun cas des logiciels adaptatifs mais plutôt des applications dédiées à leur domaine d'activité, c'est pourquoi les SEs sont des progiciels² au sens pur. Ils sont généralement conçus pour résoudre des problèmes de classification ou de décision.

La mise en place d'un SE dans un milieu professionnel se fait en 5 étapes :

1. Etude de faisabilité

La mise en place d'un SE ne peut se faire que dans le cadre d'un domaine d'expertise dont les connaissances et le savoir-faire est formalisable. C'est à dire un domaine qui n'a pas trop

¹ Partie d'un système expert qui effectue la sélection et l'application des règles en vue de la résolution d'un problème donné.

² Contraction de produit et logiciel, est un logiciel applicatif commercial "prêt-à-porter", standardisé et générique, prévu pour répondre à des besoins ordinaires.

attirait à la sensibilité humaine. L'investissement en temps et donc en argent, pour la mise en place d'un SE, est énorme et nombreux sont les entreprises qui souhaiteraient s'en doter mais tous les domaines d'expertise ne sont pas formalisables.

Le premier travail du cogniticien³ est donc d'évaluer le domaine et les risques d'échecs de la mise en place et de succès de l'outil auprès des professionnels et futurs utilisateurs.

2. Extraction des données

Une fois l'assurance que cette mise en place est possible, la partie la plus importante de la mise en place va commencer. Il s'agit d'un dialogue entre le cogniticien et l'expert afin d'extraire de ce dernier toutes ses connaissances et son savoir-faire.

Un tel objectif est évidemment impossible à atteindre, mais le cogniticien va tenter de s'en approcher au maximum. Il devra pour cela faire preuve d'une grande compréhension des informations qui lui seront transmises (les experts n'étant pas forcément bons pédagogues) et d'un certain sens de la psychologie pour faire parler un expert qui aura tout naturellement le sentiment de se faire très prochainement remplacer par un système informatique.

3. Formalisation

Après et pendant l'extraction des données, le cogniticien devra formaliser les connaissances qu'il a glanées. Pour cette partie, il peut alors commencer à se tourner vers les développeurs et autres professionnels techniques de l'informatique afin de commencer à définir le cahier des charges précis, la base de connaissance et les règles d'inférence. À partir de cette étape, on a déjà un pied dans la technique.

4. Design et développement

Une fois la base de connaissances et les règles d'inférences définies, le cogniticien peut alors se retourner vers l'équipe technique qui va définir l'architecture technique nécessaire. Le cogniticien aura à partir de là le rôle de lien entre l'équipe d'experts et l'équipe de développement afin de peaufiner le cahier des charges et d'optimiser les métadonnées⁴ et métarègles⁵.

³ Personne ou ingénieur spécialisé dans l'extraction des connaissances.

⁴ Les métadonnées, ou données sur les données, renseignent sur la nature, les caractéristiques et la disponibilité des données. Elles rendent les données compréhensibles et partageables pour les utilisateurs dans le temps.

⁵ Règles contrôlant la sélection des règles à appliquer.

5. Tests et optimisations

Naturellement la mise en place se termine par une série de tests auprès des experts mais aussi auprès d'utilisateurs lambdas⁶ qui sont sensés à partir de cet outil fournir les résultats d'un expert débutant.

Avec le recul du temps et la confrontation à la réalité, l'utilisation des systèmes experts avait dégagé un ensemble de limitations [Tomsovic et al., 2000], dont la plus importante est celle du raisonnement incertain. En fait, un des plus grands problèmes que rencontre le cogniticien lorsqu'il tente de formaliser le savoir d'un expert, c'est que celui-ci est capable de raisonner sur des connaissances incertaines ou imprécises et qu'on ne dispose que de très peu d'outils pour rendre compte de cette capacité [Främling et al., 1992]. Il s'agit d'une problématique différente de la science objective. On assiste, alors, à la naissance d'une discipline scientifique appelée logique floue s'intégrant du raisonnement humain.

3.1.2 Systèmes experts flous

L'introduction de la logique floue dans les systèmes experts a pour but de faire face aux limitations des systèmes experts ordinaires surtout celle concernant le raisonnement incertain. Cette introduction évoque la notion des Systèmes Experts Flous. La logique floue a été introduite, dans les systèmes experts depuis une trentaine d'années [Ketata et al., 2005], dans plusieurs disciplines qui manipulent des données incertaines et des informations imprécises.

Elle peut être intégrée sous différents aspects en fonction du domaine de l'application et du degré de complexité du problème. La structure générale d'un système expert flou est représentée par la Figure 1 qui précise également dans quels niveaux la logique floue peut intervenir.

⁶ Un utilisateur qui ne fait pas usage de fonctionnalités avancées, qui ne cherche pas à comprendre le fonctionnement du système, ou qui n'a pas une connaissance poussée dans le domaine concerné.

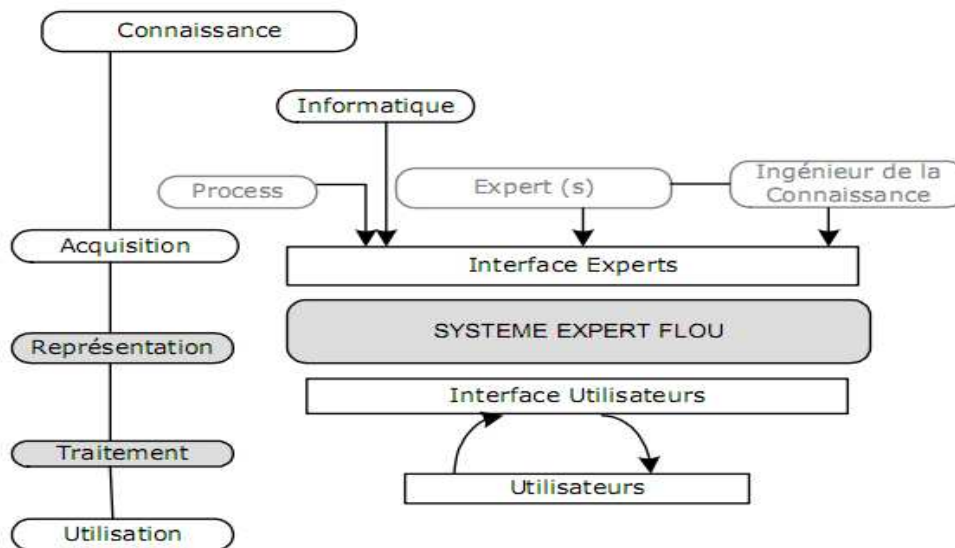


Figure 1. Principe d'un système expert Flou [Ketata et al., 2005]

En fait, cette discipline touche l'étape de la formalisation et du traitement des connaissances et permet de générer un ensemble de conclusions pondérées. La structure interne du système expert flou dépend de son domaine d'application et de la nature des données à représenter.

Ainsi, des systèmes experts flous ont été conçus dans le domaine de l'électronique, essentiellement pour faire le diagnostic des différentes pannes qui peuvent affecter un circuit de transmission [Lee et al., 2000]. La logique floue a été utilisée, également, dans le domaine de diagnostic pour l'agriculture pour la détection et l'identification du type de maladies d'insecte à partir des symptômes constatés sur les plantes [Saini et al., 2002]. Un prototype de logiciel a été, justement, développé dans le but d'étudier la possibilité d'utiliser la logique floue dans des systèmes d'aide à la décision (SAD) pour l'élevage de la vache laitière [Strasser, 1997].

Il a été intéressant, certainement, d'utiliser un système expert flou dans la conception d'architecture d'un réseau informatique. En fait, il s'est avéré que 70% de la connaissance dans ce domaine est de nature floue [Fahmy et al., 1997].

Enfin, le domaine du diagnostic médical avait beaucoup utilisé les systèmes experts flous notamment pour l'analyse du sang de la corde embryonnaire [Garibaldi et al., 2000], dans un logiciel d'aide pour l'enseignement des infirmières [Zajpt et al., 1997] et dans le diagnostic des symptômes du cancer de la peau.

Les systèmes experts flous ont pu, dans plusieurs domaines, prouver leurs efficacités et pallier les problèmes de l'incertitude et l'imprécision des données et des connaissances intervenant

dans l'expertise. L'étude des différents systèmes déjà cités nous a prouvé la diversification de l'utilisation de la logique floue. En effet, la manière de son intégration dans un système expert diffère d'une application à une autre. Elle dépend de la nature de l'incertitude et de sa modélisation. La dotation et la propagation des coefficients de certitude diffèrent lorsque la connaissance du système est modélisée sous modélisation de Sugeno ou bien de Mamdani.

Les systèmes experts sont une des applications de l'intelligence artificielle qui ont quitté les laboratoires de recherche pour être utilisées dans le monde professionnel. De nombreux systèmes experts ont été implantés avec succès pour résoudre des problèmes concrets.

Néanmoins, les grandes difficultés rencontrées pendant l'extraction des connaissances des experts puis pendant leur formalisation forment peut être un point faible trop difficile à contourner dans les Systèmes Experts. La structuration des connaissances, notamment incertaines, reste parfois floue et mal implantée dans les SEs, en plus les contraintes liées à la nature des systèmes diagnostiqués rendent ces approches insuffisantes. Cela remet en question le modèle à la base de connaissance qui pourrait migrer vers d'autres modèles (modèles connexionnistes, systèmes adaptatifs, etc.).

Au début des années quatre-vingts, d'autres approches ont permis de surmonter certaines limites inhérentes à ces méthodes. En particulier, l'utilisation des modèles des systèmes a ouvert de nouvelles voies de recherche. Ces approches sont appelées approche de diagnostic à base de modèles.

3.2 Diagnostic à base de modèles

Le raisonnement à base de modèle est une technique de l'intelligence artificielle qui est applicable pour résoudre de nombreux problèmes comme le contrôle, le diagnostic, etc. [Wotawa, 1999].

Le diagnostic à base de modèles a été introduit dans les années quatre-vingts par [Reiter, 1987] et [Kleer et al., 1987], et a, depuis, été largement diffusé et repris dans de nombreux travaux connexes. Selon ce formalisme, une théorie logique décrit le comportement normal ou anormal d'un système physique. Ensuite, à partir d'observations mesurées sur le système réel, des hypothèses, ou encore des diagnostics, peuvent être émises afin d'expliquer son éventuel mauvais fonctionnement.

Cette approche, qui est née aux Etats Unis d'Amérique dans la communauté Intelligence Artificielle, présente l'avantage de se baser sur le modèle du bon fonctionnement du système et les observations de pannes ne sont utiles qu'au moment du diagnostic. Elle s'appuie uniquement sur la donnée d'un modèle de fonctionnement correct du système et procède par comparaison des comportements du modèle et du système réel.

Le diagnostic à base de modèles est utilisé dans plusieurs applications notamment, dans le diagnostic des fautes graves dans les marchandises [Davis, 1984], [Davis et al., 1988], dans la médecine [Lucas, 1997], ou encore pour la réparation des systèmes [Stumptner et al., 1998], [Friderich et al., 1992], [Friderich et al., 1992a].

Dans cette catégorie de méthode de diagnostic à base de modèle, l'utilisation de réseaux bayésiens comme modèle pour le diagnostic semble une approche intéressante.

3.2.1 Diagnostic par Réseaux Bayésiens

L'intelligence artificielle a exploité des formalismes mathématiques et des propriétés issues de la théorie de graphes pour améliorer la modélisation et la représentation de données incertaines.

Les recherches dans le domaine du diagnostic à base de modèles ont bénéficié de ces avancées. Ainsi, plusieurs travaux ont intégré des calculs de probabilités pour faciliter la tâche de maintenance et de réparation. Des approches classiques ont été augmentées, d'autres ont intégré dans le modèle les connaissances incertaines. Ceci facilite la génération de résultats de nature probabiliste. Les réseaux bayésiens font partie de ces méthodes. Ces modèles intègrent l'aspect incertain dès la création du modèle. Chaque nœud comprend une distribution de probabilités qui est mise à jour en fonction des observations acquises.

Ainsi, les réseaux bayésiens constituent un outil performant pour la modélisation et le diagnostic. En effet, ils permettent, d'une part, d'exploiter pleinement la connaissance *a priori* des systèmes à diagnostiquer et, d'autre part, de fournir l'information *a posteriori* nécessaire dans le contexte de la recherche de diagnostics.

Par conséquent, un grand nombre d'outils utilisant des raisonnements bayésiens existent pour la modélisation des systèmes et le diagnostic médical [Onisko et al., 1998], [Onisko et al., 1999] ou technique. Une description détaillée des réseaux bayésiens sera présentée ultérieurement dans ce rapport.

4 Conclusion

Dans ce chapitre, nous avons présenté quelques techniques de diagnostic ainsi que leurs domaines d'application.

Nous avons détaillé aussi des méthodes de diagnostic. Ces méthodes peuvent être classées en deux catégories : approches à base de connaissances et approches à base de modèles.

Parmi les méthodes de modélisation de système nous avons distingué la logique floue et les réseaux bayésiens. Il s'avère donc intéressant de présenter ces deux techniques. Ceci fera l'objet du chapitre suivant.

Chapitre2

Logique Floue et Réseau Bayésien

*"Il est dans la probabilité que mille
choses arrivent qui sont contraires à
la probabilité." Henry Louis Menckem
(1880-1956).*

1 Introduction

La tâche de diagnostic est devenue de plus en plus pénible surtout face aux problèmes de complexités croissantes des systèmes actuels. Ceci nous impose le recours à des méthodes plus sophistiquées tels que la logique floue et les réseaux bayésiens.

Il est cependant utile de détailler les notions de base de ces deux méthodes, qui sont parmi les techniques les plus intéressantes pour la gestion des connaissances et d'aide à la décision ou plus spécifiquement pour l'aide au diagnostic. C'est ce que propose ce deuxième chapitre.

2 Logique Floue

2.1 Introduction

La plupart des problèmes rencontrés sont modélisables mathématiquement. Mais ces modèles nécessitent des hypothèses parfois trop restrictives, rendant délicate l'application au monde réel. Les problèmes du monde réel doivent tenir compte d'informations imprécises et incertaines. Prenant l'exemple d'une climatisation : si on veut obtenir une température *fraîche*, on peut se demander quelle gamme de températures conviendra (la demande est imprécise). On voit apparaître la difficulté d'interprétation des variables linguistiques comme *frais, chaud,....*

Ainsi, cette limite de la logique classique à représenter l'aspect ambigu du monde réel, a constitué une grande motivation pour la naissance de la logique floue.

2.2 Définition

La logique floue [4] est une branche de l'intelligence artificielle qui aide les ordinateurs à toucher et manipuler des représentations vagues et incertaines.

C'est une approche développée par Lotfi Zadeh⁷, basée sur sa théorie des sous-ensembles flous (*fuzzy sets* en anglais), généralisant la théorie des ensembles classiques. Dans la nouvelle théorie de Zadeh, un élément peut plus ou moins appartenir à un certain ensemble.

Les imprécisions et les incertitudes peuvent ainsi être modélisées, et les raisonnements acquièrent une flexibilité que ne permet pas la logique classique.

⁷ Né le 4 février 1921 à Bakou en Azerbaïdjan, de mère russe et de père azéri iranien, et a ensuite étudié à l'Université de Téhéran. C'est un scientifique connu pour ses travaux en informatique et en automatique.

2.3 Historique

Le concept flou est né en 1965 avec Lotfi Zadeh qui a déclaré qu' « *un contrôleur électromécanique doté d'un raisonnement humain serait plus performant qu'un contrôleur classique* ». En 1980, FLSmidt & Co. A/S ⁸ a mis en application la théorie de la logique floue dans le contrôle de fours à ciment, en fait c'était la première mise en œuvre pratique de cette théorie.

À partir de cette année plusieurs applications ont commencé à émerger, notamment au Japon, jusqu'à l'année 1987 qui a marqué l'explosion de la logique floue au Japon.

2.4 Application

Il n'y a pas en fait de domaine où l'on ne puisse appliquer les raisonnements propres à la logique floue car elle a été conçue pour s'adapter aux techniques de pensée humaine. Tout d'abord il y a les secteurs privilégiés de la commande floue qui sont l'automatisme et la robotique.

Dans ces deux secteurs la commande floue a déjà connu un franc succès et est devenue ainsi la source de nombreuses recherches sur la logique floue. Le troisième secteur très important est l'informatique (aussi bien l'intelligence artificielle que les bases de données et la programmation) où l'on doit souvent traiter des informations vagues et imprécises. Ensuite, viennent l'ingénierie, la gestion et la prise de décision.

Cependant, la logique floue est souvent utilisée en météorologie à cause de la complexité des phénomènes et de l'imprécision naturelle de toute prévision. Elle est déjà utilisée en médecine pour l'aide au diagnostic. Elle sert beaucoup les mathématiques appliquées. Mais aussi les sciences humaines telles que la sociologie et la psychologie car l'homme est tout ce qu'il y a de plus flou.

2.5 Théorie des sous-ensembles flous

Considérons une personne désirant compléter le niveau d'un réservoir contenant de l'eau à une température donnée de façon à remplir ce réservoir d'eau à une température souhaitée T à l'aide d'un mitigeur⁹. Dans un premier temps, nous considérons trois températures possibles de l'eau d'alimentation : froide T_F , chaude T_C et tiède T_T . L'eau du réservoir pourra être

⁸ Fournisseur d'équipement et de services de ciments et d'industries minérales.

⁹ Robinetterie permettant le réglage manuel ou thermostatique de la température et éventuellement du débit du mélange d'eau froide et d'eau chaude.

appréciée comme froide, tiède ou chaude avec une certaine marge d'incertitude si on ne dispose pas d'appareils de mesure de température.

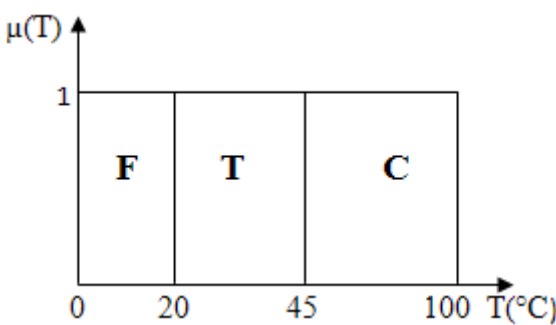
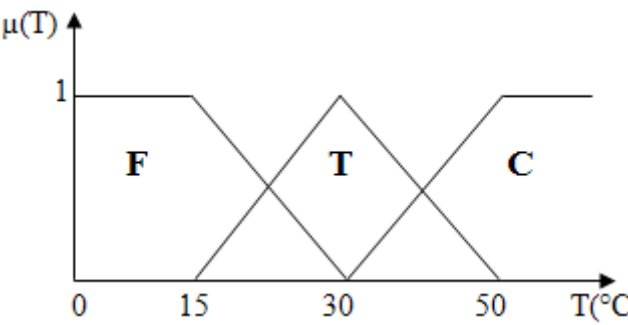
La commande du mitigeur sera simple :

- Si l'eau du réservoir est froide, mettre de l'eau chaude.
- Si l'eau du réservoir est tiède, mettre de l'eau tiède.
- Si l'eau du réservoir est chaude, mettre de l'eau froide.

Sans appareils de mesure précis, on a peut-être donné une information plus nuancée, tel que « tiède et plutôt chaude » et également défini des intervalles de température pour lesquels il n'y a pas d'incertitude comme « l'eau est nettement chaude » ou « réellement froide » et des zones pour lesquelles on peut hésiter.

Pour mieux ressortir la différence entre la théorie classique et la théorie logique floue, nous présentons dans le Tableau 1 une illustration de notre exemple selon ces deux théories.

Tableau 1. Fonctions d'appartenances entre la théorie classique et la théorie floue

Théorie Classique	Théorie Floue
Mathématiquement : $\mu(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{sinon} \end{cases}$	Mathématiquement : $\mu(x) \in [0, 1]$
Représentation graphique : 	Représentation graphique : 

D'après le Tableau 1, la logique classique ne peut utiliser que le 0 et le 1 ainsi l'eau est d'abord totalement froide puis tiède et enfin chaude. Cependant, dans la logique floue, nous pouvons observer la représentation graphique de trois fonctions d'appartenance Froid, Tiède et

Chaud. Ces fonctions se superposent sur des plages de température ayant les qualificatifs froid et tiède ainsi que tiède et chaud. On se rapproche donc du raisonnement humain.

Egalement, en se basant sur le Tableau 1, dans la théorie classique si la température est de 19°C alors elle appartient nécessairement au sous-ensemble Froid avec un degré 1 c'est à dire l'eau est Froide à 100%. Par contre en s'appuyant sur la théorie floue, l'eau de température 19°C est en même temps froide et tiède avec des degrés d'appartenance respectivement 0.8 et 0.2.

La logique floue est basée sur des variables floues dites variables linguistiques à valeurs linguistiques dans l'univers du discours U, où chaque valeur linguistique constitue un ensemble flou de l'univers du discours.

Dans l'exemple précédent nous avons une seule variable linguistique température qui peut prendre l'une des valeurs linguistiques froide, chaude et tiède dans l'univers du discours [0, 100] qui exprime l'intervalle de valeurs que peut avoir la température de l'eau.

2.6 Opérateurs en logique floue

Il s'agit de la généralisation des opérateurs négation, intersection, et union de la théorie classique des ensembles.

2.6.1 L'opérateur NON (complément)

Il est défini mathématiquement par : $\bar{A} = \{x | x \notin A\}$ et représenté par la fonction: $\text{non}(\mu_A(x)) = \mu_{\bar{A}}(x) = 1 - \mu_A(x)$. Ainsi, la Figure 2 montre les deux fonctions $\mu_A(x)$ et $\text{non}(\mu_A(x))$.

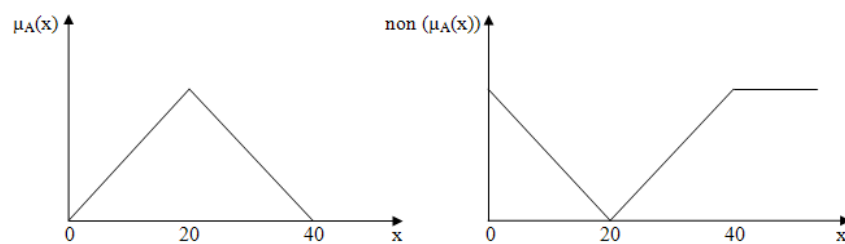


Figure 2. Représentation des fonctions $\mu_A(x)$ et $\text{non}(\mu_A(x))$

2.6.2 L'opérateur ET (intersection)

Il est défini mathématiquement par : $A \cap B = \{x | x \in A \wedge x \in B\}$ et représenté par la fonction $\mu_{A \cap B}(x) = \mu_A(x) \wedge \mu_B(x) = \min(\mu_A(x), \mu_B(x))$. Ainsi la Figure 3 illustre la fonction intersection.

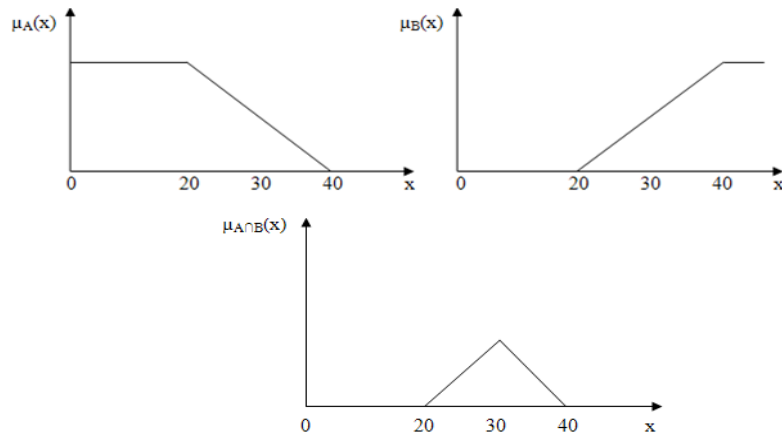


Figure 3. Illustration de la fonction intersection

2.6.3 L'opérateur OU (union)

Il est défini mathématiquement par : $A \cup B = \{x | x \in A \vee x \in B\}$ et représenté par la fonction : $\mu_{(A \cup B)}(x) = \mu_A(x) \vee \mu_B(x) = \max(\mu_A(x), \mu_B(x))$. Ainsi, la Figure 4 illustre la fonction union.

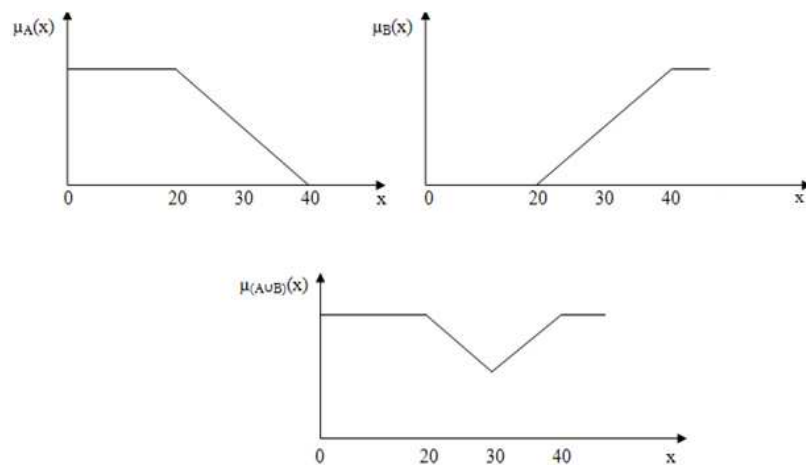


Figure 4. Illustration de la fonction union

2.7 La commande floue

Une loi de commande est habituellement construite à partir d'une approche système basée sur l'utilisation des informations disponibles. Dans plusieurs cas le modèle mathématique du processus est difficile à obtenir surtout lorsque la connaissance provient uniquement de mesures sur des variables caractéristiques et d'une description linguistique du fonctionnement du processus.

Les approches conventionnelles sont inadaptées à l'utilisation de telles connaissances car leur conception nécessite généralement une modélisation mathématique. Par contre, les contrôleurs flous (*fuzzy controller* en anglais) [Raymond et al., 1995], sont recommandés

pour ce type de problème. En effet, la loi de commande est exprimée avec des règles linguistiques, déduites des informations énoncées dans un langage naturel.

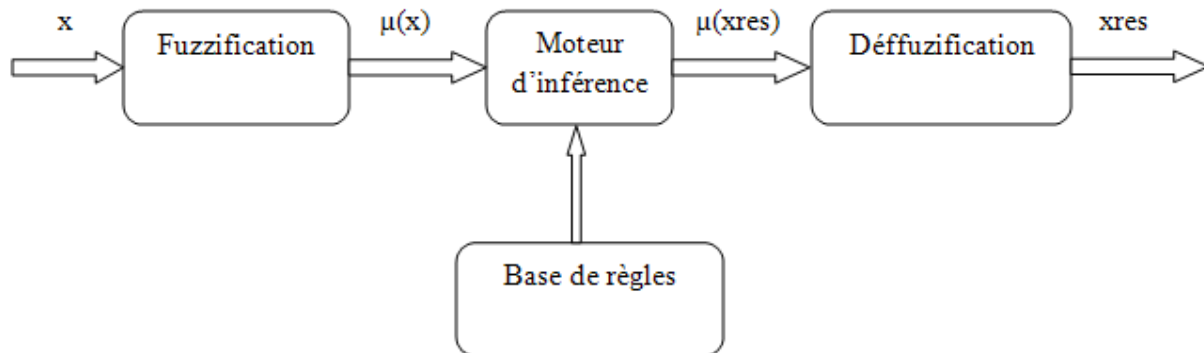


Figure 5. Schéma d'un contrôleur flou

Légende :

x représente le vecteur des entrées, $xres$ celui des commandes, $\mu(x)$ et $\mu(xres)$ les fonctions d'appartenances correspondantes.

Ainsi, la conception d'un système flou passe par trois étapes principales comme indique la Figure 5. La première étape est la *fuzzification* qui consiste à convertir les valeurs d'entrées, qui sont sous formes de grandeurs physiques, en grandeurs floues. La deuxième étape est l'*inférence* (avec la base de règles) qui est en fait la prise des décisions où chaque règle activée donne un sous-ensemble flou de sortie. La troisième et la dernière étape est la *déffuzification*. Elle consiste à convertir les sous-ensembles flous de sortie en valeurs déterminées. Pour mieux expliquer ces trois étapes, nous introduisons un exemple illustratif.

Supposons que nous cherchons à déterminer le freinage qu'il faut appliquer à un véhicule compte tenu de sa vitesse V et la distance D qui le sépare du véhicule suivant, en prenant comme cas pratique $D = 25$ m et $V = 55$ Km/h.

2.7.1 Fuzzification

L'opération de fuzzification permet de passer du domaine réel au domaine du flou. Elle consiste à déterminer le degré d'appartenance d'une valeur (mesurée par exemple) à un ensemble flou. Par exemple (cf. Figure 6), si la valeur courante de la variable « Distance » est 25 m, le degré d'appartenance à la fonction d'appartenance « Faible » est égal à 0,25 et le degré d'appartenance à la fonction d'appartenance « Moyenne » est égal à 0,25.

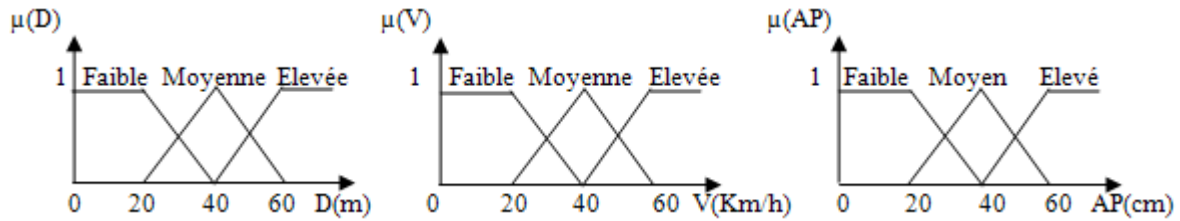


Figure 6. Illustration des sous-ensembles des trois variables Distance, Vitesse et Freinage

Légende :

D : Distance, V : Vitesse, AP : Appuie sur la pédale

2.7.2 Inférence

Tout d'abord, il faut construire la base de connaissance. Les contrôleurs flous utilisent en général une expertise exprimée sous forme de règles, dont la forme, pour un contrôleur à deux entrées et une sortie, est la suivante :

Règle1 : **Si** X1 est A1 **ET** X2 est B2, **alors** Y est C3.

Règle2 : **Si** X1 est A3 **ET** X2 est B1, **alors** Y est C2.

L'expression « X1 est A1 **ET** X2 est B2 » est la prémisse de la règle1, tandis que l'expression « Y est C3 » est la conclusion de cette règle.

Revenant maintenant à notre exemple, nous disposons en faite de deux règles :

Règle1 : **Si** la distance entre les deux véhicules est Faible **et** que la vitesse de la voiture est Moyenne, appliquer **alors** un freinage Elevé pour obtenir une réduction rapide de la vitesse.

Règle2 : **Si** la distance entre les deux véhicules est Moyenne **et** que la vitesse de la voiture est Elevée, freiner **alors** modérément.

Nous passons ensuite au traitement des règles, en traduisant les opérateurs Et, OU et l'implication par l'une des fonctions vue précédemment (Minimum, Maximum, Produit,...).

Il existe en faite plusieurs méthodes d'inférences, notamment Max-Prod, Somme-Prod et Max-Min. Ainsi, dans cette dernière méthode, au niveau de la condition le « **Et** » est représenté par la fonction « **Min** » et le « **Ou** » est représenté par la fonction « **Max** ». Alors qu'au niveau de la conclusion le « **Ou** » est représenté par la fonction « **Max** » et le « **Alors** » est représenté par la fonction « **Min** ».

En appliquant cette méthode à notre exemple, nous obtenons les résultats indiqués dans la Figure 7:

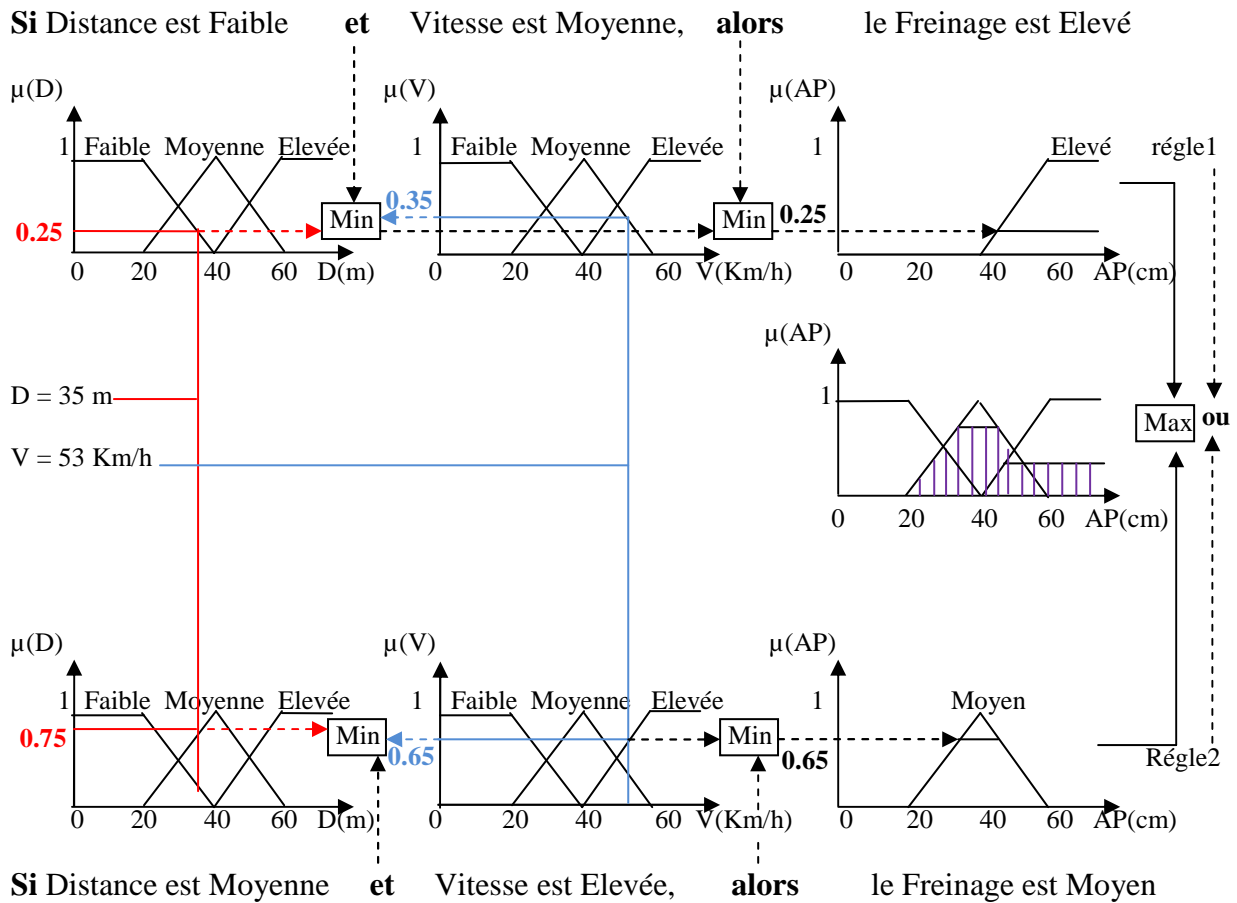


Figure 7. Illustration de l'inférence en appliquant la méthode Max-Min

En procédant à la première règle on obtient :

$D = 35$ m est Faible avec un degré de 0,25 et $V = 53$ Km/h est Moyenne avec un degré de 0,35. Au niveau de la condition on a (D est Faible **Et** V est Moyenne) donc il faut prendre la valeur minimale entre 0,25 et 0,35 ce qui donne la valeur 0,25. Au niveau de la conclusion on a un « **Alors** » qui sera remplacé par la fonction min c'est-à-dire il faut tronquer la fonction d'appartenance *Freinage est Elevé* par la valeur 0,25.

De même d'après la deuxième règle on obtient :

$D = 35$ m est Moyenne avec un degré de 0,75 et $V = 53$ Km/h est Elevé avec un degré de 0,65. Au niveau de la condition on a (D est Moyenne **Et** V est Elevée) donc il faut prendre la valeur minimale entre 0,75 et 0,65 ce qui donne la valeur 0,65. Au niveau de la conclusion on

a un « **Alors** » qui sera remplacé par la fonction min c'est-à-dire il faut tronquer la fonction d'appartenance *Freinage est Moyen* par la valeur 0,65.

Enfin, nous obtenons une fonction d'appartenance représentée par la surface hachurée. Cette fonction sera traitée lors de la troisième phase de défuzzification.

2.7.3 Défuzzification

A la fin de l'inférence, l'ensemble flou de sortie est déterminé mais il n'est pas directement utilisable pour donner une information précise. Il est nécessaire de passer du « monde flou » au « monde réel » qui est assuré par la défuzzification. En faite, il existe plusieurs méthodes, notamment la méthode du maximum, la méthode de la moyenne des maxima, la méthode de la somme pondérée et la méthode du centroïde qui est le plus souvent rencontrée.

Ainsi, dans cette méthode, la sortie correspond à l'abscisse du centre de gravité de la surface de la fonction d'appartenance résultante.

2.8 Conclusion

Les outils fournis par la logique floue [Sur et al., 1998] permettent une modélisation des phénomènes pouvant, en un certain sens, s'approcher du raisonnement humain. Le fait de transcender le « tous ou rien » des ordinateurs introduit une souplesse faisant la puissance des outils flous dans de nombreux domaines.

Au milieu des années 80, plusieurs applications industrielles utilisant la logique floue ont vu le jour, et ce essentiellement en Asie du Sud Est, l'Europe et l'Amérique. Elles vont du contrôle du métro automatique à l'élimination du tremblement pour les caméras vidéo en passant par le réglage de cycle sur une machine à laver.

En outre, la flexibilité des modèles flous a permis également des applications dans des domaines tels que la médecine (aide au diagnostic), la finance (prévision boursière, opération de change), la météorologie, etc.

Mais même bénéficière d'un effet de mode, les algorithmes flous ne sont pas nécessairement les meilleurs. D'autres méthodes, par exemple les réseaux bayésiens, sont aussi performants n'opèrent pas de la même façon et donnent des résultats satisfaisants dans un contexte incertain.

3 Réseaux Bayésiens

3.1 Introduction

Le domaine de la gestion des connaissances, qui connaît un intérêt croissant, est donc également un champ d'application potentiel pour les Réseaux Bayésiens, qui sont actuellement l'une des techniques les plus intéressantes de l'Intelligence Artificielle, dans la mesure où ceux-ci offrent un formalisme riche et intuitif de représentation de la connaissance. En fait, les Réseaux Bayésiens constituent aujourd'hui l'un des formalismes les plus complets et les plus cohérents pour l'acquisition, la représentation et l'utilisation de connaissances par ordinateurs.

3.2 Définition

Les réseaux bayésiens sont des modèles graphiques qui représentent les relations entre les différentes variables qui expriment des événements avec leurs probabilités de réalisation en tenant compte des liens qui existent entre les variables.

Un réseau bayésien est un outil de représentation des connaissances, qui permet de calculer des probabilités conditionnelles, apportant ainsi des solutions à différentes sortes de problématiques. La structure de ce type de réseau est simple : En fait, selon Judea Pearl « *les réseaux bayésiens sont des graphes acycliques orientés pour lesquels les nœuds représentent des variables aléatoires et les arcs représentent l'indépendance conditionnelle entre les différents nœuds* » [Pearl, 1988].

Formellement, un réseau bayésien est défini dans [Naïm et al., 2004] par :

- un graphe acyclique orienté G , $G = (V, E)$, où V est l'ensemble des nœuds de G , et E l'ensemble des arcs de G ,
- un espace probabiliste fini (Ω, Z, p) ,
- un ensemble de variables aléatoires associées aux nœuds du graphe et définies sur (Ω, Z, p) , tel que :

$$p(V_1, V_2, \dots, V_n) = \prod_{i=1}^n p(V_i | C(V_i)), \text{ où } C(V_i) \text{ est l'ensemble des causes (parents) de } V_i$$

dans le graphe G .

3.3 Historique

Encore du domaine de la recherche au début des années 90, cette théorie connaît de plus en plus d'applications. Le domaine d'application varie du contrôle de véhicules autonomes à la modélisation des risques opérationnels, en passant par le data mining et la localisation des gènes.

Les réseaux bayésiens, sont le résultat de recherches effectuées dans les années 80, dues à J. Pearl à UCLA (University of California, Los Angeles) et à une équipe de recherche danoise à l'université d'Alborg.

Cependant, les réseaux bayésiens doivent leur nom aux travaux de Thomas Bayes au XVIII^e siècle sur la théorie de probabilités, par son célèbre théorème.

3.4 Le théorème de Bayes

Thomas Bayes (1702-1761) est né à Londres en Angleterre et a développé un théorème qui porte sur le calcul de la probabilité d'un événement basé sur une connaissance *a priori*. Le théorème, de façon indirecte, fut publié à titre posthume en 1763 sous le titre de « *Essay Towards Solving a Problem in the Doctrine of Chance* ». En gros, l'application du théorème permet d'établir que la probabilité d'un événement est le résultat conditionnel d'une probabilité connue. La formule du théorème est la suivante :

$$P(M_i | A) = \frac{P(A | M_i)P(M_i)}{\sum_k P(A | M_k)P(M_k)}$$

$P(M_i)$: probabilité *a priori* de M_i .

$P(A | M_i)$: probabilité de A conditionnellement à M_i .

$P(M_i | A)$: probabilité *a posteriori* de M_i conditionnellement à A.

3.5 Construction du réseau bayésien

La construction d'un réseau bayésien passe par trois étapes principales [Maalej, 2006]. La première est l'étape qualitative qui correspond à l'identification des variables et de leur espace d'états en tenant compte des relations d'influence qui existent entre les variables.

La deuxième étape est l'étape probabiliste qui consiste à distribuer les probabilités sur les variables et l'appliquer au graphe qui représentera le réseau bayésien.

Finalement, l'étape quantitative qui consiste à spécifier numériquement les distributions de probabilités conditionnelles.

3.6 Modélisation par réseau bayésien

Pour mieux comprendre l'intérêt des réseaux bayésiens, nous allons définir les notions de graphe causal, de d-séparation et d'indépendance conditionnelle puis nous décrivons les algorithmes d'inférence. Nous abordons également l'inférence et l'apprentissage dans les réseaux bayésiens [Naïm et al., 2004].

3.6.1 Liens d'influences et graphes causaux

Un graphe causal est une représentation graphique dans laquelle les causes sont liées aux effets par des flèches orientées (cf. Figure 8).



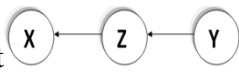
Figure 8. Graphe causal entre A et B

S'il existe une relation causale de A vers B (Figure 8), toute information sur A peut modifier la connaissance que nous avons de B, et réciproquement, toute information sur B peut modifier la connaissance que nous avons de A. Le sens des flèches indique un lien de causalité mais pas un sens de propagation de la connaissance.

Les connexions entre les nœuds définissent des lois de circulation de l'information dans le graphe. On distingue trois types de connexions :

- *Connexion convergente ou Connexion en V* : dans ce cas, l'information ne peut circuler de X à Y que si Z est connu.
- *Connexion en série* : l'information ne peut circuler de X à Y que si Z n'est pas connu, sinon c'est directement la connaissance sur le nœud Z qui influe Y.
- *Connexion divergente* : dans ce cas, l'information ne peut circuler de X à Y que si Z n'est pas connu.

Ainsi, la circulation de l'information à l'intérieur d'un graphe causal dépend du type de la connexion, plutôt que du sens des flèches. Des exemples plus détaillés sont dans [Naïm et al., 2004].

Dans une connexion série ou divergente, le graphe donne une information sur l'indépendance de X et Y : X et Y sont indépendants sachant Z . Ces deux cas (et d'ailleurs également ) peuvent être associés à la même distribution de probabilité sur X , Y et Z . Alors que la connexion convergente a nécessairement une distribution de probabilités différente.

3.6.2 D-séparation et indépendance conditionnelle

La notion de *d-séparation* est essentielle pour le calcul des probabilités car elle permet de définir l'indépendance conditionnelle entre certains nœuds : Soient X , Y et Z trois nœuds du graphe. On dit que X et Y sont d-séparés par Z (et on note $\langle X|Z|Y \rangle$), si pour tous les chemins entre X et Y , l'une au moins des deux conditions suivantes est vérifiée :

- Le chemin converge en un nœud W , tel que $W \neq Z$, et W n'est pas une cause directe de Z .
- Le chemin passe par Z , et est soit divergent, soit en série au nœud Z .

On dit que X et Y sont *indépendants conditionnellement* à Z , et on note $X \perp Y | Z$, si l'une des propriétés équivalentes suivantes est vérifiée :

- $P(X | Z, Y) = P(X | Z)$
- $P(X, Y | Z) = P(X | Z) \cdot P(Y | Z)$

Dans le cas général, l'indépendance conditionnelle ($X \perp Y | Z$) possède les quatre propriétés suivantes :

- Symétrie $X \perp Y | Z \Leftrightarrow Y \perp X | Z$
- Décomposition $X \perp (Y \cup W) | Z \Rightarrow X \perp Y | Z$
- Union faible $X \perp (Y \cup W) | Z \Rightarrow X \perp W | (Z \cup Y)$
- Contraction $X \perp Y | Z \wedge X \perp W | (Z \cup Y) \Rightarrow X \perp (Y \cup W) | Z$

Une cinquième propriété peut être ajoutée [Pearl, 2000] dans le cas où toutes les distributions de probabilités sont strictement positives :

- Intersection $X \perp W | (Z \cup Y) \wedge (X \perp Y | (Z \cup W)) \Rightarrow (X \perp (Y \cup W) | Z)$

Enfin, voici le théorème fondamental des réseaux bayésiens.

« Si X et Y sont d-séparés par Z , alors X et Y sont indépendants sachant Z » :

$$\langle X|Z|Y \rangle \Rightarrow P(X|Y,Z) = P(X|Z)$$

Ainsi la structure du graphe d'un réseau bayésien permet de prendre en compte des indépendances conditionnelles qui seront très utiles pour les calculs d'inférence. Il faut cependant noter que le réseau bayésien ne révèle pas toutes les indépendances conditionnelles. Plusieurs structures de graphes donnent une même distribution de probabilités jointes, certaines étant meilleures que d'autres, il faut en tenir compte lors de la construction du modèle.

Une description plus détaillée sur la d-séparation et l'indépendance conditionnelle est développée dans [Naïm et al., 2004].

3.7 Inférence dans les réseaux bayésiens

L'inférence, ou *la mise à jour des probabilités*, dans un réseau bayésien se résume à un calcul de probabilités *a posteriori*. Connaissant les états de certaines variables (appelées variables d'observation), on détermine les probabilités des états de certaines autres variables (appelées variables cibles) conditionnellement aux observations.

Il existe plusieurs algorithmes d'inférence dans les réseaux bayésiens classés en deux groupes [Kotchi et al., 2003]. D'un côté nous avons les méthodes d'inférence exactes qui exploitent les indépendances conditionnelles contenues dans les réseaux et donnent à chaque inférence les probabilités *a posteriori* exactes. Par exemple l'algorithme Clustering [Lauritzen et al., 1988] effectue l'inférence en transformant le réseau en un arbre pour lequel chaque nœud regroupe plusieurs nœuds du réseau initial. De l'autre côté nous avons les méthodes approchées qui estiment les probabilités *a posteriori*. Pour ces méthodes, deux exécutions d'une inférence peuvent donner des probabilités *a posteriori* différentes.

Comme exemple de méthodes approchées, nous pouvons citer les algorithmes d'échantillonnage stochastique (Logic sampling [Henrion, 1988], Likelihood weighting [Fung et al., 1989], Backward sampling [Fung et al., 1994], Self importance [Shachter et al., 1989] et Heuristic importance [Shachter et al., 1989]) qui estiment les probabilités en effectuant plusieurs tirages dans l'ensemble des combinaisons possibles des états des variables du réseau.

3.8 Apprentissage dans les réseaux bayésiens

Nous avons déjà précisé qu'un réseau bayésien est constitué à la fois d'un graphe (aspect qualitatif) et d'un d'ensemble de probabilités conditionnelles (aspect quantitatif). L'apprentissage d'un réseau bayésien doit donc rependre aux deux questions suivantes :

- Comment estimer les lois de probabilités conditionnelles ?
- Comment trouver la structure du réseau bayésien ?

Le problème de l'apprentissage sera donc séparé en deux parties :

- *L'apprentissage des paramètres*, où la structure du réseau est fixée, et où il faudra estimer les probabilités conditionnelles de chaque nœud du réseau.
- *L'apprentissage de la structure*, dont le but est de trouver le meilleur graphe représentant la tâche à résoudre.

Comme pour tout problème d'apprentissage, différentes techniques sont possibles selon la disponibilité de données concernant le problème à traiter, ou d'expert de ce domaine. Ces techniques peuvent se partager en deux grandes familles :

- **apprentissage à partir de données**, complètes ou non, par des approches statistiques ou bayésiennes,
- **acquisition de connaissances** avec un expert du domaine.

Une description plus détaillée sur l'apprentissage est développée dans [Naïm et al., 2004].

3.9 Applications utilisant les réseaux bayésien

Les domaines d'applications des réseaux bayésiens, et les types d'applications sont très variés. D'une manière générale, un réseau bayésien sert à représenter la connaissance que l'on a d'un système (technique, informatique, biologique, sociologique, économique, etc.) en vue de :

- **Prévoir** (le comportement d'un système)
- **Diagnostiquer** (les causes d'un phénomène observé dans un système)
- **Contrôler** (le comportement d'un système)
- **Simuler** (le comportement d'un système)
- **Analyser des données** (relatives au système)
- **Prendre des décisions** (concernant un système)
- Etc.

Une des raisons du boom actuel dans l'utilisation des réseaux bayésiens, réside dans leur convivialité et leur efficacité. Ce qui fait qu'ils sont présents dans une multiplicité d'applications dans les domaines de l'industrie, du marketing, de la santé, de la banque, de la finance, du droit, etc. Le "système" dont on représente la connaissance au moyen d'un réseau bayésien, peut être aussi bien le contenu du caddie d'un client de supermarché, un navire de la marine, le patient d'une consultation médicale, le moteur d'une automobile, un réseau électrique, l'utilisateur d'un logiciel, etc. [3].

En effet, on peut citer deux applications particulièrement ambitieuses sur l'un des aspects de l'utilisation des réseaux bayésiens : la détection de fraude¹⁰, pour ce qui est de l'apprentissage, et l'aide à la décision en situation critique¹¹ pour l'inférence [Naïm et al., 2004].

Afin d'illustrer la puissance des réseaux bayésiens nous allons étudier un exemple de représentation et de manipulation.

3.10 Exemple simple faisant intervenir des faits incertains

Le petit exemple [3] que nous allons utiliser pour introduire le concept de réseau bayésien est une adaptation d'un extrait du manuel de la méthode SERENE¹².

Imaginons que nous devons modéliser la connaissance suivante : « Fantasio et Gaston vont à leur travail en utilisant des moyens de transport différents. Gaston utilise sa voiture, alors que Fantasio voyage en train. Fantasio manque rarement son train qui est presque toujours à l'heure, sauf les jours de grève. Toutefois, une grève de train n'implique pas forcément que Fantasio soit en retard (il peut partir tôt en voiture). Une grève de train peut aussi retarder Gaston car elle provoque des embouteillages. Mais Gaston est de toute façon souvent en retard parce qu'il n'entend pas la sonnerie de son réveil, et de ce fait, une grève n'augmente la probabilité de son retard que d'une faible quantité. En cas de grève, Gaston a moins de chances d'être en retard que Fantasio. ».

¹⁰ L'une des applications qui fait référence pour l'utilisation des réseaux bayésiens par le Data Mining mis en production à la fin des années 1990 par la société américaine de télécommunication ATT.

¹¹ Application qui a été développée par la NASA en collaboration avec la société californienne Knowledge Industries.

¹² C'est un projet qui regroupe, dans le cadre du programme de recherche européen Esprit, plusieurs partenaires cherchant à développer une méthodologie d'utilisation des réseaux bayésiens dans le cadre du contrôle qualité du logiciel pour les systèmes critiques.

Maintenant, étant donné cette connaissance, comment pourrions-nous modéliser les inférences suivantes, issues d'un raisonnement intuitif ?

- Si nous savons que Fantasio est en retard, nous pensons qu'il y a une grève des trains, et donc que Gaston risque (un peu) plus que d'habitude d'être en retard.
- Supposons que nous sachions que Gaston est en retard. Cette constatation augmente notre croyance en les deux causes possibles de ce retard (grève, réveil non entendu). Mais si nous apprenons que Fantasio est également en retard, nous serons tentés d'en déduire qu'une grève de train est en cours, et a été la cause du retard de Gaston, ce qui fait retomber quelque peu notre croyance en le fait qu'il n'a pas entendu son réveil.

Exemple de représentation de connaissances par Réseau Bayésien :

Un réseau bayésien est un graphe constitué de nœuds et d'arcs, associé à un ensemble de tables de probabilités conditionnelles appelées aussi tables de probabilités de nœuds (TPN), ainsi nommées car il y en a une et une seule par nœud du graphe.

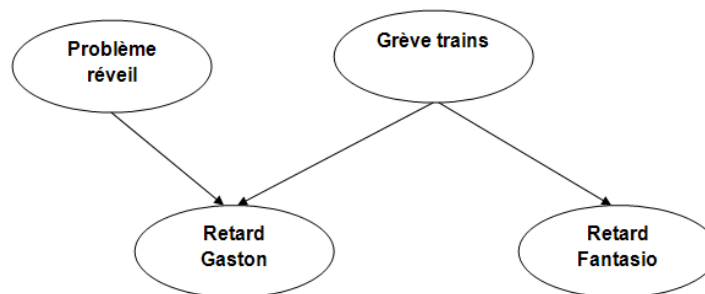


Figure 9. *Graphe du réseau bayésien*

Les nœuds représentent des variables aléatoires discrètes (il y a quelques extensions des réseaux bayésiens vers le domaine des variables continues, mais elles sont soumises à de fortes limitations sur les types de distributions utilisables : c'est pourquoi nous nous limiterons au cas des variables discrètes). Dans notre exemple, les quatre variables ont seulement deux états : 'Vrai' et 'Faux'.

Les arcs représentent des relations de cause à effet entre variables. Comme une grève peut provoquer le retard de Fantasio, nous modélisons cette relation par un arc allant du nœud 'Grève trains' au nœud 'Retard Fantasio' (cf. Figure 9).

Le grand avantage des réseaux bayésiens est de permettre de modéliser des relations non déterministes. Dans notre exemple, voici le Tableau 2 qui pourrait être la table de probabilités modélisant la dépendance entre le retard de Fantasio et la grève des trains :

Tableau 2. *Table de probabilité (Retard Fantasio)*

	Grève trains = Vrai	Grève trains = Faux
P (Retard Fantasio = Vrai)	0.6	0.1
P (Retard Fantasio = Faux)	0.4	0.9

C'est en fait la distribution de probabilité de la variable 'Retard Fantasio', conditionnelle à la variable 'Grève trains' : $P(\text{Retard Fantasio} \mid \text{Grève train})$.

Cette table exprime d'une manière formelle et précise le fait que Fantasio a très peu de chances d'être en retard en temps ordinaire, mais que s'il y a une grève des trains, au contraire, il risque fort d'être en retard (la probabilité est de 0.6).

Afin de formaliser notre connaissance décrite plus haut, nous modélisons la relation entre le retard de Gaston et ses deux causes possibles par la table de probabilité représentée dans le Tableau 3:

Tableau 3. *Table de probabilité (Retard Gaston)*

Pb. réveil	Vrai		Faux	
Grève trains	Vrai	Faux	Vrai	Faux
P (Retard Gaston = Vrai)	0.7	0.5	0.4	0.1
P (Retard Gaston = Faux)	0.3	0.5	0.6	0.9

Les tables de probabilités associées aux nœuds 'Grève trains' et 'Pb. réveil' ont une nature quelque peu différente. Ces nœuds n'ont pas de nœud parent dans ce modèle (ce sont des nœuds racines), et nous n'avons donc qu'à leur assigner des probabilités pour leurs deux valeurs 'Vrai' et 'Faux'. En fait, nous supposons que $P(\text{Grève trains} = \text{Vrai}) = 0.1$, et que $P(\text{Pb. réveil} = \text{Vrai}) = 0.4$.

Les nœuds racines modélisent des variables indépendantes entre elles. C'est bien le cas des nœuds Pb réveil et grève train.

Il peut y avoir diverses manières de déterminer les probabilités des TPNs. Par exemple, nous pourrions tirer $P(\text{Grève trains} = \text{Vrai})$ d'un historique sur les jours de grève. Mais en absence de telles données de retour d'expérience, il est toujours possible de faire appel à des valeurs de probabilités subjectives, évaluées par des experts.

L'avantage des réseaux bayésiens est de pouvoir mêler dans un cadre théorique unique (la théorie des probabilités) les probabilités issues d'un traitement statistique de retour d'expérience, et les probabilités subjectives.

Afin de mettre en évidence les inconvénients et les avantages des réseaux bayésiens nous présentons dans ce qui suit une étude comparative de la technique RB vis-à-vis d'autres techniques.

3.11 Comparatif de la technique réseau bayésien vis-à-vis d'autres techniques

Tableau 4. Comparatif des réseaux bayésiens à d'autres techniques [Naïm et al., 2004]

Connaissances	Analyse de données	Réseaux neuronaux	Arbre de décision	Système experts	Réseaux bayésiens
<u>ACQUISITION</u>					
Expertise seulement				*	
Données seulement	+	*	+		+
Mixte	+	+	+		*
Incrémental		+			*
Généralisation	+	*	+		+
Données incomplètes		+			*
<u>REPRESENTATION</u>					
incertitude				+	*
Lisibilité	+		+	+	*
Facilité		+	*		
Homogénéité					*
<u>Utilisation</u>					
Requête élaborées	+			+	*
Utilité économique	+	+			*
Performance	+	*			

Un réseau bayésien peut être avantageusement substitué à différents modèles d'évaluation, de prévision, de diagnostic, d'aide à la décision ou de data mining. On peut donc citer comme techniques concurrentes, à titre d'exemple, les réseaux neuronaux, les chaînes de Markov, les arbres de décision, les arbres de défaillances.

Du point de vue des applications, les avantages et les inconvénients des réseaux bayésiens par rapport à quelques-unes des techniques concurrentes peuvent se résumer dans le Tableau 4 .

Nous avons regroupé les avantages et les inconvénients selon les trois rubriques : Acquisition, Représentation et Utilisation de connaissances. La représentation adoptée est la suivante :

- A chaque ligne correspond une caractéristique, qui peut être un avantage, ou la prise en compte d'un problème spécifique.
- Si la technique considérée permet de prendre en compte ce problème, ou présente cet avantage, un signe + est placé dans la case correspondante.
- Un signe * est placé pour la meilleure technique du point de vue de la caractéristique considérée.

En faite, cette étude comparative montre les avantages des réseaux bayésiens par rapport aux techniques concurrentes. On peut résumer ces avantages par :

- La possibilité de rassembler et de fusionner des connaissances de diverses natures dans un même modèle : données de retour d'expérience, expertise (exprimée sous forme de règles logiques, d'équations, ou de probabilités subjectives), et observations.
- La « convivialité » : un réseau bayésien est un modèle graphique, compréhensible et manipulable par un non-spécialiste.
- La « versatilité » : on peut se servir d'un même modèle pour évaluer, prévoir, diagnostiquer, et optimiser des décisions.

De l'autre côté, d'après le Tableau 4, on peut remarquer la limite des RBs au niveau de la performance en terme de complexité des algorithmes de mise à jour. En effet, la

généralité du formalisme des réseaux bayésiens aussi bien en termes de représentation que d'utilisation les rend difficiles à manipuler à partir d'une certaine taille.

3.12 Conclusion

Les réseaux bayésiens sont actuellement une des techniques les plus intéressantes de l'IA car ils permettent la représentation de la connaissance par un graphe causal intuitif et compressible. De plus comme ils sont basés sur des probabilités, ils intègrent l'incertitude dans le raisonnement.

4 Conclusion

Dans ce chapitre nous avons présenté deux techniques utilisées en diagnostic : la logique floue et les réseaux bayésiens. Nous avons relaté leurs avantages et les atouts qu'ils offrent, surtout pour le diagnostic.

Cependant, malgré la flexibilité des modèles flous et leurs efficacités, les réseaux bayésiens peuvent les surpasser dans la résolution de problème grâce à leur pouvoir expressif et facilité d'interprétation.

De l'autre côté, l'une des limites majeure que rencontrent les utilisateurs des réseaux bayésiens est d'exprimer les connaissances non précises surtout dans les diagnostics où beaucoup d'informations sont flous. En faite, nous ne pouvons pas exprimer toutes sortes de connaissances dans un langage précis. La solution est donc de combiner les réseaux bayésiens avec la logique floue, d'où la naissance des Réseaux Bayésiens Flous.

CHAPITRE 3

Présentation des Réseaux Bayésiens Flous : Fondement et outil

*"Il est dans la probabilité que mille
choses arrivent qui sont contraires à
la probabilité." Henry Louis Menck en
(1880-1956).*

1 Introduction

Jusqu'à maintenant nous avons détaillé quelques notions de base de la logique floue et des réseaux bayésiens. Bien que toutes les deux présentent un outil d'aide à la décision, ces deux théories de l'intelligence artificielle sont différentes dans plusieurs aspects. Une étude comparative, entre la théorie de probabilité (sur laquelle reposent les Réseaux Bayésiens) et la logique floue, s'avère donc intéressante à mener :

Tableau 5. Comparaison entre la théorie de probabilité et la logique floue

Probabilités	Logique Floue
Exprime l' incertitude par rapport à l' occurrence d'un événement.	Exprime l' ambigüité par rapport à la nature d'un événement.
$P(A \cup \bar{A}) = P(U) = 1$	$\mu_{(A \cup \bar{A})}$ pas nécessairement = $\mu(U)$
$P(A \cap \bar{A}) = P(\emptyset) = 0$	$\mu_{(A \cap \bar{A})}$ pas nécessairement = $\mu(\emptyset)$

D'après le Tableau 5 la théorie des probabilités est différente de la théorie de la logique floue bien que toutes deux décrivent une notion de doute, d'incertain et ce à l'aide de nombre compris entre 0 et 1.

Généralement, l'incertitude est groupée dans deux catégories : L'aspect aléatoire et le manque de précision.

D'une part, pour surmonter le problème du manque de précision ou encore de l'ambigüité, qui est causé par un concept mal défini de l'observation, on a recours à la théorie de la logique floue.

D'autre part, pour remédier au problème de l'aspect aléatoire qui est causé par les événements imprévisibles, la théorie de probabilité est plus appropriée.

D'après *Shijium Qui*, « La fusion des deux techniques Réseau Bayésien et Logique Floue aboutit à un résultat plus précis et robuste que l'utilisation de l'un ou l'autre outil seul » [Qiu et al., 2001].

Nombreux travaux s'orientent vers un panachage de ces deux techniques, et c'est ainsi que nous assistons à l'apparition des réseaux bayésiens flous.

Ce troisième chapitre est consacré à la présentation de Réseau Bayésien Flou (RBF), notamment, sa définition, les notions de base de cette nouvelle théorie et quelques applications.

2 Définition

Les Réseaux bayésiens Flous sont une généralisation de réseaux bayésiens classiques aux réseaux bayésiens avec des variables à états flous [Fogelberg, 2008].

3 La théorie de Réseau Bayésien Flou

Comme nous avons déjà signalé, les recherches et les applications dans les RBFs sont rares que nous classons en deux axes.

D'une part, Le premier axe concerne la **discrétisation des variables continues**. En faite, l'utilisation de ces variables dans les RBs exige une discrétisation de leur domaine. Cependant, la discrétisation classique exige la division du domaine d'une variable en un nombre finis d'intervalles puis l'affectation d'une valeur discrète à chaque intervalle. D'après cette méthode tous les points d'un intervalle auront la même valeur discrète et seront traités de la même façon s'ils sont dans le centre ou dans les bords de l'intervalle.

Pour cette raison, les différentes méthodes de choix d'échantillonnage peuvent donner des résultats différents selon les valeurs considérées. Ainsi, la méthode la plus lisse d'après [Baldwin et al., 2003] est l'utilisation d'une division floue, qui doit couvrir le domaine de la variable avec le chevauchement de fonctions floues dont la somme des degrés d'appartenances égale à 1.

Pour mieux visualiser l'utilité des RBFs dans la discrétisation des variables continues, des références peuvent être consultées tel que : [Baldwin et al., 2003], [Lin et al., 2006], [Park et al., 2006], et [Pan et al., 1999].

D'autre part, le deuxième axe de recherche dans les RBFs concerne **l'utilisation des variables ou événements flous** qui résout le problème de l'ambiguïté.

En effet, nous pouvons rencontrer des problèmes, comme le cas de diagnostic, où nous ne sommes pas capables d'exprimer toutes les sortes de connaissances dans un langage logique précis. Dans ce cas l'utilisation des RBFs s'avère une solution très intéressante.

Afin d'expliquer cette technique, nous détaillons dans la suite deux approches traitant des RBFs [Fogelberg, 2008].

3.1 Première approche de Réseau Bayésien Flou

3.1.1 Probabilité Flou

Soit $\mu_{\bar{A}} : U \rightarrow [0,1]$ avec « $\mu_{\bar{A}}$ » est une fonction d'appartenance du domaine « U », ou encore l'univers du discours, vers l'intervalle [0, 1]. Alors, $\mu_{\bar{A}}(u)$ est le degré d'appartenance de « u » au sous-ensemble flou « \bar{A} ».

Prenons l'exemple [Tang et al., 2007] suivant : supposons que « \bar{A} » est un état flou dans un espace de probabilité (X, B, P) et la fonction d'appartenance $\mu_{\bar{A}}(u)$ est une fonction mesurable. Ainsi, nous pouvons considérer « \bar{A} » comme un événement flou dans X de probabilité flou :

$$P(\bar{A}) = \sum_{x \in X} \mu_{\bar{A}}(x).P(x)$$

3.1.2 Equation Bayésienne Floue

Proposons maintenant un exemple qui introduit l'équation du concept bayésien flou.

On suppose : B1 exprime : âge > 50 ; B2 exprime : $50 \geq \text{âge} > 35$; B3 exprime : $35 \geq \text{âge} \geq 20$; B4 exprime : âge < 20.

A1 exprime : revenu annuel > 10000 ; A2 exprime : $10000 \geq \text{revenu annuel} \geq 3000$; A3 exprime : revenu annuel < 3000.

Et les deux tableaux Tableau 6 et Tableau 7 qui présentent respectivement les résultats du recensement de la population et des données du bureau fiscal.

Tableau 6. *Résultat du recensement de la population*

B1	B2	B3	B4
20%	25%	30%	25%

Tableau 7. *Données du bureau fiscal*

$P(A_j B_i)$	B1	B2	B3	B4
A1	0.1	0.2	0	0
A2	0.5	0.6	0.8	0
A3	0.4	0.2	0.2	1.0

Supposons maintenant qu'un revenu annuel d'un homme est plus de 10000, et les probabilités suivantes : $P(A_1) = 0.07$; $P(A_2) = 0.49$; $P(A_3) = 0.44$.

Quelle est la probabilité de chaque phase d'âge ?

Nous cherchons donc « $P(B_i|A_1)$ » pour tous $i \in \{1, 2, 3, 4\}$.

Selon l'équation bayésienne nous aurons :

$$P(B_1 | A_1) = \frac{P(A_1 | B_1).P(B_1)}{P(A_1 | B_1).P(B_1) + P(A_1 | B_2).P(B_2) + P(A_1 | B_3).P(B_3) + P(A_1 | B_4).P(B_4)}$$

$$= 0.2857$$

De même nous trouvons : $P(B_2 | A_1) = 0.7143$; et $P(B_3 | A_1) = P(B_4 | A_1) = 0.00$

Passons maintenant au domaine flou.

En premier lieu, si nous considérons l'événement flou \bar{B} de la variable B_i qui présente l'état « non trop vieux », nous obtiendrons l'équation bayésienne floue :

$$P(\bar{B} | A_j) = \frac{\sum_{i \in I} \mu_{\bar{B}}(B_i).P(A_j | B_i).P(B_i)}{P(A_j)}$$

En deuxième lieu, si nous considérons l'événement flou \bar{A} de la variable A_j qui présente l'état « non trop bas », nous aurons l'équation bayésienne floue :

$$P(B_i | \bar{A}) = \frac{\sum_{j \in J} \mu_{\bar{A}}(A_j).P(A_j | B_i).P(B_i)}{P(\bar{A})}$$

En dernier lieu, si nous considérons les deux événements flous \bar{A} et \bar{B} , nous obtenons comme équation bayésienne floue :

$$P(\bar{B} | \bar{A}) = \frac{\sum_{i \in I} \sum_{j \in J} \mu_{\bar{B}}(B_i). \mu_{\bar{A}}(A_j).P(A_j | B_i).P(B_i)}{P(\bar{A})}$$

3.2 Deuxième approche de Réseau Bayésien Flou

3.2.1 Structure du Réseau Bayésien Flou

Un RBF $G = \langle \eta, \theta \rangle$ est caractérisé par une structure η et des paramètres θ . Ainsi, la spécification des RBFs et des RBs est la même, cependant, la seule différence est au niveau de la propagation de croyance.

3.2.2 Propagation de croyance dans le Réseau Bayésien Flou

La propagation de croyance consiste à calculer les mises à jour des probabilités des variables du RB.

3.2.2.1 Quelques notions de base

Un état flou est composé d'un ou plusieurs composants, pour chacun on associe un degré d'appartenance $\mu : \{x_1, x_2, x_3\}_{\mu_i}$ est un composant avec x_1, x_2, x_3 sont des probabilités flous et μ_i est le degré d'appartenance de x_1, x_2, x_3 à un sous-ensemble flou.

Par exemple, $S = [hi_{0.7}, mid_{0.3}]$ est une variable ayant deux composants hi et mid de degré d'appartenance respectivement 0.7 et 0.3.

Supposons que les valeurs floues que peuvent prendre une variable sont : lo, mid et hi . Ainsi, le composant $\{0.2, 0.1, 0.7\}_{0.2}$ peut être soit $lo_{0.2}$, soit $mid_{0.2}$, soit $hi_{0.2}$ avec les probabilités respectivement 0.2, 0.1, et 0.7.

3.2.2.2 Quelques suppositions

Cette approche part de quelques suppositions qui facilitent la présentation dans l'espace disponible. En outre, ces propositions sont raisonnables et ne limitent pas l'utilité générale des RBFs.

La première supposition est que la $\sum \mu = 1$. Cette supposition peut simplifier la combinaison entre les états des variables flous. En effet, si par exemple le degré d'appartenance de hi est 0.5 alors les degrés d'appartenances de lo et mid appartiennent à l'intervalle $[0, 0.5]$.

La deuxième supposition, concernant le RBF durant la propagation de croyance, est l'indépendance des composants. En faite, si une variable a seulement un seul parent alors il aura le même nombre de composants que son parent avec les mêmes degrés d'appartenances (μ). Par exemple, le fils de la variable $S = [hi_{0.7}, mid_{0.3}]$ aura deux composants l'une avec

$\mu = 0.7$ et l'autre avec $\mu = 0.3$. Cependant, si la variable possède plusieurs parents alors les composants des parents seront combinés avant la propagation.

Dans la suite nous présentons la propagation de croyance dans le cas d'un seul parent et des parents multiples.

3.2.2.3 Propagation de croyance dans le cas d'un seul parent

Nous présentons dans la Figure 10 un exemple de RBF notamment les tables de probabilités de chaque nœud. Ainsi, le Tableau 8 présente la distribution de probabilités *à priori* de A, le Tableau 9 désigne la distribution de probabilités conditionnelles de B, le Tableau 10 montre la distribution de probabilités conditionnelles de C, et enfin le Tableau 11 exprime la distribution de probabilités conditionnelle de E.

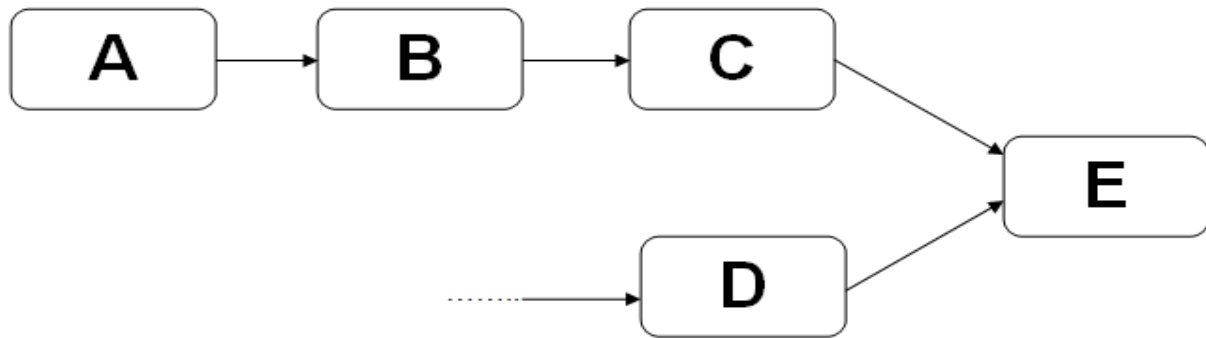


Figure 10. Exemple de RBF

Tableau 8. Distribution de probabilités *à priori* de A

A	A=lo	A=mid	A=hi
	0.7	0.1	0.2

Tableau 9. Distribution de probabilités conditionnelles de B

A→B	B=lo	B=mid	B=hi
A=lo	0.6	0.2	0.2
A=mid	0.1	0.1	0.8
A=hi	0.1	0.2	0.7

Tableau 10. Distribution de probabilités conditionnelles de C

B→C	C=lo	C=mid	C=hi
B=lo	0.1	0.1	0.8
B=mid	0.1	0.8	0.1
B=hi	0.7	0.2	0.1

Tableau 11. *Distribution de probabilités conditionnelles de E*

C, D→E		E=lo	E=mid	E=hi
C=lo	D=lo	0.6	0.2	0.2
C=lo	D=mid	0.1	0.1	0.8
C=lo	D=hi	0.1	0.1	0.8
C=mid	D=lo	0.6	0.2	0.2
C=mid	D=mid	0.1	0.6	0.3
C=mid	D=hi	0.1	0.6	0.3
C=hi	D=lo	0.1	0.2	0.7
C=hi	D=mid	0.1	0.2	0.7
C=hi	D=hi	0.8	0.1	0.1

Supposons que les observations dans notre RBF indique que $A = [\text{mid}_{0.2}, \text{hi}_{0.8}]$. Avec cette information nous pouvons calculer les distributions de probabilités floues de B et C.

Ainsi, puisque A est le seul parent de B, alors nous aurons $B = [\{0.1, 0.1, 0.8\}_{0.2}, \{0.1, 0.2, 0.7\}_{0.8}]$.

En outre, la distribution de probabilités floues de C est calculée de la même façon. Ainsi, nous aurons $C = [\alpha_{0.2}, \beta_{0.8}]$ avec α et β seront calculés en utilisant la propagation standard de RB et en se basant sur la distribution de probabilités conditionnel de C :

$$P(C|B=\text{lo}) = \{0.1, 0.1, 0.8\} ;$$

$$P(C|B=\text{mid}) = \{0.1, 0.8, 0.1\} ;$$

$$\text{Et } P(C|B=\text{hi}) = \{0.7, 0.2, 0.1\}.$$

$$\text{Par la suite, } \alpha = \{0.1, 0.1, 0.8\} * 0.1 + \{0.1, 0.8, 0.1\} * 0.1 + \{0.7, 0.2, 0.1\} * 0.8$$

$$= \{0.58, 0.25, 0.17\}.$$

$$\text{Et } \beta = \{0.1, 0.1, 0.8\} * 0.1 + \{0.1, 0.8, 0.1\} * 0.2 + \{0.7, 0.2, 0.1\} * 0.7$$

$$= \{0.52, 0.31, 0.17\}.$$

$$\text{Finalement, } C = [\{0.58, 0.25, 0.17\}_{0.2}, \{0.52, 0.31, 0.17\}_{0.8}].$$

3.2.2.4 Propagation de croyance dans le cas des parents multiples

Dans le paragraphe précédent nous avons détaillé la propagation de croyance dans le RBF lorsque la variable possède un seul parent.

Passons maintenant au cas des parents multiples et calculons la distribution de probabilités floues de E, en supposant que :

$$D = [\{0.45, 0.30, 0.25\}_{0.3}, \{0.1, 0.8, 0.1\}_{0.7}].$$

Et $C = [\{0.58, 0.25, 0.17\}_{0.2}, \{0.52, 0.31, 0.17\}_{0.8}]$ (d'après les calculs développés dans le paragraphe précédent). En combinant les composants des parents C et D, nous obtenons $E = [\alpha_{0.06}, \beta_{0.14}, \gamma_{0.24}, \delta_{0.56}]$.

Avec

$$\begin{aligned} \alpha &= 0.58 * 0.45 * \{0.6, 0.2, 0.2\} + 0.58 * 0.30 * \{0.1, 0.1, 0.8\} + 0.58 * 0.25 * \{0.1, 0.1, 0.8\} \\ &+ 0.25 * 0.45 * \{0.6, 0.2, 0.2\} + 0.25 * 0.3 * \{0.1, 0.6, 0.3\} + 0.25 * 0.25 * \{0.1, 0.6, 0.3\} + \\ &0.17 * 0.45 * \{0.1, 0.2, 0.7\} + 0.17 * 0.3 * \{0.1, 0.2, 0.7\} + 0.17 * 0.25 * \{0.8, 0.1, 0.1\} \\ &= \{0.3165, 0.2189, 0.4647\} \end{aligned}$$

De la même manière nous obtenons β , γ , et δ .

4 Avantage des Réseaux Bayésiens Flous

Les avantages des réseaux bayésiens flous sont très variés. Avant de les présenter, nous citons tout d'abord quelques avantages des réseaux bayésiens et de la logique floue.

Comme nous l'avons mentionné, d'un côté les réseaux bayésiens présentent une rationalité statistique et une rigoureuse capacité pour l'inférence causale, de l'autre côté, la logique floue est robuste face aux données bruitées.

En outre, l'utilisation de terme linguistique facilite la compréhension humaine du modèle. En addition, cette technique est particulièrement utile quand les données sont insuffisantes pour formuler un modèle précis. En effet, elle présente une connaissance supplémentaire venant des états flous pour aider à l'interprétation humaine et la conception du système flou.

Quant aux réseaux bayésiens flous, ils comportent la richesse des réseaux bayésiens classiques et de la logique floue.

En faite, l'analyse théorique montre que l'utilisation des variables flous est plus efficace que celle des variables discrètes ou continues. Conséquemment, pour résoudre des problèmes complexes on a eu recours plutôt au RBF qu'au RB classique.

5 Recherches effectuées dans les réseaux bayésiens flous

Bien que les réseaux bayésiens flous soient très efficaces surtout face aux problèmes complexes, jusqu'à présent il y a eu très peu de recherches dans cette nouvelle théorie.

Les RBFs ont été appliqués en médecine, en sécurité, en diagnostic... L'article [Lin et al., 2006] discute l'application des RBFs pour la prévision du degré de l'astrocytome¹³. En faite, cette application réalise une analyse dans la prévision du niveau malveillant de l'astrocytome et fournit un outil d'aide au diagnostic pour les jeunes docteurs.

Cependant, d'après [Lin et al., 2006], une exactitude de 81.67 % a été obtenue par 60 échantillons de tests, ceci satisfait l'exigence des neurologistes.

Egalement, l'article [Ren et al., 2005] propose un modèle de RBF qui traite les risques et l'analyse de sécurité marine (par exemple les accidents causés par les roches).

Parmi les recherches exploitant ces modèles, une étude de diagnostic de défauts d'impression à distance utilisant les RBFs illustré dans [Qiu et al., 2001]. Ainsi, les auteurs ont affirmé que les valeurs de confiances résultantes pour le diagnostic en utilisant le RBF sont plus précises que les valeurs issues de l'utilisation d'un RB classique ou la LF.

En outre, d'autres recherches [Park et al., 2006] ont proposé un système de recommandation de musique utilisant les RBFs et ont affirmé que, d'après l'analyse du processus de recommandation et la comparaison des résultats issus de l'utilisation des RBs classiques et celle des RBFs, la satisfaction des utilisateurs a augmenté.

6 Notre approche pour l'inférence dans les Réseaux bayésiens flous

6.1 Présentation de notre algorithme

Les observations que nous propageons sont de nature floues, ceci n'est pas le cas dans les approches classiques d'inférence dans les RBs. Pour développer cet algorithme nous avons utilisé l'algorithme arbre de jonction combiné avec soft évidence modifié.

6.1.1 Arbre de jonction

Cet algorithme est applicable dans tous types de réseaux, arbre ou non arbre. En premier lieu il transforme le graphe en un arbre de jonction, puis il initialise ses potentiels, ensuite il utilise

¹³ Tumeur du système nerveux central (cervelet, cerveau, moelle épinière plus rarement), ou tout simplement, c'est le cancer du cerveau.

la méthode message passing pour la propagation des messages et le calcul des probabilités *a posteriori*.

En faite, l'algorithme se comporte de la façon suivante [Huang et al., 1996] :

- La phase de construction (ou transformation du graphe) : elle nécessite un ensemble de sous-étapes permettant de transformer le graphe initial en un arbre de jonction, dont les nœuds sont des cliques (regroupement) de nœuds du graphe initial. Cette transformation est nécessaire, d'une part pour éliminer les boucles du graphe, et d'autre part, pour obtenir un graphe plus efficace quant au temps de calcul nécessaire à l'inférence, mais qui reste équivalent au niveau de la distribution de probabilité représentée. Cette transformation se fait en trois étapes :
 - ✓ la moralisation du graphe.
 - ✓ la triangulation du graphe et l'extraction des cliques qui formeront les nœuds du futur arbre.
 - ✓ la création d'un arbre de recouvrement minimal, appelé arbre de jonction.
- La phase d'initialisation : il s'agit d'initialiser les potentiels des cliques et séparateurs.
- La phase de propagation : il s'agit de la phase de calcul probabiliste à proprement parler où les nouvelles informations concernant une ou plusieurs variables sont propagées dans l'ensemble du réseau, de manière à mettre à jour l'ensemble des distributions de probabilités du réseau. Ceci se fait en passant des messages contenant une information de mise à jour entre les cliques de l'arbre de jonction précédemment construit.
- La phase de marginalisation : l'arbre de jonction contiendra la distribution de probabilité sachant les nouvelles informations, c'est-à-dire $p(\mathbf{U} | \mathbf{e})$ où \mathbf{U} représente l'ensemble des variables du réseau bayésien et \mathbf{e} l'ensemble des nouvelles informations sur les variables.

6.1.2 Soft evidence

L'observation classique dans un RB touche seulement l'un des états du nœud observé, autrement dit, si on a une évidence dans un nœud X , seulement l'un de ses états qui sera observé. Ainsi, on associe la valeur 1 à cet état et 0 pour les autres.

Cependant, on ne peut pas représenter le cas où notre évidence n'est pas exacte : c'est-à-dire si notre observation est ambiguë et touche en même temps plusieurs états selon des pourcentages bien définis.

Pour surmonter cette limite, [Tu et al., 2004] développe le concept de « soft evidence » qui permet de mettre en évidence l'ambiguïté au niveau d'une observation.

Cependant, et vu l'importance de ce nouveau concept, Smile¹⁴ a incorporé récemment « soft evidence » que l'on trouve intégré dans les nouvelles versions de Genie¹⁵.

Toutefois, l'inférence probabiliste aura, après une observation d'une évidence (e), quatre tâches [Tu et al., 2004] :

1. Mise à jour de croyance, $P(X=x|e)$
2. Trouver l'explication la plus probable

Dans le cas où le réseau contient des nœuds utilités l'objet de l'inférence est aussi de :

3. Maximiser une probabilité postérieure d'état du réseau
4. Maximiser l'utilité attendue

L'évidence classique pour un nœud particulier est l'observation de l'un de ses états, appelée encore évidence dure ou simplement observation. Cependant, « soft evidence » est le type le plus général de l'évidence qui introduit l'incertitude de l'observation.

Ainsi, ce concept consiste à propager des valeurs quantifiées (voir Figure 11) pour chaque état du nœud observé. On peut résumer cet algorithme dans les trois étapes suivantes :

1. *Calculer les valeurs quantifiées q_i pour chaque état i du nœud observé.*
2. *Pour chaque état i du nœud observé, on fait une inférence classique en observant l'état i .*
3. *Pour chaque nœud N du RB*

Pour chaque état j du nœud N

*Etat $j = \sum \text{valeur de l'état } j \text{ issu de l'inférence } i * q_i$*

Fin Pour

Fin Pour

Pour mieux expliquer ce concept, on peut illustrer un exemple (cf. Figure 11) détaillé dans [Tu et al., 2004].

¹⁴ SMILE (Structural Modeling, Inference, and Learning Engine) est une bibliothèque de classes C++ implémentant des modèles graphiques probabilistes tels que les réseaux bayésiens et les diagrammes d'influence.

¹⁵ GeNIe (Graphical Network Interface) est une interface graphique permettant de manipuler la plupart des fonctionnalités offertes par SMILE.

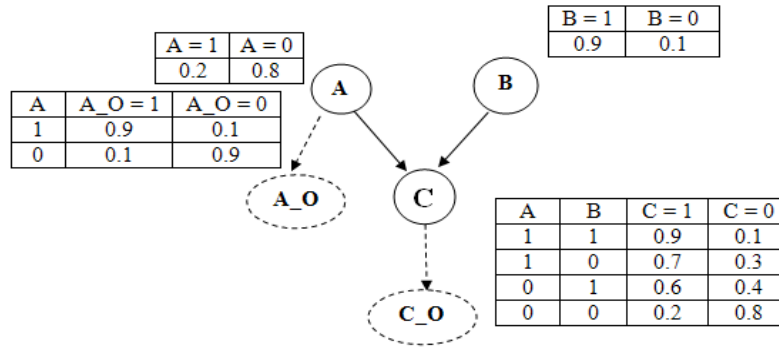


Figure 11. Exemple pour Mise à jour de Croyance avec soft evidence

Ainsi, si l'observation est effectuée sur le nœud A avec les pourcentages 90% pour l'état 1 et 10% pour l'état 2 alors l'algorithme se déroule comme suit :

1. Calculer les valeurs quantifiées des états de A :

$$\begin{aligned}
 Q(A=1) &= Q(A=1|A_o=1) \\
 &= [P(A_o=1|A=1) * P(A=1)] / [P(A_o=1|A=1) * P(A=1) + P(A_o=1|A=0) * P(A=0)] \\
 &= [0.9 * 0.2] / [0.9*0.2 + 0.1 * 0.8] \\
 &= 0.6923. \\
 Q(A=0) &= Q(A=0|A_o=1) \\
 &= [P(A_o=1|A=0) * P(A=0)] / [P(A_o=1|A=1) * P(A=1) + P(A_o=1|A=0) * P(A=0)] \\
 &= [0.1 * 0.8] / [0.9*0.2 + 0.1 * 0.8] \\
 &= 0.3077.
 \end{aligned}$$

On remarque que $Q(A=1) + Q(A=0) = 1$.

2. En premier lieu on réalise la 1^{ère} inférence classique en prenant comme observation l'état ($A = 1$) puis on multiplie tous les états du RB par $Q(A=1)$. En deuxième lieu on effectue la 2^{ème} inférence classique en prenant comme observation l'état ($A = 0$) puis on multiplie tous les états du RB par $Q(A=0)$.
3. Enfin, pour chaque état on fait la somme de toutes les valeurs qu'il a prises lors de l'étape 2.

6.1.3 Algorithme proposé

Comme on vient de signaler, notre approche d'inférence floue est une amélioration du concept « soft evidence ». Cependant, on peut résumer l'algorithme dans les deux étapes suivantes :

1. *Pour chaque état i du nœud observé, on fait une inférence classique en observant l'état i .*
2. *Pour chaque nœud N du RB*

Pour chaque état j de N

*Etat $j = \sum$ valeur de l'état j issu de l'inférence i * degré d'appartenance de la valeur observée à l'état i*

Fin Pour

Fin Pour

7 Conclusion

Au cours de ce troisième chapitre nous avons mis en évidence l'intérêt de la théorie de RBF, cependant, nous avons valorisé ses avantages, cité les applications qui l'utilisent, et détaillé les notions de base de cette nouvelle théorie en conception et en inférence. Nous avons proposé un algorithme d'inférence qui améliore les approches existantes et fournit un résultat fiable pour des observations floues.

CHAPITRE 4

Réalisation et test

"Les questions les plus importantes de la vie ne sont en fait, pour la plupart, que des problèmes de probabilités."

Pierre-Simon de Laplace (1749-1827).

1 Introduction

Au cours d'un projet de fin d'étude (2007-2008), nous avons réalisé un éditeur graphique "Baysian Editor " pour la représentation des réseaux bayésiens. Ce travail a été réalisé suite à une recherche bibliographique sur le domaine ainsi qu'une étude sur les éditeurs graphiques existants.

Cependant, l'éditeur a été repris dans un autre projet afin d'ajouter un algorithme d'inférence pour assurer la propagation de l'information dans le réseau bayésien.

"Baysian Editor" est utilisé comme point de départ pour développer le module flou décrit dans le chapitre 3.

2 Editeur Réseau Bayésien Flou

2.1 Editeur Réseau Bayésien classique

2.1.1 Aspect graphique

Nous décrivons en premier lieu quelques principales fonctionnalités de notre éditeur (cf. Figure 12).

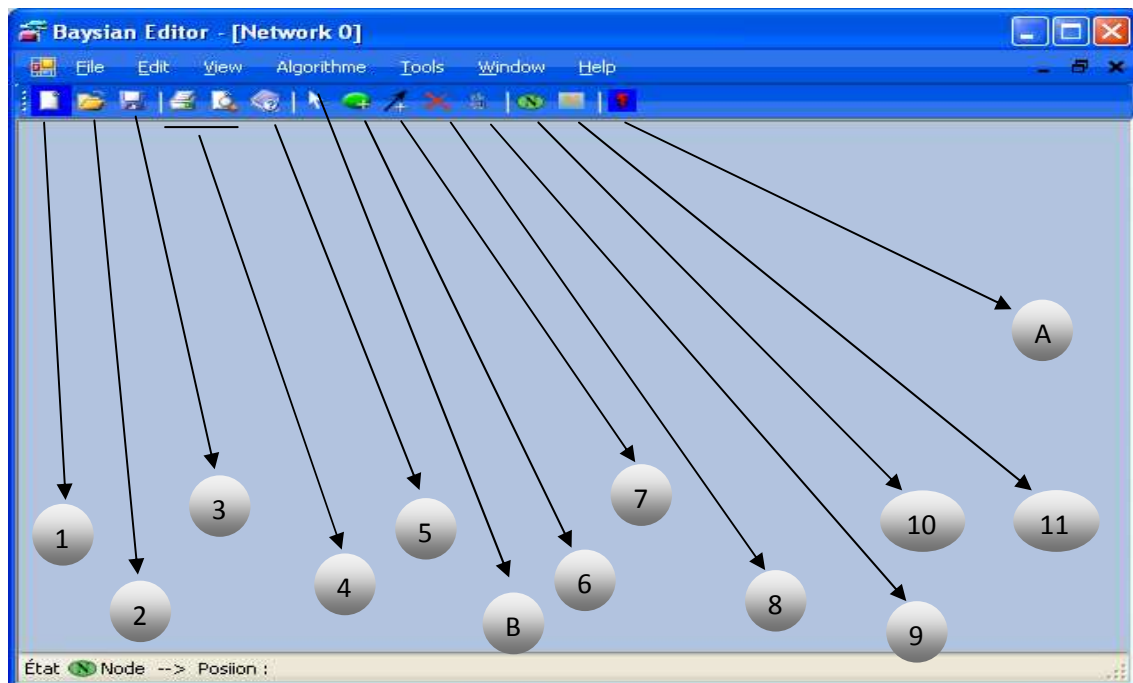


Figure 12. *Présentation générale de l'éditeur*

1 : ajouter un nouveau réseau, 2 : ouvrir un réseau, 3 : enregistrer un réseau, 4 : imprimer/aperçu réseau, 5 : aide, 6 : ajouter un nœud, 7 : relier deux nœuds à l'aide d'un

arc, 8 : supprimer un nœud, 9 : ajouter un commentaire, 10 : changer la forme d'un nœud de la forme rectangulaire à la forme ellipse, 11 : changer la forme d'un nœud de la forme ellipse à la forme rectangulaire, A : choisir l'algorithme d'inférence, B : mise à jour les probabilités du réseau selon l'algorithme choisi.

2.1.2 Algorithme d'inférence

Nous présentons maintenant le déroulement des étapes de l'algorithme arbre de jonction en commençant par la création du graphe sur lequel on va appliquer notre algorithme (cf. Figure 13).

Une fois qu'on a fini la création du graphe avec les outils disponibles dans notre éditeur, on fait entrer les probabilités de chaque nœud. Ces probabilités sont décrites à l'aide d'un expert ou à partir d'une source d'information. Ensuite, on sélectionne nos évidences si elles existent, comme la Figure 13 nous montre, puis on choisit l'algorithme d'inférence par un simple clique sur **Algorithme** puis sur **Arbre de Jonction**.

À ce stade, la page dans laquelle l'algorithme va se dérouler s'ouvre à côté de notre réseau ouvert afin de permettre de voir les changements apportés sur notre graphe par l'algorithme.

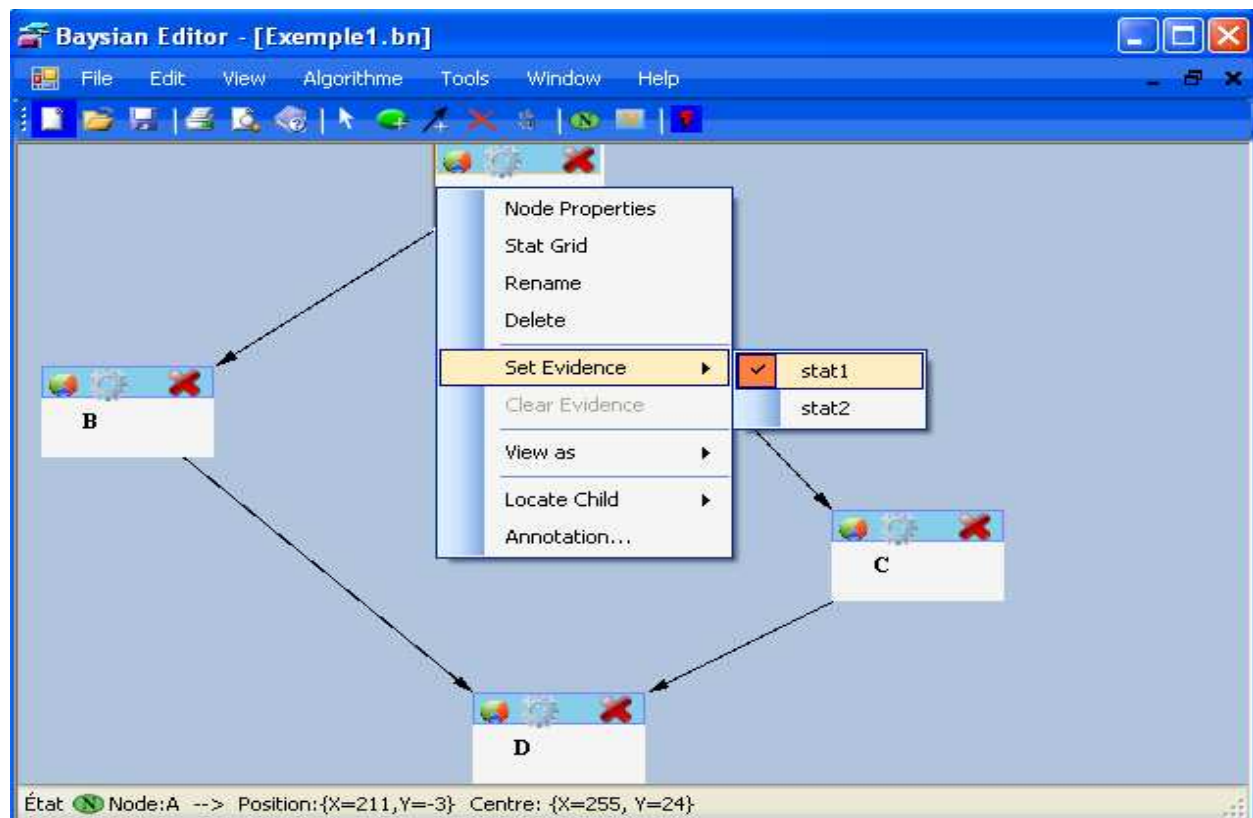


Figure 13. Graphe initial

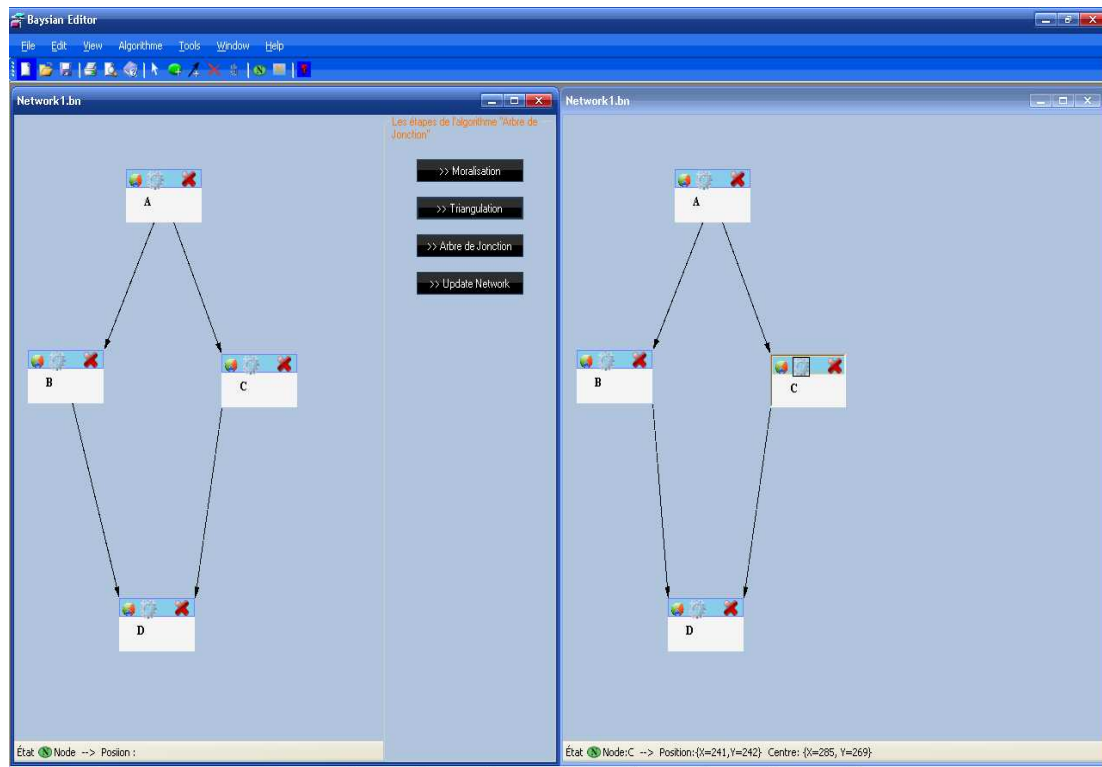


Figure 14. Réseau après choix de l'algorithme

Nous remarquons maintenant deux réseaux ouverts, l'un contient le graphe initial et l'autre où va se dérouler l'algorithme (cf. Figure 14).

Quatre boutons s'affichent dans la fenêtre de gauche, chacun décrit une étape de l'algorithme Arbre de jonction.

Pour réaliser la première phase on clique sur **>> Moralisation**.

Pour réaliser la deuxième phase on clique sur **>> Triangulation**, mais dans ce cas le graphe est triangulé, donc le graphe reste lui-même.

Pour réaliser la troisième phase on clique sur **>> Arbre de Jonction**.

On remarque :

- ✓ la construction de l'arbre de jonction qui est constituée de cliques et de séparateurs.
- ✓ l'initialisation des potentiels des cliques et des séparateurs.

Afin de voir les résultats de cet algorithme, c'est-à-dire pour appliquer la dernière phase sur le graphe, on clique sur **>> Marginalisation**, puis se déplacer vers le nœud choisi pour afficher ses probabilités *a posteriori* (cf. Figure 15).

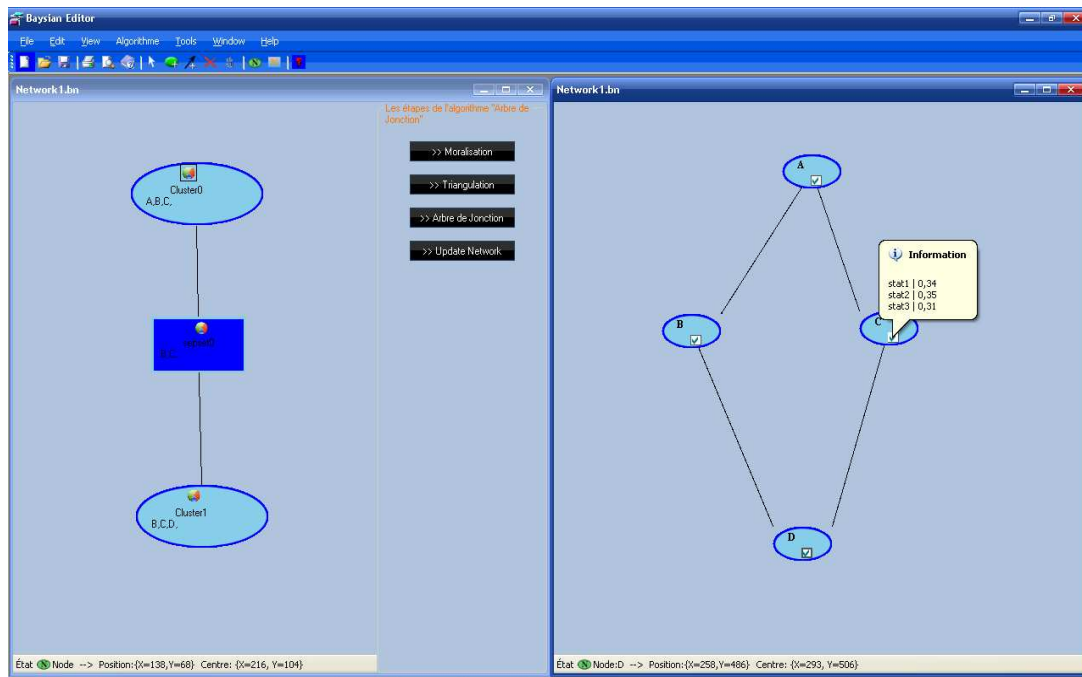



Figure 15. Résultat final

Nous avons montré comment se déroule l'algorithme phase par phase, nous indiquons que notre éditeur permet aussi d'appliquer l'algorithme sur le graphe d'un seul coup en cliquant sur le bouton .

2.2 Représentation des données floues

Afin d'étendre l'utilisation de cet éditeur au RBF, nous pouvons accéder au module de représentation des variables (cf. Figure 16).

Cette interface inclut un tableau contenant les différentes données liées à la variable floue tel que son nom, sa fonction, l'intervalle de définition de cette fonction et la couleur qui sera choisie dans sa représentation graphique.

Dans ce tableau, nous pouvons ajouter, insérer ou supprimer des lignes à travers les boutons "Ajouter", "Insérer" et "Supprimer".

À propos des fonctions de la variable floue, elles peuvent inclure plusieurs types de fonctions mathématiques notamment les nombres négatifs, les parenthèses, les racines carrées, les puissances, sinus, cosinus, tangente, +, -, * et /.

Pour faciliter la tâche à l'utilisateur de notre éditeur, le bouton "Utilisation" contient les différents détails concernant les fonctions que nous venons de citer.

Pour mieux visualiser les sous-ensembles flous, nous disposons de la fonctionnalité "Représentation graphique" accessible par les trois boutons : "Range", "Ecran" et "Graphe". Ainsi, le premier bouton "Range" nous permet de fixer les différentes caractéristiques de l'échelle de graduation afin d'obtenir une meilleure représentation graphique. Les deux autres

boutons permettent l'affichage des courbes des fonctions saisies dans le tableau.

Finalement, lors d'une évidence, nous pouvons saisir la valeur observée dans le champ "Valeur Observée" et cliquer sur le bouton "Calculer les degrés d'appartenance" pour obtenir les degrés d'appartenance de chaque variable floue dans le deuxième tableau.

3 Exemples d'inférence flou

Nous allons à présent étudier de plus près comment fonctionne l'inférence floue dans notre éditeur. Pour cela, nous allons illustrer deux exemples de RBF, un exemple simple pour mieux visualiser les aspects général de notre éditeur, notamment les caractéristiques des variables flous et l'inférence floue, puis un deuxième exemple plus détaillé concernant le diagnostic du cancer de la prostate.

3.1 Exemple simple

Dans cet exemple, détaillé dans [Heckerman, 2004], nous avons cinq variables : Gaz, Fraude, Bijou, Age et Sexe, comme indique la Figure 17.

Ainsi, supposons que la variable « Age » est une variable floue qui comporte trois sous-ensembles flous « Jeune », « Adulte » et « Vieux » caractérisés par des fonctions bien déterminées qu'on peut saisir dans le tableau spécifique de la variable floue (cf. Figure 16).

The screenshot shows the 'Propriety' application window with the 'Variable Floue' tab selected. The 'Valeur Observée' field contains the value 23. Below it, there are buttons for 'Utilisation', 'Ajouter', 'Insérer', and 'Supprimer'. A table lists fuzzy functions for the 'Age' variable:

Variable Floue	Fonction	Min	Max	Couleur
Jeune	1	0	20	rouge
Jeune	$n(0.1)^x + 3$	20	30	rouge
Adulte	$0.1 \cdot x - 2$	20	30	vert
Adulte	1	30	40	vert
Adulte	$n(0.1)^x + 5$	40	50	vert


To the right of the table is a 'Mode Graphique' section with buttons for 'Range', 'Ecran', and 'Graphe'. At the bottom, there is a table showing the results of the calculation:

Variable Floue	Résultat
Jeune	0,7
Adulte	0,3

At the bottom right of the window are 'OK' and 'Cancel' buttons.

Figure 16. Représentation de la variable floue Age

Supposons que l'âge observé est 23 ans, on peut saisir cette valeur directement dans le champ spécifique, puis on clique sur le bouton pour obtenir les degrés d'appartenance de la valeur observée à chacun des sous-ensembles flous comme indique la Figure 16.

Ensuite, on clique sur le bouton pour enregistrer les données puis sur le bouton  afin de réaliser l'inférence.

Finalement, on obtient les résultats finals fournis par cet algorithme. Ainsi, on peut se déplacer vers le nœud choisi pour afficher ses probabilités *à posteriori* (cf. Figure 17).

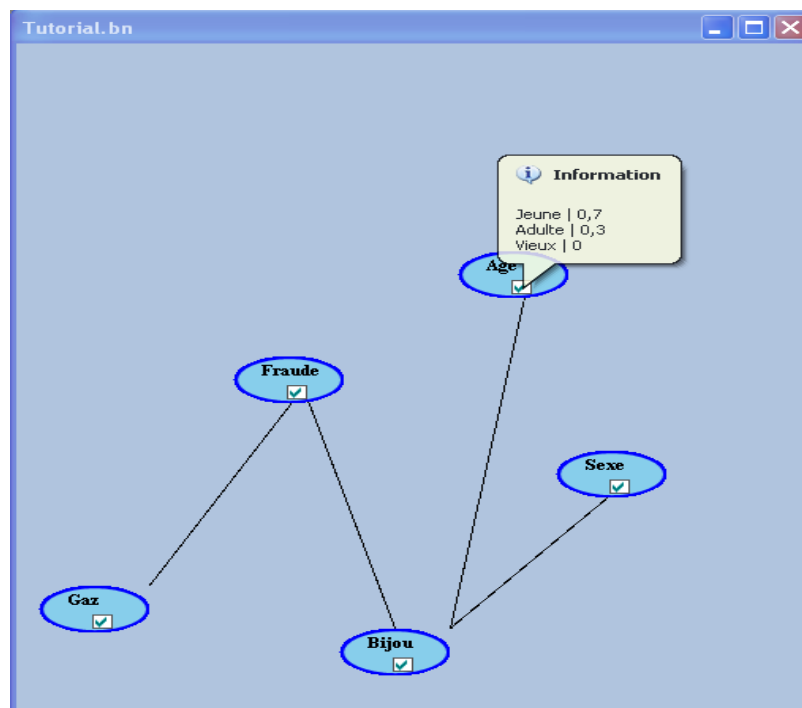


Figure 17. Résultat final

3.2 Exemple de diagnostic du cancer de la prostate

3.2.1 Cancer de la prostate

Le cancer de la prostate est un cancer fréquent touchant la prostate et donc exclusivement le sexe masculin. Ce cancer se développe à partir des tissus de la prostate, une glande de l'appareil reproducteur masculin, quand des cellules y mutent pour se multiplier de façon incontrôlée.

Celles-ci peuvent s'étendre en migrant de la prostate jusqu'à d'autres parties du corps.

Le cancer de la prostate peut provoquer des douleurs, une difficulté d'uriner, un dysfonctionnement érectile et d'autres symptômes.

Le traitement se fait par chirurgie, radiothérapie¹⁶, thérapie hormonale et parfois la chimiothérapie¹⁷, ou en combinant plusieurs de ces méthodes [5].

Dans la plupart des cas, le cancer de la prostate est asymptomatique, c'est-à-dire qu'il est découvert, alors qu'il n'entraîne aucune manifestation lui étant propre. Il est le plus souvent découvert :

- ✓ Lors d'analyses sanguines incluant l'étude de l'APS (Antigène Prostatique Spécifique), communément connu en terme anglophone par PSA (Antigène Prostatique Spécifique), dont la valeur prédictive et l'utilisation, sans bénéfice prouvé en termes de santé publique, ont récemment été remises en cause. L'APS est une protéine normalement sécrétée par les cellules prostatiques, mais une cellule cancéreuse en sécrète 10 fois plus qu'une cellule normale. Cette propriété a suscité de nombreux espoirs en termes de dépistage.

Le taux sanguin de l'APS peut toutefois être augmenté par de très nombreux autres facteurs (le volume prostatique, les infections et/ou inflammations, les contraintes mécaniques (toucher rectal, autre...) ou diminué par certains traitements de l'hypertrophie bénigne.

- ✓ Lors d'un toucher rectal réalisé à titre systématique, ou en raison de symptômes liés à une autre maladie (en particulier l'hypertrophie bénigne de la prostate).
- ✓ Fortuitement, sur des pièces de résection prostatique lors du traitement chirurgical de l'adénome prostatique.

Lorsqu'il est symptomatique, le cancer de la prostate est le plus souvent à un stade avancé. Il peut entraîner :

- ✓ une rétention aiguë d'urine,
- ✓ une hématurie¹⁸,
- ✓ une impuissance sexuelle,
- ✓ une altération de l'état général,

¹⁶ La radiothérapie est une méthode de traitement locorégional des cancers, utilisant des radiations pour détruire les cellules cancéreuses en bloquant leur capacité à se multiplier.

¹⁷ La chimiothérapie est l'usage de certaines substances chimiques pour traiter une maladie.

¹⁸ L'hématurie est un terme médical désignant la présence de sang dans les urines. En fait on dépiste la présence de globules rouges en quantité anormalement élevée.

- ✓ des douleurs et/ou le dysfonctionnement ou la défaillance d'autres organes liés à la présence de métastases¹⁹.

3.2.2 Problématique et solution

Actuellement, le cancer de la prostate est le cancer le plus fréquent et la deuxième cause de mortalité par cancer chez les hommes. Pour suivre l'évolution des patients traités pour ce cancer, les médecins utilisent depuis plusieurs années le dosage sanguin de l'APS. En faite, l'APS est le meilleur indicateur de l'évolution du cancer de la prostate. Une augmentation de son taux dans le temps signale une progression de la maladie.

Le dosage de l'APS est donc essentiel pour suivre tous les patients qui ont subi un traitement pour un cancer de la prostate ou qui ont opté pour l'observation en présence d'une tumeur localisée.

Cependant, les seuils de significativité sont difficiles à établir. Il est admis, toutefois, qu'un taux de l'APS (APS libre / APS total) compris entre 10 et 20 est douteux, mais qu'il est nettement significatif au-delà.

Ainsi, une valeur du taux de l'APS supérieure à 20 oriente le diagnostic vers un adénome de la prostate (tumeur bénigne de la prostate) qui ne nécessite pas de démarrer d'autres explorations qui sont invasives.

D'un autre côté, une valeur de l'APS inférieur à 10 oriente le diagnostic vers un cancer de la prostate qui sera confirmé et identifié par une biopsie de la prostate (exploration invasive et traumatique).

Cependant, si la valeur du taux de l'APS est située entre 10 et 20, le médecin aura besoin d'un pourcentage qui l'oriente vers l'un des chemins en s'appuyant aussi sur d'autres paramètres diagnostic. C'est un vrai flou pour le médecin.

Sur le plan pratique, le jugement d'un taux de l'APS en faveur d'un adénome et d'un cancer, est une valeur subjective que nous pouvons l'objectiver par un prototype issu de l'étude des statistiques des cas déjà vue (cf. Figure 18).

¹⁹ Une métastase est la croissance d'un organisme pathogène ou d'une cellule tumorale à distance du site initialement atteint.

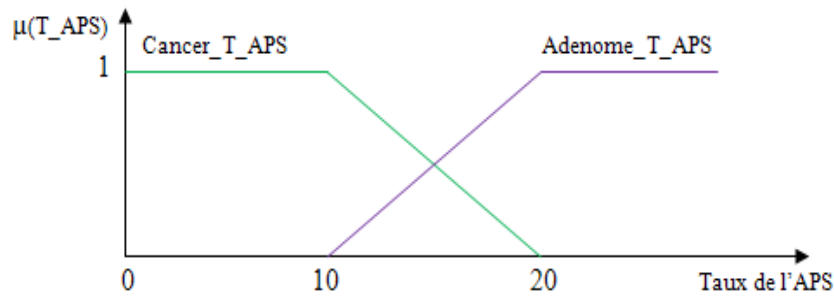


Figure 18. *Sous-ensembles flous de la variable Taux de l'APS*

3.2.3 Réseau Bayésien Flou pour le diagnostic du cancer de la prostate

Le cancer de la prostate occupe la deuxième place en termes de fréquence des cancers touchants l'homme. D'où l'intérêt d'un diagnostic précoce de ce cancer pour l'éradiquer à un stade primaire.

En effet, un diagnostic précoce permet une prise en charge rapide qui permet un traitement efficace pouvant amener à une guérison parfois complète.

Toutefois, il n'y a pas vraiment un tableau clinique clair et évident qui permet ce diagnostic mais ce sont des ensembles de critères qui donnent l'impression au médecin que son patient à une grande probabilité d'avoir ce cancer.

Récemment des travaux sur le diagnostic du cancer de la prostate ont vu le jour [Mahjoub et al., 2010]. Cette étude présente un outil d'aide à la décision pour le diagnostic, sauf que le modèle utilise des Tables de Probabilités Conditionnelles (TPCs) classiques mais ne traite pas les connaissances ambiguës.

Par ailleurs, cette démarche, qui reste toujours floue pour le clinicien, lui impose parfois de demander beaucoup d'examen complémentaires qui sont parfois de manière abusifs et aussi traumatique comme la biopsie de la prostate.

Notre objectif est d'éclaircir un chemin diagnostique qui guide le médecin vers l'une des trois voix : soit que son patient à une grande probabilité d'avoir ce cancer, soit par contre il s'agit plutôt d'une tumeur bénigne ou adénome. Ce sont les deux diagnostics qui occupent le plus un praticien. Soit le troisième chemin qu'il s'agit d'un autre diagnostic dont on s'intéresse pas beaucoup dans notre étude c'est la prostatite²⁰.

²⁰ Infection aiguë ou chronique de la prostate.

Pour atteindre un diagnostic il faut rassembler les critères cliniques tels que les plaintes du patient, les motifs de consultation et les critères biologiques comme la valeur de l'APS et le taux de l'APS.

Cependant, le médecin se confronte à un autre déficit qui concerne le taux de l'APS qui peut avoir une valeur non significative quand elle est comprise entre 10 et 20. Dans cet intervalle, le médecin se sent hésitant et ne peut pas être affirmative ni par le bien ni par le mal, d'où l'intérêt d'un pourcentage qui lui guide dans sa prise de décision.

Afin d'étudier ce cancer, nous proposons le modèle en réseau bayésien flou (cf. Figure 19). Ce modèle est synthétisé à partir de l'étude de ce cancer et en tenant compte des recommandations d'un ensemble de praticien.

Dans notre RBF (cf. Figure 19) nous avons pris en considération les paramètres les plus frappants guidant aux diagnostics et nous avons essayé de substituer tout ce qui est flou et tout ce qui est intuition par un degré d'appartenance qui représente une valeur plus réaliste et plus scientifique qu'une valeur ambiguë.

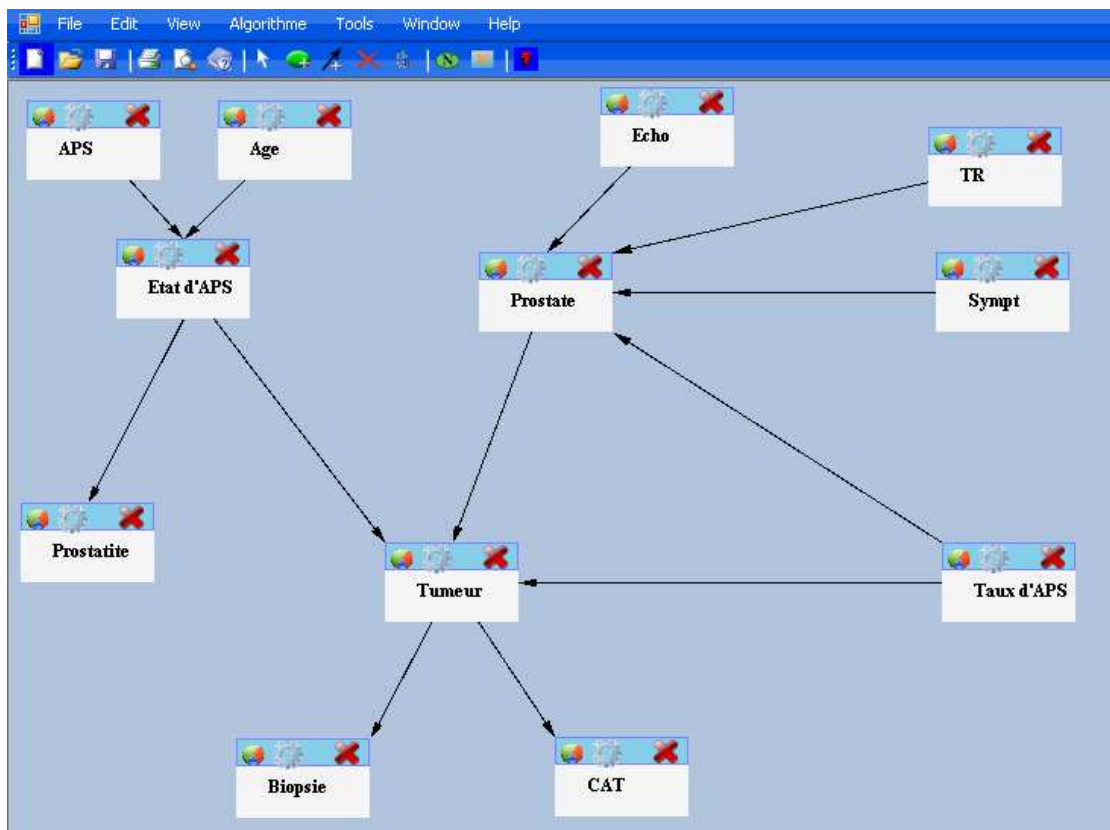


Figure 19. RBF pour le diagnostic du cancer de la prostate

Pour atteindre notre objectif nous avons suivi le raisonnement classique des médecins. En effet, ces derniers recueillent les informations selon un ordre bien précis avec des alternances parfois centripètes et parfois centrifuges selon la présence ou l'absence d'indice alarmant orientant vers un diagnostic ou vers un autre.

Ainsi, nous avons construit un RBF qui reflète la stratégie diagnostic du médecin où il trouve les réponses à la plupart des questions.

Toutefois, la question la plus précoce qu'il cherche est l'**âge** de son patient. En effet, la probabilité d'avoir une **tumeur de la prostate** augmente avec la progression de l'âge et oriente le médecin vers ce chemin. Par contre, plus son patient est jeune plus le médecin pense à une **prostatite**.

Cependant, pour avoir une vision plus claire, le docteur a recours au dosage de l'**APS** donnant ainsi une valeur qui ne peut être interprétable qu'en fonction de l'âge du patient. Nous parlons donc d'un **état de l'APS** que par tranche d'âge.

La valeur de l'APS, qui est interprétée par le laboratoire et récupérée par le médecin, peut être soit élevée soit normale. Ainsi, une valeur élevée de l'APS inquiète le clinicien sur l'**état de la prostate** de son patient qui dépend en fait d'autres paramètres.

En effet, ce sont les données de l'**échographie** qui se résument en quatre possibilités (états) :

- ✓ Prostate de volume normal et Aspect normal.
- ✓ Prostate de volume normal et Aspect anormal.
- ✓ Prostate de volume anormal et Aspect normal.
- ✓ Prostate de volume anormal et Aspect anormal.

Ces données seront confrontées, d'une part avec le **toucher rectal (TR)**, qui donne les informations sur le volume de la prostate et son aspect, et d'autre part avec la **symptomatologie** du patient qui a une relation directe avec l'état de la prostate, c'est-à-dire soit présence de symptomatologie urinaire qui confirme que la prostate a vraiment augmenté de volume, soit l'absence de symptomatologie qui n'élimine pas une modification de l'état de la prostate.

L'idéal pour le médecin est d'avoir un contexte bien établi où tous les données suivent le même sens et par la suite mènent à l'un des deux résultats : soit d'un adénome de la prostate, soit d'un cancer avec une grande probabilité.

En revanche, sur le plan pratique ce n'est pas toujours le cas et les données observées sont difficiles à interpréter et amènent nulle part. D'où il fait appel à une variable plus fiable pour ce type de diagnostic qui est le **taux de l'APS**.

Ce dernier est très significatif s'il appartient à l'un des intervalles (<10 ou >20) mais dans le cas échéant où il tombe sur l'intervalle de valeur flou (entre 10 et 20), un autre déficit confronte le médecin.

À ce stade, nous intervenons pour surmonter l'ambiguïté en représentant la variable taux de l'APS comme variable floue, tout en définissant les fonctions d'appartenances spécifiques des sous-ensembles flous (taux menant à un adénome et taux menant à un cancer) (cf. Figure 20).

Propriety:

General | Format | Variable Floue

Valeur Observé : 12 Calculer les degrés d'appartenances

Utilisation Ajouter Insérer Supprimer

	Variable Floue	Fonction	Min	Max	Couleur
▶	Taux menant à un cancer	1	0	10	vert
	Taux menant à un cancer	$n(0,1)^*x+2$	10	20	vert
	Taux menant à un adénome	$0,1^*x-1$	10	20	violet
	Taux menant à un adénome	1	20	40	violet

Mode Graphique

Range Ecran Graphe

	Variable Floue	Résultat
	Taux menant à un cancer	0,8
▶	Taux menant à un adénome	0,2

OK Cancel

Figure 20. Sous-ensembles flous de la variable Taux de l'APS

Ainsi, cette variable floue participe à établir un pourcentage décisif à la prise en charge ; c'est le degré d'appartenance, issu de la projection de la valeur observée du taux de l'APS sur les sous-ensembles flous (cf. Figure 21) ou encore obtenu en cliquant sur **Calculer les degrés d'appartenances** dans notre éditeur (cf. Figure 20), qui nous aide à connaître s'il s'agit d'une tumeur ou d'un adénome qui sont en faite les deux états de la variable **Tumeur**.

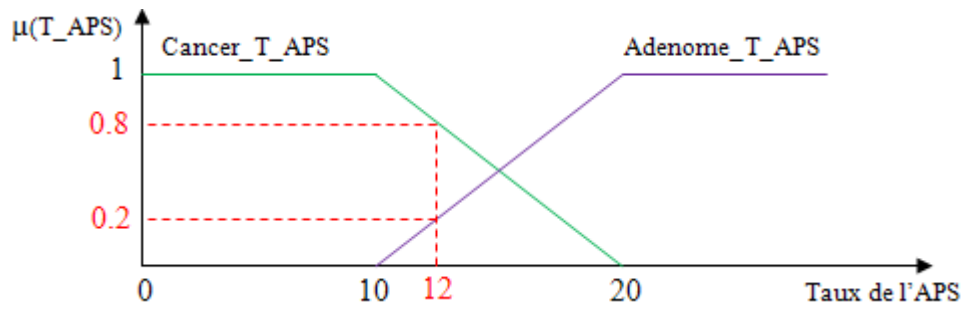


Figure 21. Degré d'appartenance aux sous-ensembles flous de la variable Taux de l'APS

Si encore la problématique n'est pas résolue, un dernier examen, qui est la **biopsie**, va participer à la décision.

En faite, une biopsie positive pour un cancer confirme le diagnostic à 100% et la conduite à tenir (**CAT**) sera dans ce sens. Par contre, une biopsie négative n'est pas concluante mais c'est un bon soulagement pour l'inquiétude du médecin et du patient surtout.

Ainsi, nous distinguons deux états pour la CAT : soit une biopsie positive qui nécessite un *traitement du cancer et suivi*, soit une biopsie négative qui exige un *traitement d'adénome et surveillance*.

4 Conclusion

Ce dernier chapitre présente l'éditeur graphique dédié à la présentation et l'inférence des RBFs. L'objectif de cet éditeur est d'illustrer le déroulement d'un diagnostic à partir de la modélisation en passant par l'inférence et en arrivant aux résultats.

La modélisation des nœuds flous fait appel notamment aux degrés d'appartenance aux sous-ensembles flous, l'inférence quant à elle se base sur l'algorithme JLO en tenant compte des degrés d'appartenances de la valeur observée. Ces deux aspects sont présentés à travers deux exemples. Le premier est un exemple classique de la littérature, le deuxième, plus détaillé, concernant le diagnostic du cancer de la prostate.

Conclusion générale

"L'homme raisonnable s'adapte au monde ; l'homme déraisonnable s'obstine à essayer d'adapter le monde à lui-même. Tout progrès dépend donc de l'homme déraisonnable."

George Bernard Shaw (1856-1950).

Tout au long de ce projet de recherche, nous sommes passés par plusieurs étapes. Une première phase de recherche bibliographique concernant les techniques de diagnostic nous a permis de constater que malgré leurs diversités, leurs efficacités restent en lieu étroit avec la nature du système à étudier.

Ainsi, cette première phase nous a permis d'identifier la technique des réseaux bayésiens en tant que outil de diagnostic très intéressant. Cependant, l'une des limites majeures que rencontrent les utilisateurs de ce type de réseaux, est d'exprimer les connaissances non précises surtout si plusieurs informations sont ambiguës. Étant donné que, nous ne pouvons pas exprimer toutes sortes de connaissance dans un langage précis.

Après avoir détaillé les deux techniques à savoir réseau bayésien et logique floue, dans un deuxième temps, nous avons pu démontrer l'apport que peut offrir le panachage de ces deux techniques qui a conduit à la naissance de la théorie de réseau bayésien flou.

Suite à cette étude, nous avons mis en évidence cette théorie en détaillant ses avantages, citer les applications qui l'utilisent et détaillant son fondement théorique.

En s'inspirant des résultats de toutes ces recherches, nous avons pu en premier lieu modéliser quelques systèmes par réseaux bayésiens flous, ce qui nous a permis de développer un module supplémentaire de représentation des variables flous à un éditeur graphique de représentation de RB en deuxième lieux, ainsi que de développer un algorithme exploitant la richesse de cette récente théorie, puis tester et comparer l'algorithme que nous avons développé par des exemples illustratifs et par rapport à d'autres travaux.

Bibliographie

- [Baldwin et al., 2003] Jim F. Baldwin, Enza Di Tomaso. « Inference and learning in fuzzy Bayesian networks ». *Fuzzy Systems. FUZZ' 03*, volume 1, pp. 630-635, 2003.
- [Davis, 1984] Randall Davis. « Diagnostic Reasoning Based on Structure and Behavior ». The Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A. Elsevier Science Publishers B.V., North-Holland, 1984.
- [Davis et al., 1988] Randall Davis, Walter C.Hamscher. Model-Based Reasoning : Troubleshooting. Massachudetts Institute Of Technology Artificial Intelligence Laboratory. July 1988.
- [Fahmy et al., 1997] Hany I. Fahmy, Christos Douligeris. « *Applications of Hybrid Fuzzy Expert Systems in Computer Networks Design* ». Dept. of Electrical and Computer Engineering University of Miami, 1997.
- [Fogelberg, 2008] Christopher Fogelberg. « *Belief Propagation in Fuzzy Bayesian Networks* ». *Combinations of Intelligent Methods and Approaches Workshop (CIMA)* in Patras, Greece, 2008.
- [Främling et al., 1992] Kary Främling, Albert Mathon. « Les réseaux de neurones comme outils d'aide à la décision floue ». Rapport de D.E.A, Ecole Nationale Supérieure des Mines de Saint-Etienne. Juillet 1992.
- [Friderich et al., 1992] Gerhard Friderich, Georg Gottlob, Wolfgang Nejdl. « Formalizing the Repair Process ». *Published in Proceedings of the European Conference on Artificial Intelligence*. Vienna, August 1992.
- [Friderich et al., 1992a] Gerhard Friderich, Wolfgang Nejdl. Choosing Observations and Actions in Model-Based Diagnosis/Repair Systems. *Published*

in Proceedings of the International Conference on Knowledge Representation and Reasoning. Cambridge, 1992.

- [Fung et al., 1989] Fung R., Chang K.C. *Weighting and integrating evidence for stochastic simulation in Bayesian Network*. In Proc. of UAI, vol. 5, page 209-219, New York Elsevier science publishing company, Inc 1989.
- [Fung et al., 1994] Fung R., Del Favero B. *Backward simulation in Bayesian network*. In R. Lopez de Mantaras and D. Poole Proc. of UAI, page 227-234, July 1994.
- [Gaag et al., 1994] LC.van der Gaag and M.Wessels. efficient multiple-disorder diagnosis by strategy focusing. Utrecht University No UU-CS-1994-23, 1994.
- [Garibaldi et al., 2000] Jonathan M. Garibaldi and Emmanuel C. Ifeakor. The Development of a Fuzzy Expert System for the Analysis of Umbilical Cord Blood. School of Electronic, Communication and Electrical Engineering, University of Plymouth, 2000.
- [Heckerman, 2004] David Heckerman. *Bayesian Networks for Data Mining. Data Mining and Knowledge Discovery*. Éditeur : Springer Netherlands. Novembre 2004.
- [Henrion, 1988] Henrion M. « *Propagating uncertainty in Bayesian network by probabilistic logic sampling* ». vol. 2, eds. J Lemmer and L. Kanal, 149-163. New York : Elsevier science, 1988.
- [Huang et al., 1996] Cecil Huang and Adnan Darwiche. *Inférence in belief networks: A procedural guide. International Journal of Approximate Reasoning*, 15(3) : 225-263, 1996.
- [Ketata et al., 2005] Raouf Ketata et Yosra Najar. « Générateur de Systèmes Experts Flous pour la Maintenance Industrielle ». Institut National des Sciences Appliquées et de Technologie, B.P. N° 676, Tunis, 2005.

- [Kleer et al., 1987] Johan de Kleer and Brian C. Williams. « Diagnosing Multiple Faults ». Elsevier Science Publishers B.V., North-Holland, 1987.
- [Kotchi et al., 2003] Camille Séka Kotchi, Véronique Delcroix, Sylvain Piechowiak. « Etude de la performance des algorithmes d'inférence dans les réseaux bayésiens ». Laboratoire d'Automatique, de Mécanique et d'Informatique industrielles et Humaines, Université de Valenciennes et du Hainaut Cambrésis, 2003.
- [Lauritzen et al., 1988] Lauritzen S. L. and Spiegelhalter D. J. « Local computation with probabilities on graphical structure and their applications to expert system ». Proc. of the royal statistical society, serie B., 50, 154-227, 1988.
- [Lee et al., 2000] Heung-Jae Lee, Deung-Yong Park, Bok-Shin Ahn, Young-Moon Park, Jong-Keun Park, and S.S.Venkata. *IEEE Transactions On Power Delivery*, VOL.15,NO.2, APRIL 2000.
- [Lin et al., 2006] Chun-Yi Lin, Jun-Xun Yin, Li-Hong Ma, and Jian-Yu Chen. « Fuzzy Bayesian Network-Based Inference in Predicting Astrocytoma Malignant Degree ». *Proceedings of the 6 th World Congress on Intelligent Control and Automation*, June 21-23, 2006, Dalian, China.
- [Lucas, 1997] Peter J.F. Lucas. « Editorial Model-based diagnosis in medicine. Department of Computer Science », Utrecht University, 13 February 1997.
- [Maalej, 2006] Mohamed-Amine Maalej. « Diagnostic multiple des systèmes complexes à base de réseaux bayésiens ». Thèse présentée à l'Université de Valenciennes et du Hainaut-Cambrésis en vue de l'obtention du grade de DOCTEUR, 30 novembre 2006.
- [Mahjoub et al., 2010] Mohamed Ali Mahjoub, Imed Boudriga. « Outil interactif d'aide à la décision basé sur les réseaux bayésiens : application au

- diagnostic du cancer de la prostate ». *5^{èmes} Journées Francophones sur les Réseaux Bayésiens*. Nante : France. 2010.
- [Naïm et al., 2004] Patrick Naïm, Pierre-Henri Wuillemin, Philippe Leray, Olivier Pourret, Anna Becker, « *Réseaux bayésiens* », Edition Ellipse, 2004.
- [Onisko et al., 1999] Agnieszka Onisko, Marek J. Druzdzel, and Hanna Wasyluk. « A Bayesian Network Model for Diagnosis of Liver Disorders ». *In Proceedings of the Eleventh Conference on Biocybernetics and Biomedical Engineering*, pages 842-846, Poland, December 2-4, 1999.
- [Onisko et al., 1998] Agnieszka Onisko, Marek J. Druzdzel and Hanna Wasyluk. « A Probabilistic Causal Model for Diagnosis of Liver Disorders ». *Intelligent Information Systems VII Proceedings of the Workshop held in Malbork*, Poland, June 15-19, 1998.
- [Pearl, 1988] J. Pearl. « Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference ». *Morgan Kaufmann Publishers, Inc.* San Mateo, CA. 1988
- [Pan et al., 1999] Heping Pan and Lin Liu. « Fuzzy Bayesian Networks - A General Formalism for Representation, Inference and Learning with Hybrid Bayesian Networks ». *Cooperative Research Centre for Sensor Signal and Information Processing, IEEE*, Australia, 1999.
- [Park et al., 2006] Han-Saem Park, Ji-Oh Yoo, and Sung-Bae Cho. « A Context-Aware Music Recommendation System Using Fuzzy Bayesian Networks with Utility Theory ». *FSKD 2006, Springer-Verlag Berlin Heidelberg*, 2006.
- [Pearl, 2000] J. Pearl. « Causality: Models, Reasoning and inference ». *Published by Cambridge University Press*, March 2000.
- [Qiu et al., 2001] Shijun Qiu, Alice M. Agogino, Shuang Song, Jialong Wu, Shankaran Sitarama, « A Fusion Of Bayesian and Fuzzy

- Analysis for Print Faults Diagnosis », *Proceedings of the ISCA 16th International Conference Computers and Their Applications*, March 28-30, 2001, Seattle, Washington, USA 2001.
- [Raymond et al., 1995] Raymond Christophe, Titli A. « Commande floue monovariante et multivariante selon l'expertise et par modélisation floue ». Thèse nouveau doctorat. Institut national des sciences appliquées de Toulouse, Toulouse, France, 1995.
- [Reiter, 1986] R. Reiter. « A theory of diagnosis from first principles ». *Technical Report, Computer science department university of Toronto*, 1986.
- [Reiter, 1987] Raymond Reiter. « A Theory of Diagnosis from First Principles ». *Elsevier Science Publishers B.V.*, North-Holland, 1987.
- [Ren et al., 2005] Jun Ren, Jin Wang, Ian Jenkinson. « Fuzzy Bayesian Modelling In Maritime Risk Analysis ». *Project is funded by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant No: GR/S85504/01 and GR/S85498/01*, 2005.
- [Saini et al., 2002] Harvinder S.Saini, Raj Kamal, A.N.Sharma. « Web Based Fuzzy Expert System for Integrated Pest Management in Soybean ». *International journal of Information Technology*, Vol.8, No.1, August 2002.
- [Shachter et al., 1989] Shachter R. D. and Peot M. A. « Simulation approaches to general probabilistic inference on belief networks ». In *Proc. of UAI*, vol. 5, pages 311-318, 1989.
- [Strasser, 1997] Mark Strasser. « The Development of a Fuzzy Decision-Support System for Dairy Cattle Culling Decisions ». A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfilment of the requirements of the degree of Masters of Science, May 1997.

- [Stumptner et al., 1998] Markus Stumptner and Franz Wotawa. « Model Based Reconfiguration ». Institut für Informations Systeme Abteilung für Datenbanken und Experten Systeme Technische Universität Wien Paniglgasse 16, A-1040 Wien, Austria, 1998.
- [Sur et al., 1998] Frédéric Sur, jérôme Lacaille. « Présentation de la Logique Floue ». Ecole Normal Supérieure de Cachan, Magistère de Mathématiques, Mémoire de première année, 1998.
- [Tang et al., 2007] Hao Tang and Shi Liu. « Basic Theory of Fuzzy Bayesian Networks and Its Application in Machinery Fault Diagnosis ». *Fourth International Conference On Fuzzy Systems and Knowledge Discovery (FSKD 2007)*, IEEE 2007.
- [Tomsovic et al., 2000] K. Tomsovic, M.Y. Chow. « Tutorial on Fuzzy Logic Applications in Power Systems ». Prepared for the IEEE-PES Winter Meeting in Singapore January, 2000.
- [Tu et al., 2004] Haiying Tu, Jeffrey Allanach, Satnam Singh, Krishna R.Pattipati et Peter Willett. « The Adaptive Safety Analysis and Monitoring System ». Electrical and Computer Engineering Department, University of Connecticut, 2004.
- [Wotawa, 1999] Franz Wotawa. « Model-based reasoning ». Institut für Informations Systeme, Technische Universität Wien, Paniglgasse 16, A-1040 Wien, Austria, 1999.
- [Zajpt et al., 1997] Christophe Zajpt, Michel Curé et Béatrice Mazzucotelli. « Système expert pour l'aide à la décision et à l'enseignement des diagnostics infirmiers ». *Informatique et Gestion Médicalisée*, Paris, 1997.

Webographie

- [1] <http://fr.wikipedia.org/wiki/Diagnostic>
- [2] <http://www-igm.univ-mlv.fr/~dr/XPOSE2004/jvaldes/index.html#1>
- [3] <http://perso-math.univ-mlv.fr/users/bouissou.marc/siteRB/Page0.htm>
- [4] <http://users.abo.fi/rfuller/fuzs.html>
- [5] http://fr.wikipedia.org/wiki/Cancer_de_la_prostate