

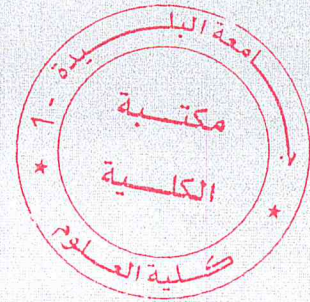


République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université SAAD DAHLAB de Blida  
Faculté des sciences  
Département informatique

En vue d'obtenir le diplôme du master

Spécialité: Ingénierie du Logiciel



**Thème:**

Analyse d'images rétiniennes par raisonnement à Base de Cas et Réseau bayésien pour l'aide au Diagnostic Médical

Réalisé par: *Maallem Abderrahmane*

*Bouhacene Houssem Eddine*

*promotrices: Benblidia Nadja*

**Le jury composé :**

*Aichja*..... Chargé de cours ,USDB

Président

*Reguiq*..... Chargé de cours, USDB

Examineur

..... Chargé de cours, USDB

Rapporteur

Blida, Décembre 2014

# ***Dédicace***

Je dédie le fruit de mes cinq (05) années d'études :

4- A mes parent qui me soutenue dans les moments les plus difficiles de ce long parcours et qui m'encouragé sans relâche à persévérer et à ne pas perdre espoir.

A ma chère sœur Rabab

A mon petit frère Abd el Malek.

A mes amis Youcef, choayb, Abd Errahmane, Zakaria qui mon aidé son relâche

A mes collègues de travail sans Exception

A mes très chers oncles Aboubakre, Khaled, ainsi qu'à son épouse.

A mes tantes Ratiba, Malika, Khalida, Amina, Safia, Salima.

Et a tous mes cousins et cousines.

# *Dédicace*

Je dédie le fruit de mes cinq (05) années d'études :

4- A mes parent qui me soutenue dans les moments les plus difficiles de ce long parcours et qui m'encouragé sans relâche à persévérer et à ne pas perdre espoir.

A ma chère sœur Amina

A mes frère Bilal, nasreddine, Abderrahim.

A mes amis Youcef, Hichem, Mohamed, Abd el Karim qui mon soutenue moralement.

A mes très chers oncles Ilyass, Mohamed, ainsi qu'à leurs épouses.

A mes tantes Wahiba, Hniffa, Kenza.

Et a tous mes cousins et cousines.

# **REMERCIEMENTS**

Il m'est agréable d'adresser l'expression de mon profond respect et d'exprimer ma gratitude et ma profonde reconnaissance à tous mes professeurs pour toutes les connaissances acquises, dans le domaine informatique, durant mon cursus universitaire.

Je voudrais ici remercier tout spécialement mes promoteurs Mlle BENBLIDIA Nadja et Mlle REGUIEG F. Zohra, Enseignantes -Chercheurs et Chargés de cours,

Je tiens à remercier en particulier les membres du jury

Je tiens à remercier en particulier M. Madani Yahya Youcef Etudiant en électronique pour sa disponibilité et son langage franc, et qui m'a été d'une aide précieuse

Enfin je tiens à souhaiter bonne chance à tous mes camarades de promotion.

Une fois encore merci à tous.

# Résumé

Dans ce travail, nous nous intéressons à l'imagerie rétinienne pour l'aide à la décision diagnostique d'une pathologie très répandue : la rétinopathie diabétique. Notre objectif est de concevoir un système pour la détection automatique des microanévrismes, premier signe de la rétinopathie diabétique, par l'intégration du traitement d'images dans la représentation des connaissances d'un système de raisonnement à partir de cas.

En phase de traitement on élabore des méthodes pour corriger l'image en utilisant la correction de non-uniformité. Cette méthode consiste à utiliser un filtre moyen sur l'image et après le soustraire de l'image originale pour améliorer le contraste. On élimine les vaisseaux vasculaires en utilisant le chapeau haut de forme.

On applique sur ce résultat une binarisation pour obtenir les liaisons les plus sombres (microanévrisme, hémorragie), après on utilise la fonction de détection de microanévrismes.

On élabore les probabilités des réseaux bayésiens qui convient le mieux avec l'incertitude médicale. Le choix du réseau bayésien naïf est de considérer comme étape de classification des images ; la probabilité de détection. Enfin on obtient des résultats de détection des microanévrismes fiables en les comparants avec une base d'images.

**Mots clés :** Rétinopathie diabétique, microanévrismes, morphologie mathématique ; réseau de Bayes

## Abstract

In this work, we focus on retinal imaging to aid in diagnostic decision of a widespread pathology of diabetic retinopathy. Our goal is to design a system for the automatic detection of microaneurysms, the first sign of diabetic retinopathy, the integration of image processing in knowledge representation of a system of reasoning from cases.

Phase treatment is developing methods to correct the image using the non-uniformity correction. This method is to use a median filter on the image and after subtracting the original image to enhance the contrast. Vascular vessels are removed using the top hat.

We apply this result binarization for the darkest links (microanévrisme, bleeding) after using the microaneurysms detection function.

It develops the probabilities of Bayesian networks is best with medical uncertainty. The choice of naïve Bayesian network is regarded as image classification step; the probability of detection. Finally we obtain reliable microaneurysms detection results by comparing them with an image database.

**Keywords:** Diabetic retinopathy, microaneurysms, mathematical morphology; Bayesian network

# Sommaire

<i>Introduction générale</i> .....	<i>1</i>
<i>Chapitre 1 : Contexte médical</i>	
1- Introduction .....	3
2- Anatomie de l'œil .....	3
3- Rétinopathie diabétique.....	6
3.1- Introduction .....	6
3.2- La Présentation de la maladie .....	6
3.3- Causes .....	7
3.4- Symptômes et Complications .....	7
3.5- Epidémiologie de la rétinopathie diabétique .....	8
3.5.1- Diabète de type 1 .....	8
3.5.2- Diabète de type 2 .....	8
3.6- Etude Clinique .....	9
3.6.1- La rétinopathie exsudative .....	9
3.6.2- La rétinopathie proliférante .....	10
3.6.2.1- La rétinopathie pré-proliférante.....	10
3.6.2.2. La rétinopathie proliférante .....	11
3.7- CLASSIFICATION DE LA RÉTINOPATHIE DIABÉTIQUE .....	12
3.7.1- Stades de la rétinopathie diabétique.....	12
3.7.2-maculopathies diabétiques .....	13
3.8- Traitement .....	14
3.9- Conclusion .....	14
4- l'Imagerie médicale (rétinienne).....	14
4.1- Introduction .....	14
4.2- Les étapes d'analyse .....	15
4.2.1. Prétraitement .....	15
4.2.1.1. Restauration d'images .....	16
4.2.1.2. Elimination de bruit .....	16
4.2.1.3. Les filtres .....	17
4.2.2. Segmentation .....	17
4.2.2.1. Segmentation par régions .....	18
4.2.2.2. Segmentation par contours .....	18
4.2.2.3. Segmentation par morphologie mathématique .....	19
5- Conclusion .....	21

## Chapitre 2 : Raisonnement à partir de cas

1.	Introduction .....	22
2.	Principe de RàPC .....	22
2.1.	Définition .....	22
2.2.	Résonnement par analogie .....	23
2.3.	Composantes d'un système de Raisonnement à Partir de Cas.....	24
3.	Model de RàPC .....	25
3.1.	Model simple .....	25
3.2.	Modèle à mémoire dynamique .....	26
3.3.	Modèle à base de catégories .....	27
4.	Cycle de RàPC .....	28
4.1.	Elaborer .....	29
4.2.	Retrouver .....	29
4.3.	Adapter .....	33
4.4.	Réviser .....	36
4.5.	Mémoriser .....	37
5.	Connaissances dans un système RàPC.....	37
6.	Les principaux systèmes RàPC en diagnostic médical .....	37
6.1.	Système CASEY .....	38
6.2.	Système PROTOS .....	39
6.3.	Système IDEM .....	39
7.	Traitement d'image et le RàPC .....	39
8.	Application du RàPC dans l'imagerie médicale .....	40
9.	Conclusion .....	41

## Chapitre 3 : Inférence bayésienne

1.	Introduction.....	43
2.	Définition.....	43
3.	Définition Formelle.....	44
4.	Représentation graphique de la causalité.....	44
4.1.	Exemple.....	45
4.2.	Circulation de l'information.....	45
4.3.	Définitions et propriétés.....	47
4.3.1.	Indépendance conditionnelle.....	47
4.3.2.	D-séparation.....	47
5.	Formule de Bayes.....	51
5.1.	Autres écritures du théorème de Bayes.....	52
5.2.	Exemple d'application de la formule de Bayes.....	52
6.	Construction des réseaux bayésiens .....	54
6.1.	Identification des variables et de leurs espaces d'états.....	54
6.2.	Définition de la structure du réseau bayésien .....	55

6.3.	Loi de probabilité conjointe des variables.....	55
7.	Propagation de messages (Algorithme Pearl).....	56
8.	Domaines d'application des réseaux bayésiens.....	57
9.	Réseau bayésien et l'incertitude médicale.....	58
10.	Avantage des réseaux bayésiens .....	59
11.	Conclusion.....	60

#### **Chapitre 4 : Détection des microanévrismes**

1.	Introduction.....	61
2.	Description de système.....	61
2.1.	phase recueil de donnée.....	61
2.1.1.	Prétraitement.....	62
2.1.1.1.	Correction de l'illumination non uniforme.....	62
2.1.1.2.	L'élimination du réseau vasculaire.....	62
2.1.2.	La segmentation d'image.....	63
2.1.2.1.	La binarisation (Seuillage).....	63
2.1.2.2.	Détection des microanévrismes.....	63
2.2.	Recherche de cas le plus probable.....	64
3.	Implémentation.....	64
3.1.	Environnement de Programmation.....	64
3.1.1.	MATLAB.....	64
3.2.	Présentation de l'application détecte_ microanévrismes.....	65
3.2.1.	Fenêtre principale.....	65
3.2.2.	Recueil d'information.....	66
3.2.3.	Fenêtre recherche de cas.....	67
3.3.	Application et interprétation des résultats.....	69
3.3.1.	Phase recueil des informations.....	69
4.	Conclusion.....	75
<b>6-Conclusion générale .....</b>		<b>76</b>

#### **Bibliographie**



# Liste des Figures

<b>Figure 1.1</b> : l'Anatomie d'œil [4].....	3
<b>Figure 1.2</b> la structure de la Rétine [4].....	5
<b>Figure 1.3</b> : Effet de la rétinopathie diabétique:(a):bonne vision, (b):mauvaise vision.....	6
<b>Figure1.4</b> :Forme exsudative. Les flèches indiquent la présence d'exsudats (lipidiques, secs) profonds périmaculaires .....	9
<b>Figure1.5</b> :4Forme exsudative. Couronne d'exsudats périmaculaires avec oedème.....	9
<b>Figure1.6</b> : Capillaropathie oedémateuse Diabétique en cours de grossesse, décompensation passagère. (M. Mouillon).....	10
<b>Figure1.7</b> :Forme proliférante. Voile néovasculaire intravitréen pédiculé sur la papille. (M. Mouillon).....	11
<b>Figure1.8</b> :Rubéose de l'iris chez un diabétique. (M. Mouillon).....	11
<b>Figure 1.9</b> : RD non proliférante modérée hémorragie en tache peu nombreuses .....	12
<b>Figure 1.10</b> : RD non proliférante sévère: nombreuses hémorragies en taches sur toute la périphérie.....	12
<b>Figure1.11</b> : Néovaisseau pré-rétinien.....	13
<b>Figure1.12</b> : Néovaisseaux pré-papillaires.....	13
<b>Figure 1.13</b> : Oedème maculaire focal localisé, entouré d'exsudats.....	13
<b>Figure 1.14</b> : Oedème maculaire diffus.....	13
<b>Figure1.15</b> : Prolifération fibrovasculaire (flèches).....	14
<b>Figure1.16</b> : les Effets de bruit.....	16
<b>Figure1.17</b> :Exemple d'une segmentation par contour.....	19
<b>Figure2.1</b> : Le carré d'analogie .....	23
<b>Figure2.2</b> . Modèle générique d'un système RàPC .....	25
<b>Figure2.3</b> : Model simple .....	26
<b>Figure2.4</b> :: Modèle à mémoire dynamique .....	29
<b>Figure2.5</b> : Modèle à base de catégories.....	29

<b>Figure 2-6</b> : Le cycle d'étapes du Raisonnement à Partir de Cas.....	29
<b>Figure 2.7</b> : Scénario de déroulement de l'algorithme KPPV sélectionnant le cas 1 (de classe solution "jaune") pour le cas cible C.....	32
<b>Figure 2.8</b> :Même scénario qu'en 3-5, mais avec un seuil plus bas ( $R_s$ plus grand) la solution "rouge" devient éligible. Elle n'est pas la plus représentée. Le cas 1 (jaune) est encore sélectionné.....	32
<b>Figure 2.9</b> :Dans ce scénario, aucune solution n'est éligible. Aucune expérience ne peut être rationnellement remémorée. Ce cas cible devra être résolu par une autre méthode.....	33
<b>Figure 2.10</b> : Les règles d'adaptation dépendent de la zone de regroupement des cas d'une même classe de solution. Ce regroupement est précisément réalisé sur la base des connaissances d'adaptation.....	34
<b>Figure 2.11</b> :Schéma général des connaissances et opérateurs d'adaptation.....	35
<b>Figure 2.12</b> : Schéma général d'adaptation d'un cas source en fonction des influences pb/sol et des opérateurs d'adaptation associés avec.....	36
<b>Figure2.13</b> : Coopération CASEY/ HFP [KOL 93].....	39
<b>Figure3.1.</b> Représentation graphique du modèle.....	45
<b>Figure3.2.</b> Modes de connexion.....	46
<b>Figure3.3.</b> Circulation de l'information dans le graphe de la Fig.3.1.....	46
<b>Figure.3.4.</b> Exemple de D-séparation.....	49
<b>Figure.3.5.</b> Connexion en série.....	49
<b>Figure.3.6.</b> Connexion divergente.....	50
<b>Figure3.7.</b> Connexion convergence.....	51
<b>Figure 3.8.</b> Représentation graphique de la relation entre la variable M (malade ou non malade) et la variable T (test positif ou négatif) dans le cas d'un test médical.....	53
<b>Figure3.9.</b> Etapes de construction d'un réseau bayésien.....	54
<b>Figure.3.10.</b> Principe de la fusion de données par réseau bayésien .....	58
<b>Figure.4.1</b> L'architecture de système.....	61
<b>Figure.4.2</b> La procédure de recueil de données .....	62
<b>Figure.4.3</b> Processus de recherche.....	64

<b>Figure.4.4</b> Fenêtre principale.....	65
<b>Figure.4.5</b> Fenêtre recueil d'information .....	66
<b>Figure.4.6</b> Recherche d'un cas .....	67
<b>Figure.4.7</b> Fenêtre construction de RB .....	68
<b>Figure.4.8</b> Sélection le cas probable .....	69
<b>Figure.4.9</b> Final résultat .....	69
<b>Figure.4.10</b> L'image après correction d'illumination non uniforme.....	70
<b>Figure.4.11</b> L'image après le supremum .....	71
<b>Figure 4.12:</b> L'image après l'application de chapeau haut de forme .....	72
<b>Figure 4.13 :</b> L'image obtenue après la binarisation.....	73
<b>Figure 4.14 :</b> Détection des microanévrismes.....	74
<b>Figure 4.15:</b> Après détection des microanévrismes.....	75

# Liste des Tables

**Tableau 1.1:** *Les éléments structurants élémentaires les plus usuels*..... 20

**Tableau 2.1:** *Adaptation de cas, recherche de cas, segmentation d'image et image interprétation* .....41

**Tableau 2.2 :** *Couleur, forme, contour, texture*..... 42

# Introduction générale

L'imagerie médicale est l'un des champs d'application privilégiés des techniques de l'analyse numérique d'images. Nous nous intéressons en particulier à l'imagerie rétinienne pour l'aide au diagnostic de la rétinopathie diabétique. En effet, avec le débarquement des systèmes d'angiographie rétinienne numérisée, il est possible d'examiner de manière automatique les images angiographiques afin de faciliter et d'améliorer le diagnostic de la rétinopathie diabétique.

En médecine spécialement, la compréhension des experts n'est pas bornée seulement à des règles, mais à un mélange de connaissances théoriques et pratiques. Ces dernières reposent sur des cas distincts et d'autres inhabituels, et le raisonnement des médecins prend en compte les deux. Les systèmes de connaissances médicales contiennent deux types de connaissances: la connaissance objective, que l'on retrouve dans les ouvrages médicaux et la connaissance subjective qui change fréquemment. Les deux types de connaissances peuvent être clairement séparés: la connaissance objective des ouvrages peut être figurée sous forme de règles ou de fonctions, tandis que la connaissance subjective est englobée dans des cas. La difficulté du traitement du changement des connaissances subjectives peut être résolue par l'assimilation progressive de nouvelles mises à jour des cas. Les expériences sont synonymes de cas concrets qui font intervenir des connaissances approximatives, mal définies ou vagues, ou exprimées en langage humain. Le raisonnement à base se pose sur la réutilisation d'expériences passées, mises dans une base de cas, pour résoudre de nouveaux problèmes.

L'information requise par les médecins pour diagnostiquer certaines pathologies est plus complexe. Dans le cas de la rétinopathie diabétique par exemple, les médecins analysent des séries d'images suivies d'information contextuelles. Les méthodes de raisonnement à base de cas dans le domaine médical nécessitent donc pouvoir intégrer de l'information symbolique, telles que des explications cliniques et des informations numériques comme des images.

Notre objectif est de décrire et développer un outil de recherche pour l'aide au diagnostic de la rétinopathie diabétique. Cet outil permettra de raisonner par similarité, à la fois pour indexer et rechercher les semblables mais il faudra aussi combiner des techniques d'indexation et de recherche conçues sur des critères visuels numériques objectifs (couleur, forme, texture etc..).

Les réponses sont composées sur le degré de similarité entre les caractéristiques de l'image requête et les images de la base, évaluées à partir de leurs signatures.

### **Organisation du mémoire :**

Notre mémoire se compose de 4 chapitres qui sont les suivants :

#### ***Chapitre 1 : Contexte médical***

Ce chapitre préliminaire décrit le cadre médical du projet : anatomie de l'œil particulièrement la rétine, la pathologie oculaire à laquelle nous nous intéressons, la rétinopathie diabétique, le dépistage et la surveillance de cette dernière, les différentes techniques d'exploration de la rétine notamment l'angiographie rétinienne à la fluorescéine.

#### ***Chapitre 2 : Raisonnement à partir de cas***

Ce chapitre est consacré à la présentation du raisonnement à partir de cas, ses principes généraux, ses modèles et les différentes phases qui le composent, les différentes sources de connaissances. Nous terminons ce chapitre par citer les différentes manières d'intégration du raisonnement à partir de cas et le traitement d'image.

#### ***Chapitre 3 : Inférence bayésienne***

Ce chapitre est consacré à la présentation des réseaux bayésiens ainsi que leur propriété notamment la représentation de l'indépendance conditionnelle et l'algorithme de propagation des messages locaux dans les poly arbres (Pearl)

#### ***Chapitre 4 : Détection des microanévrismes***

Ce chapitre est consacré à la modélisation et l'implémentation d'une application en MATLAB pour l'aide au diagnostic de la rétinopathie diabétique en utilisant le RàPC et le Réseau Bayésien pour fournir des résultats expérimentés des méthodes exprimés dans les Chapitre 2, et 3

# Chapitre 1



### 1. Introduction

Les yeux sont une merveille que dieu a donnée a l'humain, un des organes les plus perfectionnés de notre corps. Miroir de nos émotions et de nos pensées secrètes, elles permettent à un être vivant de capter la lumière pour ensuite l'analyser et interagir avec son environnement dans le monde.

Le fonctionnement de l'œil peut être rapproché à ce qui se passe quand on prend une photographie. Comme l'appareil photographique, les yeux permettent le contrôle de la focalisation d'une quantité de lumière à travers une lentille sur une surface photosensible ou une pellicule. Mais les yeux sont bien différents d'un appareil photographique, car après que l'image soit produite, elle est intégrée dans le système nerveux central. [1]

L'œil humain présente plusieurs pathologies provocantes des troubles de la vision. Parmi ces pathologies nous pouvons citer en particulier la rétinopathie diabétique. L'objectif de notre travail est le développement des outils qui servent à une aide au diagnostic en ophtalmologie, notamment pour le dépistage et le suivi de ces anomalies.

### 2. Anatomie d'œil

L'œil est situé dans l'orbite, une cavité dans la crâne : il a environ 2.5 cm de diamètre. La paroi de l'œil est formée de 3 tuniques [2] : la tunique fibreuse, la tunique uvéale, la tunique nerveuse. Ces tuniques enferment des milieux transparents : l'humeur aqueuse, le cristallin, le corps vitré.

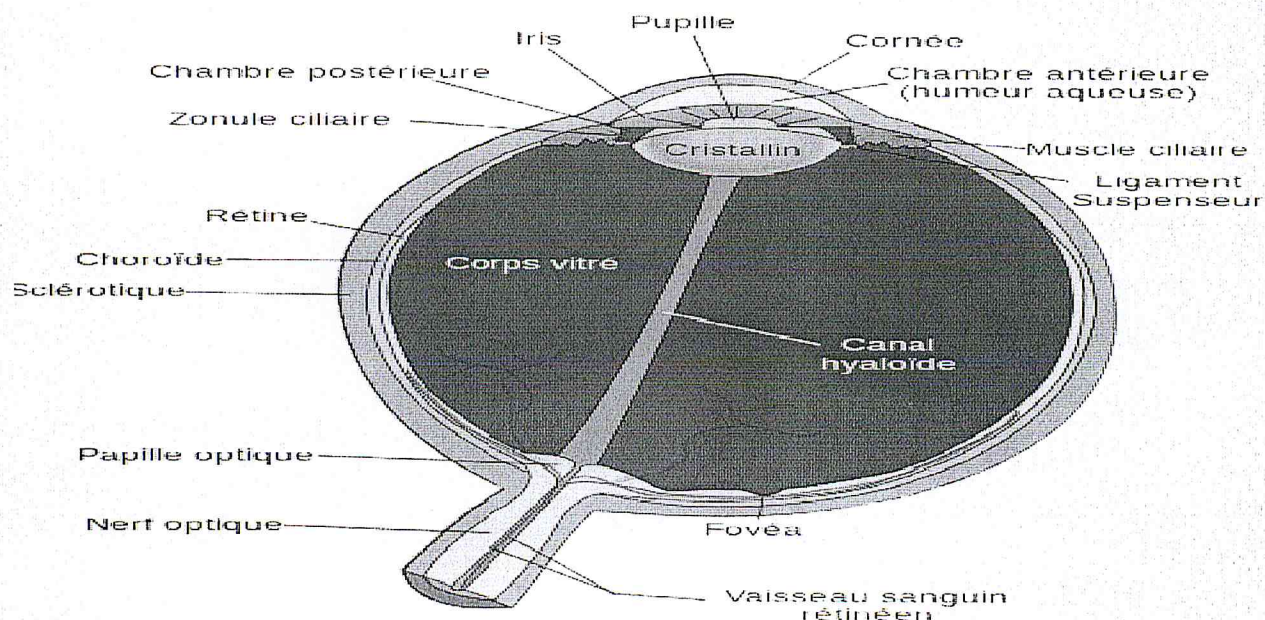


Figure 1.1 : l'Anatomie d'œil [3]

**La tunique fibreuse** : externe, se compose de la sclérotique opaque en arrière et de la cornée transparente en avant.

## Chapitre 1: Contexte médical

---

**La Sclérotique** : c'est une membrane rigide qui donne sa forme à l'œil, cette membrane blanche, fibreuse, résistante entoure et protège l'œil à l'extérieur tout en maintenant sa forme. Elle est traversée par un grand nombre de petites canaux (artères, nerfs, veines) et a l'arrière par une ouverture où passent les fibres du nerf optique.

**La cornée** : constitue la lentille principale du système optique oculaire. Pour que ce tissu puisse remplir sa fonction, il doit être transparent et tout concourt à assurer cette transparence. La cornée est enchâssée dans l'ouverture antérieure de la sclérotique, elle est la structure qui a la plus grande sensibilité tactile du corps humain.

**La tunique uvéale** : dite aussi uvée, se compose de 3 éléments : l'iris en avant, le corps ciliaire, la choroïde en arrière.

**L'iris** : est la partie colorée de l'œil, il permet de contrôler la taille de la pupille (Ouverture centrale de l'œil) et donc la quantité de lumière: diaphragme de l'œil. La contraction ou la dilatation réflexe règlent la quantité de lumière pénétrant dans l'œil par la pupille.

**Le muscle ciliaire** : est représenté par un épaississement de l'uvée situé en couronne derrière l'iris contenant des fibres musculaires lisses et des vaisseaux sanguins.

**La choroïde** : la choroïde est une couche située entre la sclérotique et la rétine. Elle est richement vascularisée ce qui permet la nutrition des tissus alentour. Notamment de la rétine. Elle contient de nombreux pigments colorés. Notamment la mélanine et l'hémoglobine. Elle forme aussi un écran et maintient ainsi l'intérieur de l'œil en chambre noire en absorbant les rayons lumineux pour la vision.

**La tunique nerveuse** : se compose de

**La rétine** : La rétine est une membrane nerveuse tapissant le fond de l'œil. Epaisse de quelques dixièmes de millimètre, d'une surface voisine de celle d'un timbre-poste, elle est constituée de plus de 130 millions de cellules nerveuses (125 millions de bâtonnets et 5 millions de cônes). Elle transforme la lumière en signaux électriques qui sont acheminés par le nerf optique vers le cerveau.

**a. les bâtonnets** : les plus nombreux, essentiellement situés au niveau de la rétine périphérique, sont responsables de la vision périphérique en lumière de faible intensité. Ils ont une très grande sensibilité à la lumière, d'où leur capacité à percevoir de très faibles lueurs la nuit. Ainsi ils ont une très faible perception des détails et des couleurs car plusieurs dizaines de bâtonnets ne sont liés qu'à une seule fibre du nerf optique.

**b. Les cônes** : dans principalement regroupés dans la rétine centrale, au sein d'une zone ovale appelée, macula, sont destinés à la vision des détails et à la vision colorée. Ainsi ils ont une très bonne sensibilité aux couleurs. Ils sont de trois types selon le pigment qu'ils contiennent et ont donc une sensibilité à des ondes lumineuses de longueurs différentes: cônes contenant de l'érythrosine (sensibles au rouge), de la chloropsine (vert), du cyanopsine (bleu).

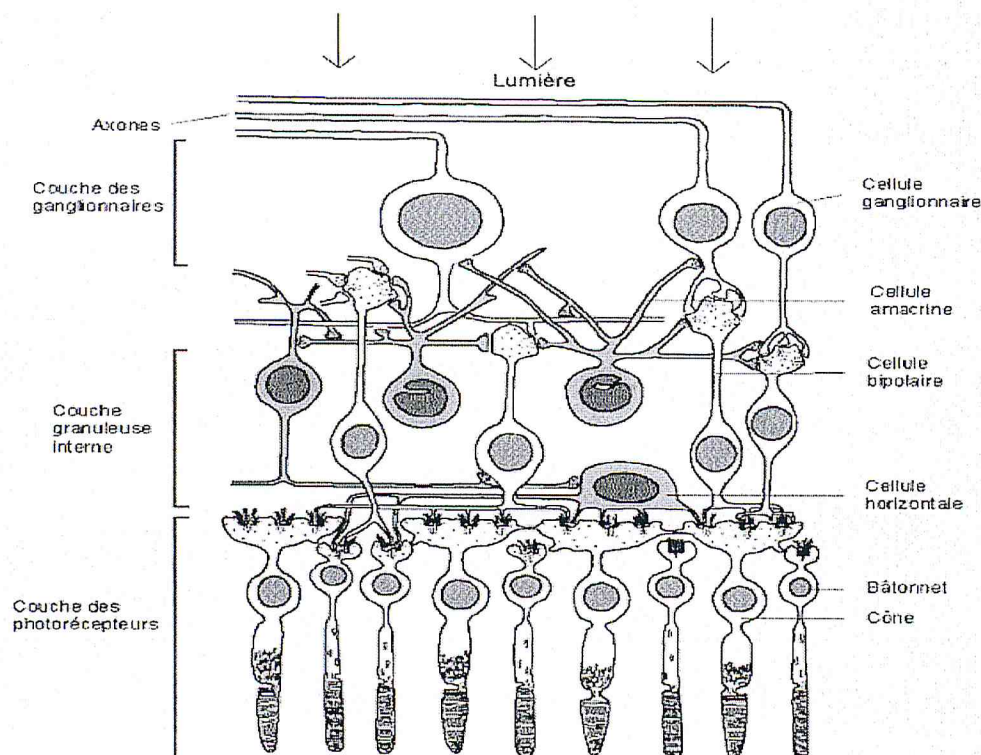


Figure 1. 2: la structure de la Rétine [3]

**La fovéa** : est le centre géométrique de la rétine plus précisément la zone centrale de la macula. Elle possède une forte densité de cônes, ce qui explique la perfection et la précision

**La tache aveugle** : quant à elle, est un endroit dépourvu de photorécepteurs car c'est le point de convergence de tous les axones des cellules ganglionnaires formant le nerf optique.

**La macula** : « la macula lutea », ou tache jaune, est la zone de la rétine caractérisée par une concentration maximale de cônes. Située au fond de l'œil, dans l'axe de la pupille, la macula a un diamètre d'environ 2 mm.

La macula contient en son centre une petite dépression, la fovéa : entièrement composée de cônes serrés les uns contre les autres, celle-ci est la zone d'acuité maximale de l'œil, c'est-à-dire celle qui donne la vision la plus précise en éclairage diurne. C'est sur elle que l'on amène l'image du point vers lequel on dirige le regard. [4]

### Conclusion

L'œil est donc constitué de trois membranes : la rétine, la choroïde et la sclérotique, et de trois milieux transparents : l'humeur aqueuse, le cristallin et l'humeur vitrée. Les membranes sont à la périphérie de l'œil, hétérogènes et opaques (sauf pour la cornée). En revanche, les milieux transparents sont au centre de l'œil, homogènes et transparents.

### 3. La rétinopathie diabétique (ou la RD)

#### 3.1. Introduction

Des grandes études épidémiologiques ont permis de mieux cerner la prévalence des complications rétiniennes de la rétinopathie diabétique (RD), après tout elle est parmi les premières causes de cécité avec la dégénérescence maculaire liée à l'âge (DMLA), le glaucome, la cataracte et la myopie dégénérative.

#### 3.2. La Présentation de la maladie

La rétinopathie diabétique regroupe des lésions de la rétine caractéristiques chez des personnes ayant un diabète sucré depuis plusieurs années. Elle apparaît lorsque l'excès de sucre dans le sang, caractéristique du diabète, a endommagé les petits vaisseaux qui approvisionnent la rétine en nutriments et en oxygène [5]. Les anomalies rétiniennes qui caractérisent la rétinopathie diabétique suivent une évolution prédictible sans variations majeures dans leur ordre d'apparition. La rétinopathie diabétique est le résultat de troubles vasculaires rétiniens. Les stades précoces se caractérisent par des occlusions et des dilatations vasculaires rétiniennes. Ensuite, elle évolue vers une rétinopathie proliférative avec apparition de néo-vaisseaux. L'apparition du l'œdème maculaire réduit souvent considérablement la vision.

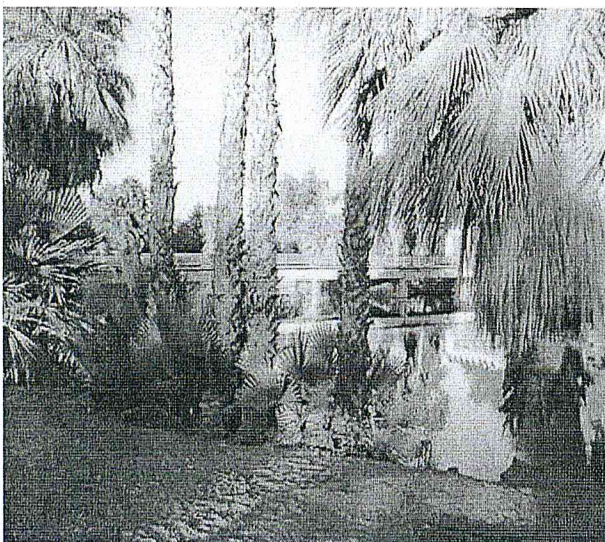


Figure 1.3 : Effet de la rétinopathie diabétique:(a):bonne vision, (b):mauvaise vision [5].

### 3.3. Causes

**La rétinopathie est habituellement le signe d'un autre état pathologique.** Bien que plusieurs troubles médicaux (par ex. la drépanocytose, le lupus) puissent causer une rétinopathie, sa présence est le plus souvent imputable au diabète ou à l'hypertension (une pression artérielle élevée). . [6]

**La rétinopathie diabétique est une complication du diabète.** Le diabète se caractérise par un taux de sucre sanguin élevé susceptible d'endommager les vaisseaux sanguins. Les vaisseaux rétiniens lésés peuvent laisser fuir des protéines et des acides gras qui forment des dépôts et finissent par gêner la vision. Les vaisseaux sanguins endommagés n'approvisionnent pas efficacement la rétine en oxygène et d'autres dommages peuvent s'ensuivre. . [6]

Dans la forme évoluée, désignée *rétinopathie proliférante*, de nouveaux vaisseaux sanguins se multiplient en abondance dans l'œil. Leur fragilité est souvent responsable de ruptures qui entraînent des saignements dans l'œil. Un saignement peut causer des cicatrices dans l'œil et perturber la vision. . [6]

**La rétinopathie hypertensive est une complication d'une pression artérielle élevée et elle évolue habituellement au cours de plusieurs années.** Une pression artérielle élevée endommage les vaisseaux sanguins qui s'épaississent et se rétrécissent sous son influence. L'irrigation sanguine de la rétine se trouve diminuée et des dommages s'ensuivent. Des fuites de sang peuvent éventuellement se produire dans la rétine et engendrer des dommages supplémentaires. [6]

### 3.4. Symptômes et Complications

**Il est possible qu'absolument aucun symptôme ne soit observable durant le stade précoce d'une rétinopathie diabétique.** À mesure que l'affection évolue, les symptômes comportent :

- une vision floue ;
- des fluctuations de la vision ;
- des mouches volantes ;
- des taches aveugles ;
- une modification de la perception des couleurs ;
- une perte de la vision subite ;
- une vision double ;
- une douleur oculaire dans les cas plus évolués.

**Des micro-anévrysmes sont les signes les plus précoces de rétinopathie diabétique que votre médecin puisse déceler.** Il s'agit de dilatations se constituant en poche sur la paroi des petits vaisseaux sanguins de la rétine. Elles apparaissent comme de minuscules points rouges à l'arrière de l'œil et quand parfois elles éclatent, elles entraînent un saignement dans la rétine et une vision trouble.

Une forme plus évoluée de la rétinopathie diabétique appelée *rétinopathie diabétique proliférante*, peut entraîner des cicatrices responsables d'une baisse de la vision. Dans un cas de rétinopathie proliférante, de nouveaux vaisseaux sanguins se multiplient sur la rétine et dans le *corps vitré* (la substance gluante entre le cristallin et la rétine). Ces vaisseaux sanguins peuvent gonfler et éclater ; le saignement qui s'ensuit endommage l'œil. Dans certains cas, le sang se trouve réabsorbé, mais bien souvent la rétine se détache et provoque une cécité totale.

Tout comme pour la rétinopathie diabétique, il est possible qu'aucun symptôme ne soit discernable dans les stades précoces de la rétinopathie hypertensive. Toutefois, à mesure que le trouble progresse, les symptômes comportent. [6]

### 3.5. Epidémiologie de la rétinopathie diabétique

Le diabète est une maladie chronique qui apparaît lorsque le pancréas ne produit pas suffisamment d'insuline ou que l'organisme n'utilise pas correctement l'insuline qu'il produit. L'insuline est une hormone qui régule la concentration de sucre dans le sang. L'hyperglycémie, ou concentration sanguine élevée de sucre, est un effet fréquent du diabète non contrôlé qui conduit avec le temps à des atteintes graves de nombreux systèmes organiques et plus particulièrement des nerfs et des vaisseaux sanguins.

347 millions de personnes sont diabétiques dans le monde. En 2004, on estimait que 3,4 million de personnes étaient décédées des conséquences d'une glycémie élevée à jeun. En 2010, le nombre de décès a été comparable. Plus de 80% des décès par diabète se produisent dans des pays à revenu faible ou intermédiaire.

Le diabète de type 2 est beaucoup plus fréquent que le diabète de type 1 puisqu'il représente environ 90 % des cas mondiaux. Pourtant, le risque de rétinopathie diabétique est plus élevé en cas de diabète de type 1 [7]

En Algérie, Plus de 2 500 000 personnes diabétiques. Le nombre des diabétiques en Algérie est passé d'un million de personnes en 1993, à plus de 2 500 000 personnes en 2007. Soit 10% de la population. Ces statistiques établies par l'OMS [7]

#### 3.5.1- Diabète de type 1

Le diabète de type 1 (précédemment connu sous le nom de diabète insulino-dépendant ou juvénile) est caractérisé par une production insuffisante d'insuline et exige une administration quotidienne de cette dernière. La cause de diabète de type 1 n'est pas connue, et en l'état des connaissances actuelles, il n'est pas évitable:

Les symptômes sont les suivants: excrétion excessive d'urine (polyurie), sensation de soif (polydipsie), faim constante, perte de poids, altération de la vision et fatigue. Ces symptômes peuvent apparaître brutalement. [8]

### 3.5.2- Diabète de type 2

Le diabète de type 2 (précédemment appelé diabète non insulino-dépendant ou diabète de la maturité) résulte d'une mauvaise utilisation de l'insuline par l'organisme. Le diabète de type 2 représente 90% des diabètes rencontrés dans le monde. Il est en grande partie le résultat d'une surcharge pondérale et de la sédentarité.[8]

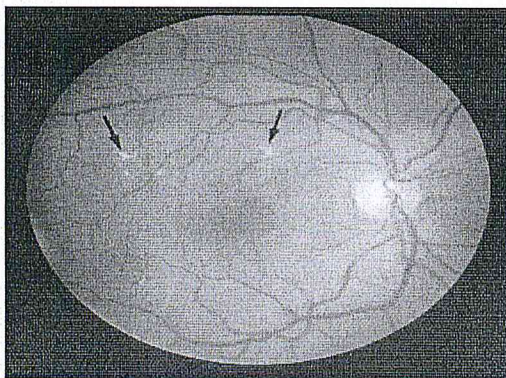
## 3.6. Etude Clinique

Repose sur un examen ophtalmologique minutieux et complet dont l'examen du fond d'œil qui doit être complété par l'angiographie fluorescéinique (essentiel +++).

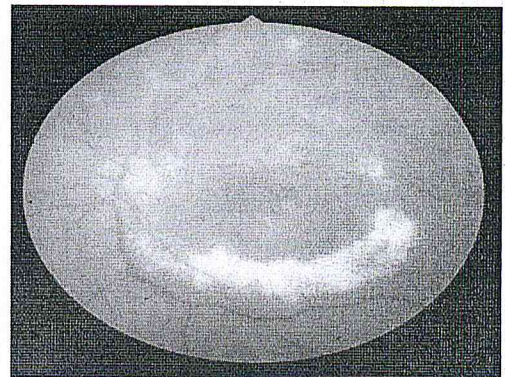
### 3.6.1- La rétinopathie exsudative

Elle est due à la capillaropathie œdémateuse. Encore appelée " non proliférante ", elle associe des anomalies microcirculatoires et des exsudats sans néovaisseaux.

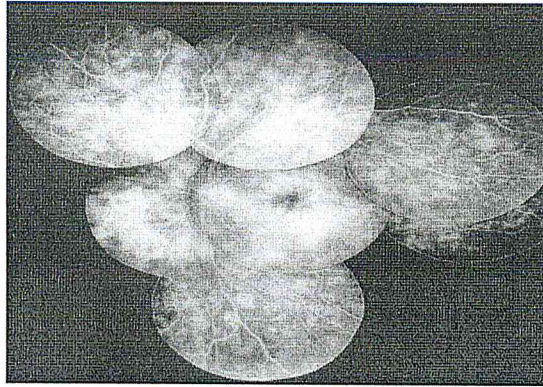
- Les micro-anévrysmes, petites taches rouges au fond d'œil, parfois difficiles à distinguer des hémorragies punctiformes, ils prennent l'aspect de point blanc fluorescent en angiographie.
- Les anomalies vasculaires : dilatation veineuse, zones de non perfusion (visibles seulement en angiographie).
- Les exsudats "durs" "lipidiques" profonds disposés en couronne, ils peuvent être confluent, ils ne sont pas fluorescents.
- Les hémorragies rétinienne de tous types peuvent disparaître (restent "noires" sur tous les temps de l'angiographie).
- L'œdème maculaire représente la principale menace fonctionnelle. D'abord diffus, il devient vite organisé et aboutit à une dégénérescence irréversible. [9]



**Figure1.4** :Forme exsudative. Les flèches indiquent la présence d'exsudats (lipidiques, secs) profonds périmaculaires[9]



**Figure1.5** :4Forme exsudative. Couronne d'exsudats périmaculaires avec oedème. [9]



**Figure1.6 :** Capillaropathie oedémateuse.

Diabétique en cours de grossesse, décompensation  
passagère [9]

### ***3.6.2- La rétinopathie proliférante***

Elle est souvent précédée du stade pré-proliférant qui doit être reconnu.

#### ***3.6.2.1- La rétinopathie pré-proliférante***

Elle associe :

- des anomalies des gros troncs veineux (boucles, oméga, etc...) en rapport avec des tractions vitréennes à ce niveau.
- des anomalies micro-vasculaires intra-rétiniennes (A.M.I.R), (Groupement de microanévrisme dilatations capillaires localisées, petits territoires ischémiques).
- des occlusions capillaires : zones "noires" de non perfusion en angiographie – images d'arrêt circulatoire (arbre mort).
- des exsudats cotonneux "mous" superficiels (nodules dysoriques), témoins d'un trouble circulatoire local. Ils peuvent disparaître. De teinte jaunâtre à bords flous, ils restent sombres sur les temps angiographiques.

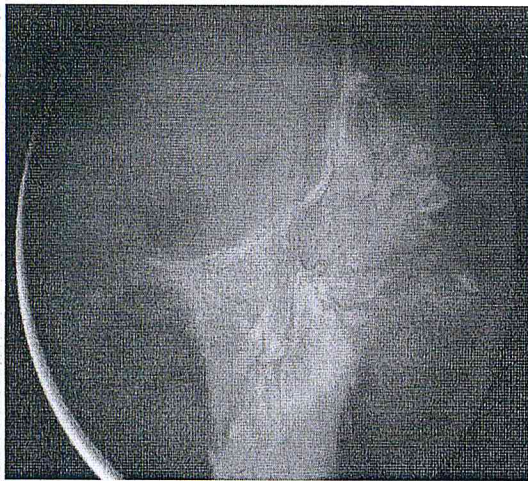


### 3.6.2.2. La rétinopathie proliférante

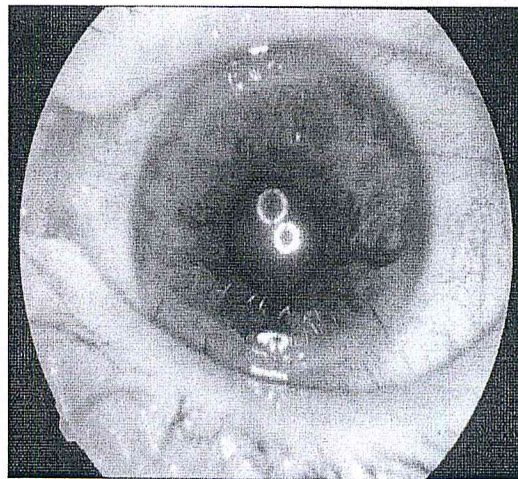
C'est la réponse à l'hypoxie rétinienne chronique et étendue.

Elle est caractérisée par l'existence de néovaisseaux, développés en bordure des territoires

ischémiques ou en avant de la papille, leur paroi est fragile et laisse diffuser le colorant sur l'angiographie. Ils peuvent proliférer au niveau du segment antérieur (rubéose de l'iris – glaucome néovasculaire). Complications : hémorragie du vitré, décollement rétinien par traction.



**Figure 1.7 :** Forme proliférante. Voile néovasculaire intravitréen pédiculé sur la papille. [9]



**Figure 1.8 :** Rubéose de l'iris chez un diabétique. [9]

### 3.7. Classification de la rétinopathie diabétique

La rétinopathie diabétique débute par un stade de rétinopathie diabétique non proliférante minime puis modérée, et évolue vers la rétinopathie diabétique non proliférante sévère (ou préproliférante) caractérisée par une ischémie rétinienne étendue, puis vers le stade de rétinopathie diabétique proliférante caractérisée par la prolifération de néovaisseaux à la surface de la rétine et/ou sur la papille.[9]

L'œdème maculaire est associé aux rétinopathies diabétiques non proliférantes ou proliférantes.

#### 3.7.1. Stades de la rétinopathie diabétique

- pas de rétinopathie diabétique
- rétinopathie diabétique non proliférante
- ♣ RD non proliférante minime (microanévrismes isolés, hémorragies rétinienne punctiformes peu nombreuses).

## Chapitre 1: Contexte médical

♣ RD non proliférante modérée (nombreux microanévrismes, hémorragies rétiniennes punctiformes, nodules cotonneux, signes d'ischémie rétinienne peu nombreux)

♣ RD non proliférante sévère (ou RD préproliférante = RDPP), définie par l'association aux signes précédents de signes ophtalmoscopiques évocateurs d'ischémie rétinienne sévère, notamment hémorragies intra-rétiniennes étendues en tache, anomalies veineuses et

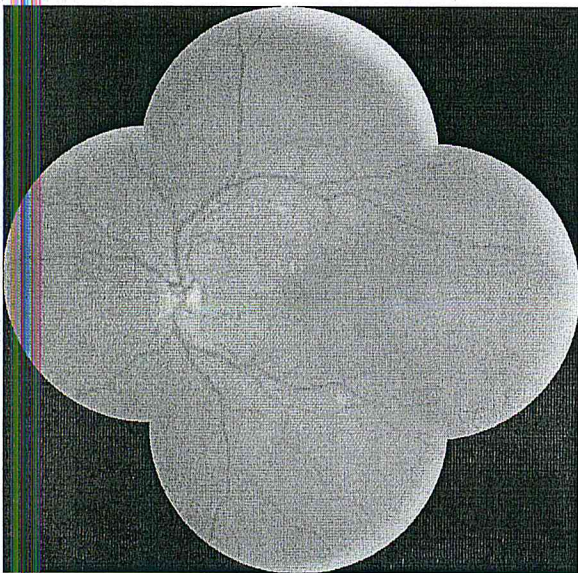
AMIRs. C'est un stade de haut risque d'évolution vers la néovascularisation, notamment dans certaines circonstances cliniques (cf. IV. Evolution)

• rétinopathie diabétique proliférante (RDP) :

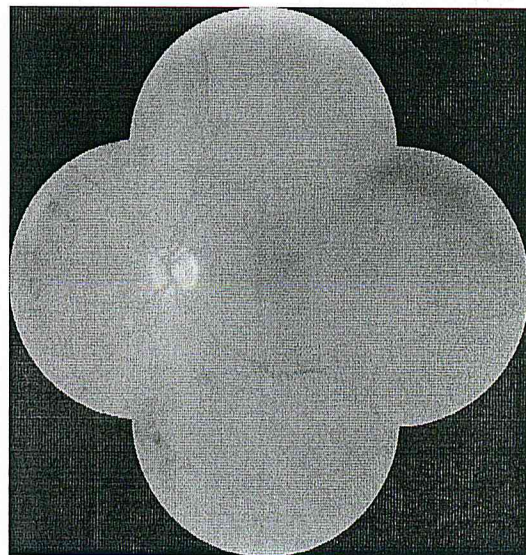
♣ RDP non compliquée : présence de néovaisseaux pré-rétiniens et/ou prépapillaires

♣ RDP compliquée (voir «Physiopathogénie») :

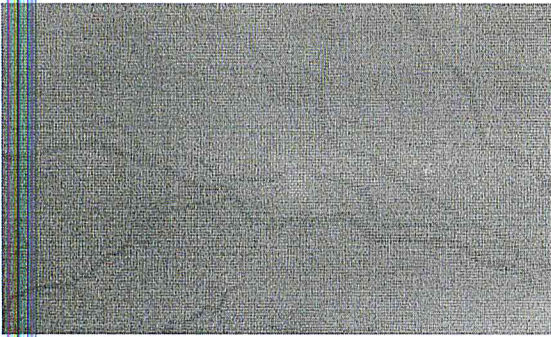
- hémorragie du vitré,
- décollement de rétine par traction,
- glaucome néovasculaire.



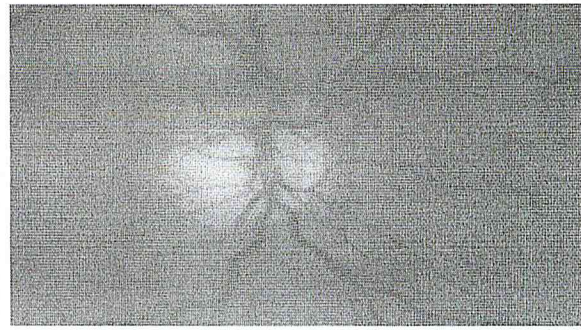
**Figure 1.9** : RD non proliférante modérée hémorragie en tache peu nombreuses [9]



**Figure 1.10** : RD non proliférante sévère: nombreuses hémorragies en taches sur toute la périphérie [9]



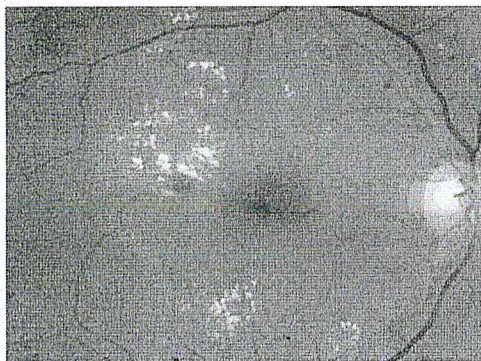
**Figure 1.11** : Néovaisseau pré-rétinien [9]



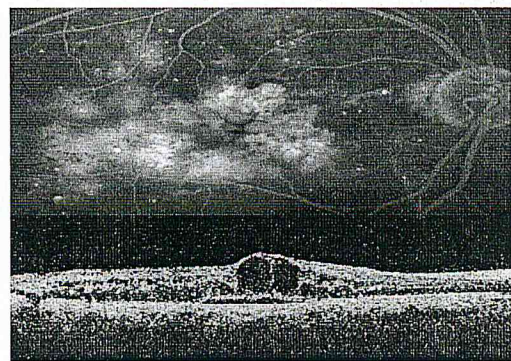
**Figure 1.12** : Néovaisseaux pré-papillaires [9]

### 3.7.2. maculopathies diabétiques

- oedème maculaire focal : oedème localisé entouré d'exsudats.
- oedème maculaire diffus : oedème maculaire étendu, lié à une hyperperméabilité généralisée à partir des capillaires du pôle postérieur, non cystoïde ou cystoïde (en fonction de la présence de kystes intrarétiniens)
- maculopathie ischémique



**Figure 1.13**: Oedème maculaire focal localisé, entouré d'exsudats [10]



**Figure 1.14** : Oedème maculaire diffus [10]

- Oedème rétinien étendu à toute la surface de la macula, s'imprégnant de fluorescéine, avec de nombreuses logettes, d'aspect kystique en angiographie, caractéristique de l'oedème maculaire diffus cystoïde. L'OCT montre l'épaississement rétinien maculaire associé à des kystes intrarétiniens

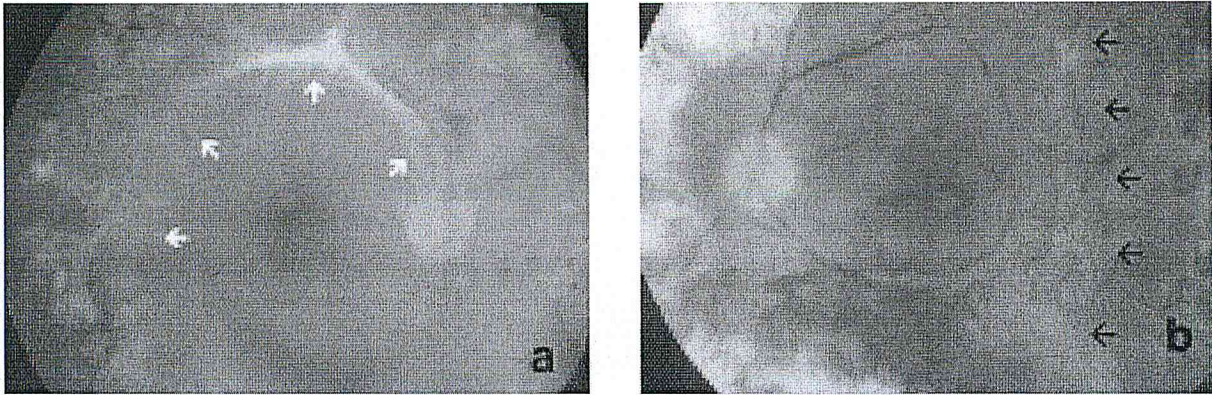


Figure 1.15 : Prolifération fibrovasculaire (flèches) [10]

### 3.8. Traitement

Le traitement de la rétinopathie diabétique dépend du type de rétinopathie que vous avez. Pour la rétinopathie non proliférative, vous n'aurez pas besoin de traitement. Cependant, vous devez continuer à faire des examens oculaires réguliers pour identifier le plus tôt possible tout développement de la maladie. Si vous avez une rétinopathie pré proliférative, vous n'aurez sans doute pas besoin de traitement non plus. Cependant, un traitement au laser peut être conseillé si les écoulements commencent à gêner votre vision.

Si vous avez une rétinopathie proliférative, un traitement au laser pourra être utilisé pour « brûler » les vaisseaux sanguins anormaux, afin d'éviter une croissance supplémentaire. Le traitement au laser ne cible pas directement les vaisseaux sanguins, mais détruit ceux qui, autour de votre rétine, n'ont plus d'oxygène. Il ne pourra pas restaurer la vision perdue, mais pourra empêcher les aggravations. Dans de rares cas, une opération chirurgicale pourra être pratiquée si les saignements de votre oeil sont importants. Ceci peut se produire si la rétinopathie proliférative a été diagnostiquée à un stade avancé. Votre généraliste, ou votre spécialiste, pourra vous expliquer les possibilités, notamment les avantages et les risques associés au traitement au laser et à la chirurgie oculaire.

### 3.9. Conclusion

Une surveillance étroite du diabétique, concertée entre médecin généraliste, diabétologue et ophtalmologiste, doit permettre un meilleur dépistage et un meilleur contrôle des lésions rétinienne. Un traitement précoce bien conduit évitera les complications gravissimes et améliorera le pronostic fonctionnel.

## 4. l'Imagerie médicale

### 4.1. Introduction

Il existe un nombre important de situations ou de maladies pouvant être suivies d'une modification de forme (tortuosité) des vaisseaux de la rétine. On peut citer l'artériosclérose, le diabète

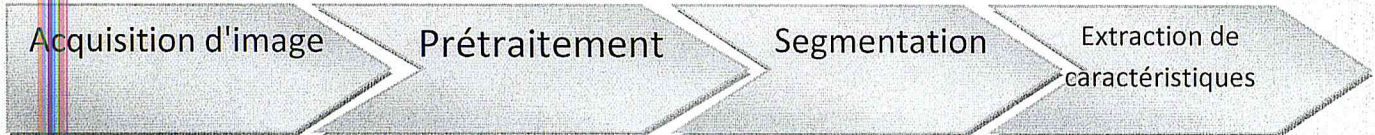
et la forme SC de la drépanocytose. Penmam et al, Binaghi et Levy ont prouvé que seul un suivi fréquent et régulier peut permettre l'évaluation des risques d'une perte sévère de la vision encourus par les patients. Jusqu'à présent, les ophtalmologistes utilisent les images analogiques d'angiographie rétinienne en fluorescence pour analyser les modifications de forme des vaisseaux de la rétine. Une telle tâche est fastidieuse et parfois subjective.

### 4.2. Les étapes d'analyse [4]

Le but d'analyse est d'extraire des informations qualitatives ou quantitatives d'une image ou d'un ensemble des images, cela prend la mise en œuvre et l'enchaînement d'un certain nombre de processus. Dans ce contexte, nous nous intéressons à la chaîne de l'analyse d'image, elle se décompose en plusieurs étapes qui partent de l'acquisition d'image, après le prétraitement, amélioration d'image et élimination du bruit pour mettre en évidence l'information utile qu'elle contient et /ou d'atténuer l'information inutile.

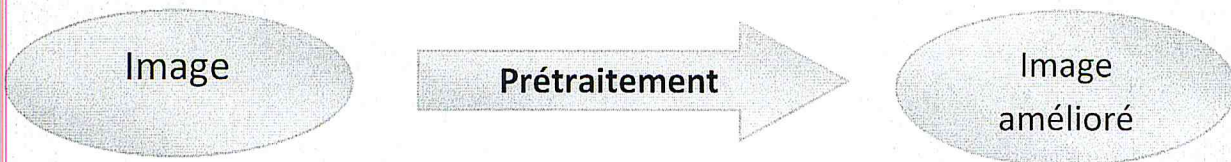
Les limites entre l'info utile et l'info inutile n'est pas simple à assurer car elle dépend de la nature d'image et de l'application envisagée. L'étape suivante concerne la segmentation qui à partir de l'image médicale originale ou l'image clarifiée, s'attache à reconstruire les régions qui correspondent aux objets présents dans l'image.

La segmentation transforme l'image médicale en une image où chaque pixel est caractérisé par une étiquette correspondant à la région à laquelle il appartient. [1]



#### 4.2.1. Prétraitement

Il permet une meilleure visualisation de l'image, corrige les effets de l'étape acquisition (corrections radiométriques et/ou géométriques), il réduit le bruit et facilite les traitements suivants (segmentation, extraction de la caractéristique). Les principales méthodes sont développées pour améliorer les images et éliminer le bruit.



L'amélioration d'images est utilisée dans plusieurs domaines, on considère deux sources de dégradation :

1. les distorsions introduites par le système d'acquisition ou de stockage.

2. les perturbations liées aux bruits (d'origine variée tels que les bruits phonétiques, de quantification, électroniques, les rayures..).

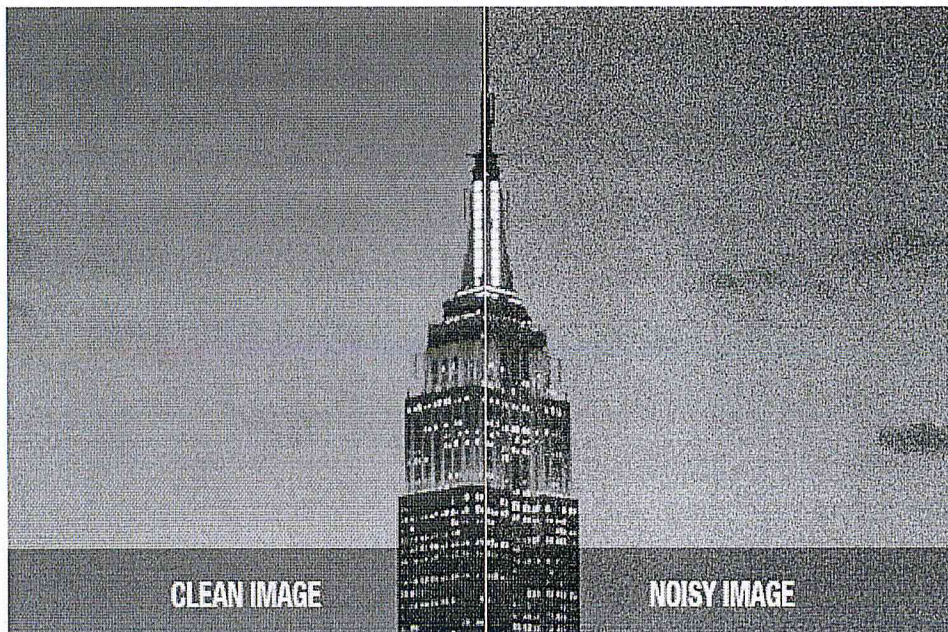
### 4.2.1.1- Restauration d'images

Elle réduit le bruit (l'élimination des distorsions introduites par le système ayant servi à acquérir l'image), ce travail doit permettre d'obtenir une image qui soit la plus proche possible de l'image idéale qui aurait été obtenue si le système d'acquisition était parfait.

### 4.2.1.2- Elimination de bruit

Il existe trois types de bruit :

- 1- Bruit lié au contexte de l'acquisition (bougé, dérive lumineuse, flou, etc...).
- 2- Bruit lié au capteur (distorsion de la gamme des niveaux de gris, distorsion géométrique, mauvaise mise au point, etc...).
- 3- Bruit lié à la numérisation (codage, quantification, échantillonnage, etc...).



**Figure1.16** : les Effets de bruit

Pour l'élimination du bruit on utilise les filtres comme suivant :

### 4.2.1.3- Les filtres

On distingue plusieurs types des filtres :

- **Filtre linéaire**

Il remplace chaque pixel par une valeur égale à une somme pondérée de ses voisins (convolution). Les noyaux de convolution sont des matrices, et leur taille dépend du voisinage (3x3, 5x5). Ce sont toujours des matrices carrées et possédant un nombre de lignes et colonnes impaire [10].

On utilise le plus souvent des filtres :

1. Passe-haut qui applique une opération dérivée sur l'image : Gradient, Laplacien.
2. Passe-bas qui réalise donc lissage : Moyenne, Gaussien.

- **Filtre lissage Moyenne (smoothing)**

Ce filtre passe-bas permet de remplacer la valeur d'un pixel par la moyenne de sa valeur et de la valeur des 8 pixels qui l'entourent. Le filtre moyen a pour but :

- 1- de lisser l'image (smoothing)
- 2- de remplacer chaque pixel par la moyenne des valeurs des pixels adjacents et du pixel central.
- 3- de réduire les parasites (bruits de mesure).
- 4- de réduire les détails non importants.

- **Le filtre non linéaire**

La famille des filtres non linéaires est vraiment large, elle regroupe les filtres morphologiques, les filtres homomorphiques, les filtres basés sur l'approche markovienne, le filtre d'ordre,...etc. La famille des filtres morphologiques est beaucoup utilisée en imagerie médicale, la famille des filtres homomorphiques est plus juste pour le rehaussement du contraste, en diminution de bruit multiplicatif. La famille des filtres d'ordre est généralement entrevue en réduction de bruit.

### 4.2.2. Segmentation

La segmentation permet de différencier les objets d'intérêt d'une image en les isolants du fond ou des autres choses. Après la segmentation d'une image, chaque pixel se voit attribuer à une région et chaque région convient à une partie sémantique de l'image. Les discontinuités entre les régions correspondent aux contours des objets, on distingue 3 types de segmentation :

1. **Segmentation par régions** cette approche cherche à détecter les zones de l'image présentant des caractéristiques homogènes en vérifiant les critères d'homogénéité et de similitude. Les pixels présentant une caractéristique commune, par exemple l'intensité seront regroupés.

2. **Segmentation par contours** cette approche tente de détecter les contours correspondant aux positions qui rendent localement maximale la variation d'un critère. Ces deux conceptions sont duales : une région définit son contour, un contour définit une région.

3. **Segmentation par morphologie mathématique** cette approche consiste à l'extraction de connaissances à partir de mesures fondées sur les formes et les structures des objets contenus dans l'image étudiée. Le principe de base consiste à parcourir cette dernière à l'aide d'éléments structurants assimilables à des filtres et présentant des propriétés géométriques connues, choisies en fonction des objectifs à atteindre [11]

### 4.2.2.1. Segmentation par régions

La segmentation d'image en région est un traitement de bas niveau qui consiste à créer des partitions dans l'image en identifiant des pixels similaires. Le critère de similarité permet de contrôler l'aspect final de la segmentation (regroupement par couleur, par homogénéité, par taille...etc.).

La segmentation par régions est une approche spécifique dans laquelle on cherche à construire des surfaces en regroupant des pixels voisins suivant un critère d'homogénéité et de connexité. Les régions doivent respecter les quatre conditions suivantes :

- 1-  $I = \cup R_i$  avec  $I$ =image et  $R_i$  appelé région .... (1)
- 2-  $R_i$  est constituée de pixels connexes pour tout  $i$  .... (2)
- 3-  $\text{Pred}(R_i) = \text{vrai}$  pour tout  $i$ ...(3)
- 4-  $\text{Pred}(R_i \cup R_j) = \text{faux}$  pour tout  $i \neq j$ ,  $R_i$  et  $R_j$  étant adjacents ... (4)

Ces conditions expriment que l'homogénéité d'une région  $R_i$  est définie par un prédicat d'uniformité noté  $\text{Pred}(R_i)$ . Ce prédicat est vrai si  $R_i$  est homogène, faux dans le cas contraire [13]

Il existe cinq approches de segmentation en régions:

- 1- Segmentation par régions en utilisant un seuil
- 2- La segmentation par croissance de régions.
- 3- La segmentation par division de régions.
- 4- La segmentation par fusion de régions.
- 5- La segmentation par division-fusion de régions.

### 4.2.2.2. Segmentation par contours

Dans certains cas, il est difficile de distinguer des zones par leur valeur de niveau de gris. Cependant, il est possible d'en distinguer les frontières. La méthode de segmentation devra tirer parti de cette propriété. En général, la détection de contour n'est qu'une étape préliminaire dans la reconnaissance d'objets car elle fournit plus d'informations sur les régions que les régions elles-mêmes et doit donc être complétée par un algorithme de segmentation par régions. Cependant, on peut dire que les contours constituent des indices riches, au même titre que les points d'intérêts, pour toute interprétation ultérieure de l'image et méritent donc d'être traités à part [12].

Le principe de la détection de contours repose donc sur l'étude des dérivées de la fonction d'intensité dans l'image :

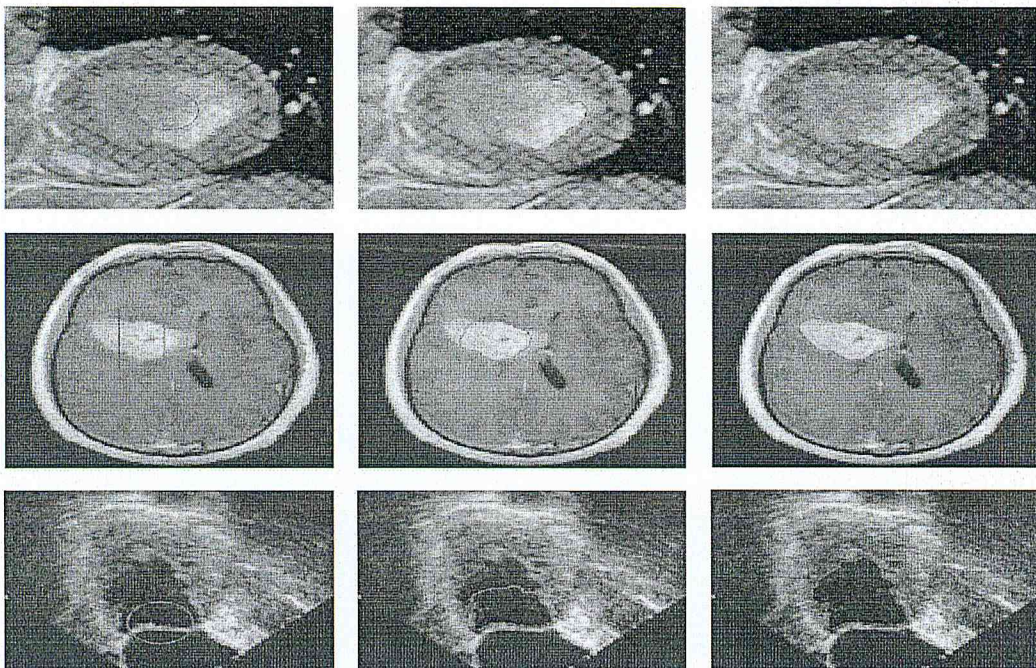


- 1- Les extrema locaux du gradient de la fonction d'intensité.
- 2- Les passages par zéro du laplacien.

Ces deux approches ont pour principe d'assimiler les contours aux discontinuités d'ordre 0 de la fonction d'intensité. Hors le calcul de la dérivée nécessite un pré-filtrage des images. Les filtres sont linéaires dans le cas de bruit de moyenne nulle ou non linéaire pour les bruits impulsionnels. Les inconvénients sont le plus souvent dus à la présence de bruit dans les images.

Les différentes approches existantes peuvent être réparties selon la manière d'estimer les dérivées de la fonction d'intensité :

- 1- par différences finies (opérateurs de Roberts, de Prewitt, de Sobel, de Kirch, de Robinson)
- 2- par filtrage optimal (filtres de Shen-C'asian, de Deriche, filtre Gaussien)
- 3- par modélisation de la fonction d'intensité.



**Figure 1.17** : Exemple d'une segmentation par contour

### 4.2.2.3. Segmentation par morphologie mathématique

La morphologie mathématique, dont les principes sont valables tant dans les espaces continus que discrets, a vu, en l'espace d'une quarantaine d'années, son usage se généraliser dans toutes les disciplines requérant des outils d'analyse quantitative ou qualitative : métrologie, astronomie, imagerie biomédicale, aérienne ou satellite.

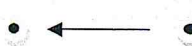





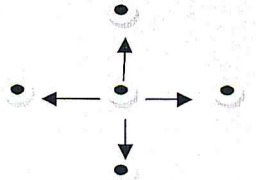
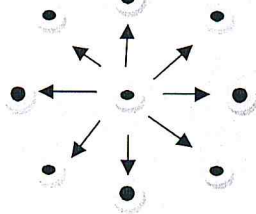
#### . Élément structurant

Un élément structurant, appelé aussi ES, est un ensemble qui possède les caractéristiques suivantes :

## Chapitre 1: Contexte médical

- 1- il possède une forme (géométrie connue)
- 2- cette forme a une taille
- 3- une ou plusieurs directions
- 4- cet élément est repéré par son origine.

L'origine appartient généralement à l'élément structurant mais ce n'est pas une obligation. Les éléments structurants élémentaires les plus usuels sont présentés dans le tableau :

Elément structurant	Notation	Représentation
Le plus proche voisin droit	V1-D	
Le plus proche voisin gauche	V1-G	
Le plus proche voisin haut	V1-H	
Le plus proche voisin bas	V1-B	
Les deux plus proches voisins horizontaux	V2-H	
Les deux plus proches voisins verticaux	V2-V	
Les quatre plus proches voisins	V4	
Les huit plus proches voisins	V8	

**Tableau 1.1:** Les éléments structurants élémentaires les plus usuels

### **5. Conclusion**

Dans ce chapitre, nous avons présenté les principes fondamentaux de la maladie de la rétinopathie diabétique, et la présentation de l'imagerie médicale.

Nous allons montrer dans le cadre de ce travail l'apport des traitements des images numérique sur le dépistage de rétinopathie diabétique.

Nous avons présenté les étapes d'analyse d'images. Le prétraitement permet l'amélioration de la qualité d'image. La segmentation des images permet d'isoler chacun des objets présents dans une image. La phase d'extraction des caractéristiques consiste à calculer un ensemble de mesures permettant de représenter chaque objet de façon unique

Les travaux on montre que l'utilisation de l'image numérique a beaucoup d'avantage

Le cout est moine élève de traitement, la qualité de l'image peut être contrôlée lors de sa prise, Finalement, la transmission dès l'imager est beaucoup plus rapide et facile.

Plusieurs algorithmes de traitement d'image numérique rétinienne ont été développés.

# Chapitre 2

### 1. Introduction

Le Raisonement à Partir de cas (RÀPC) est un paradigme de raisonnement qui s'appuie sur la remémoration de problèmes passés résolus, appelés les cas sources, pour résoudre de nouveaux problèmes, appelés les problèmes cibles (Aamodt & Plaza, 1994).

Ce paradigme a été utilisé dans de nombreux systèmes industriels pour résoudre des problèmes dans des domaines d'application variés. Le RÀPC a souvent été présenté comme une solution au goulet d'étranglement de l'étape d'acquisition des connaissances grâce à l'utilisation de connaissances expérimentales, les cas, qui sont plus faciles à recueillir. Bien que certains arguments présentent le RÀPC comme une solution nettement plus facile à mettre en oeuvre que d'autres systèmes à bases de connaissances, force est de constater qu'en pratique, il est impossible de faire l'économie de l'effort d'acquisition de connaissances. Le développement des systèmes de RÀPC se heurte toujours à un problème d'ingénierie des connaissances. En particulier, les connaissances

Historique :

Inspiré par les travaux de Minsky et Schank réalisés à la fin des années 70, Schank [49] formule pour la première fois le paradigme de raisonnement basé sur les cas. En effet, la théorie développée par Minsky [36] présente un réseau de nœuds et de relations entre ces nœuds ainsi que la notion de « frame (script, schéma) » qui correspond à une structure remémorée qui doit être adaptée pour correspondre à la réalité d'une nouvelle situation rencontrée. Cependant Schank doute de la flexibilité du raisonnement logique et d'une représentation des connaissances ordinaires sous une forme synthétique de propositions indépendamment vraies. Par conséquent, il reprend ces travaux et suppose que le processus de compréhension correspond à un processus d'explication qui s'applique d'une manière itérative [49]. D'ailleurs, Schank est considéré comme l'initiateur du terme « Case Based Reasoning ». Il introduit à travers le modèle de « mémoire dynamique » un degré de généralité varié connu sous le nom de « MOPS (Memory Organization Packets) » constituant un réseau dense d'expériences. De plus, l'auteur tente d'opérationnaliser le comportement humain et l'optimiser si possible. Dans ce cadre, Gebhardt et al. [29] définissent le raisonnement à partir d'expériences comme une façon naturelle de penser caractérisant la réflexion humaine sans doute plus encore que le raisonnement avec des règles. A la fin des années 80, les recherches dans le domaine du RàPC ont réellement commencé à prendre forme et notamment avec les conférences « DARPA » organisées aux Etats-Unis en 1988 [30], avant de s'imposer en Europe avec la première conférence Européenne en 1993 à Kaiserslautern [50], puis avec la première conférence internationale à Lisbonne en 1995 [47].

### 2. Principe de RàPC

#### 2.1. Définition

Le raisonnement à partir de cas (RàPC) (nommé en anglais case-based\_reasoning (CBR)) copie ce comportement humain. Il résout les problèmes en retrouvant des cas analogues dans sa base de connaissances et en les adaptant au cas considéré. Cette technologie est apparue il y a une quinzaine d'années mais les travaux initiaux sur le sujet remontent cependant aux expériences de Schank et Abelson en 1977 à l'Université Yale. Elle reste pourtant encore assez méconnue par rapport à d'autres technologies appartenant au domaine des sciences cognitives comme le data mining. Elle diffère de cette dernière par son approche. En effet, ici, on n'utilise qu'indirectement les données pour retrouver les cas proches, à partir desquels on va générer une solution. [15]

## 2.2. Résonnement par analogie

Conclure à une ressemblance entre une chose connue et une inconnue. Également appelé raisonnement par association d'idées ou raisonnement par analogie

Type particulier de raisonnement inductif, le raisonnement analogique consiste classiquement en la mise en correspondance d'une situation antérieure et d'une situation nouvelle qui lui ressemble, afin de déduire la nature ou des aspects de cette situation nouvelle. On le décompose généralement en 4 étapes principales auxquelles les psychologues ont rajoutés l'apprentissage par analogie, l'un des mécanismes de la catégorisation, organisateur de la pensée, longtemps considéré comme un excellent témoin du développement de l'intelligence.

L'analogie est un mode de raisonnement souvent pratiqué par l'homme. Le RàPC utilise des expériences passés validées pour résoudre les problèmes actuels.

Le RàPC repose sur le principe du raisonnement par analogie. Le principe d'analogie utilisé est le suivant « si deux problèmes sont analogues alors les solutions associées pour les résoudre sont également analogues » [17]

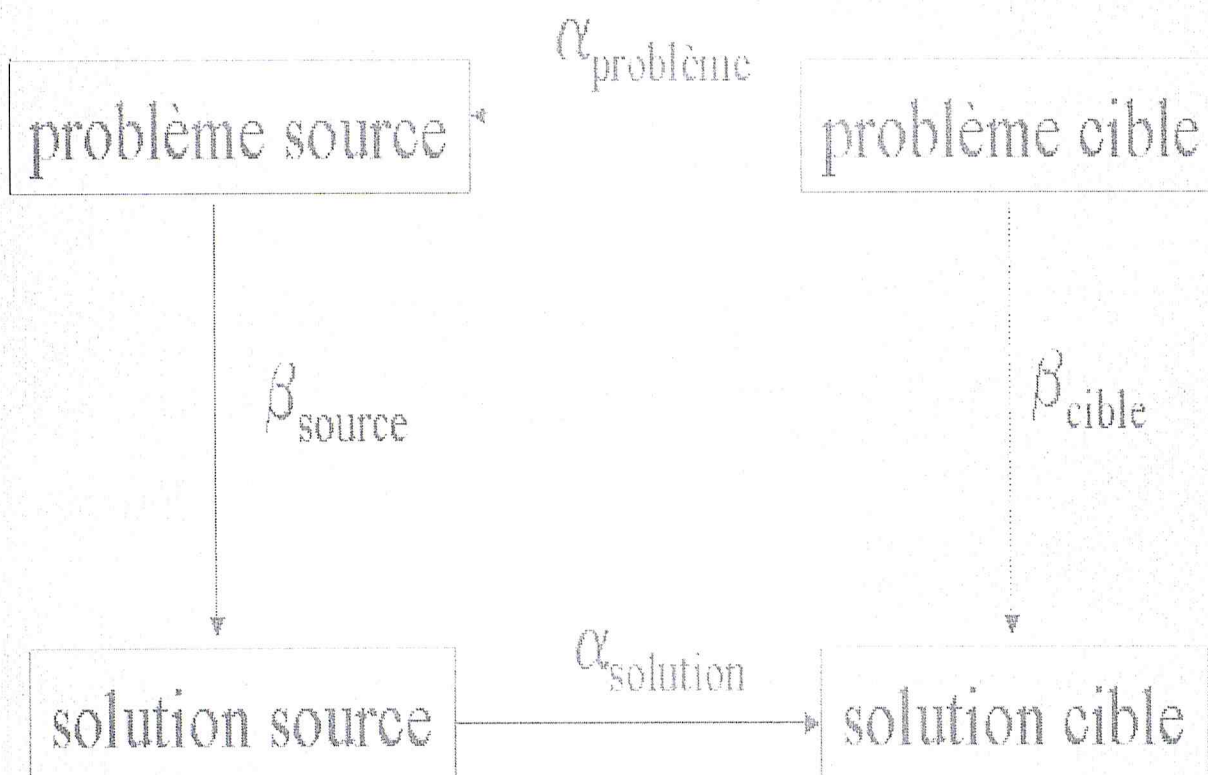


Figure 2.1 : Le carré d'analogie [17]

A partir des valeurs de descripteurs du problème cible et en utilisant la mesure de similarité alpha, un cas source est sélectionné. Les relations de dépendance Béta entre valeurs de descripteurs problème cible et valeur de descripteurs solution permettent de mettre en évidence les descripteurs solution qui doivent être adaptés car dépendant de descripteurs problèmes source différents des descripteurs problèmes cibles. En d'autres termes, si une valeur de descripteur source dépend d'une valeur de descripteur problème, une modification de la valeur du descripteur problème entraînera une modification "analogue" à la dépendance du descripteur solution correspondant. Cette connaissance est nécessaire pour l'adaptation.

En fonction de ces dépendances et des écarts alpha constatés à corriger, l'adaptation permet de proposer une solution cible candidate qui pourra être vérifiée par la vérification de sa conformité aux dépendances particulières qui pourraient exister entre problème et solution cible.

Ce carré d'analogie cache en réalité plusieurs phases dans le raisonnement à partir de cas qui sont particulièrement importantes à identifier dans le cadre de l'ingénierie des connaissances.

### 2.3. Composantes d'un système de Raisonnement à Partir de Cas

Un système RàPC est une combinaison de processus et de connaissances ("knowledge containers") qui permettent de préserver et d'exploiter les expériences passées. Pour simplifier la présentation, nous nous appuyons sur le modèle générique présenté dans la figure (Fig.2.2) les principaux processus dont la recherche ("retrieval"), l'adaptation ("reuse"), la maintenance ("retain") et la construction ("authoring"). Les structures de connaissances sont : le vocabulaire d'indexation, la base de cas, les métriques de similarité et les connaissances d'adaptation [51].

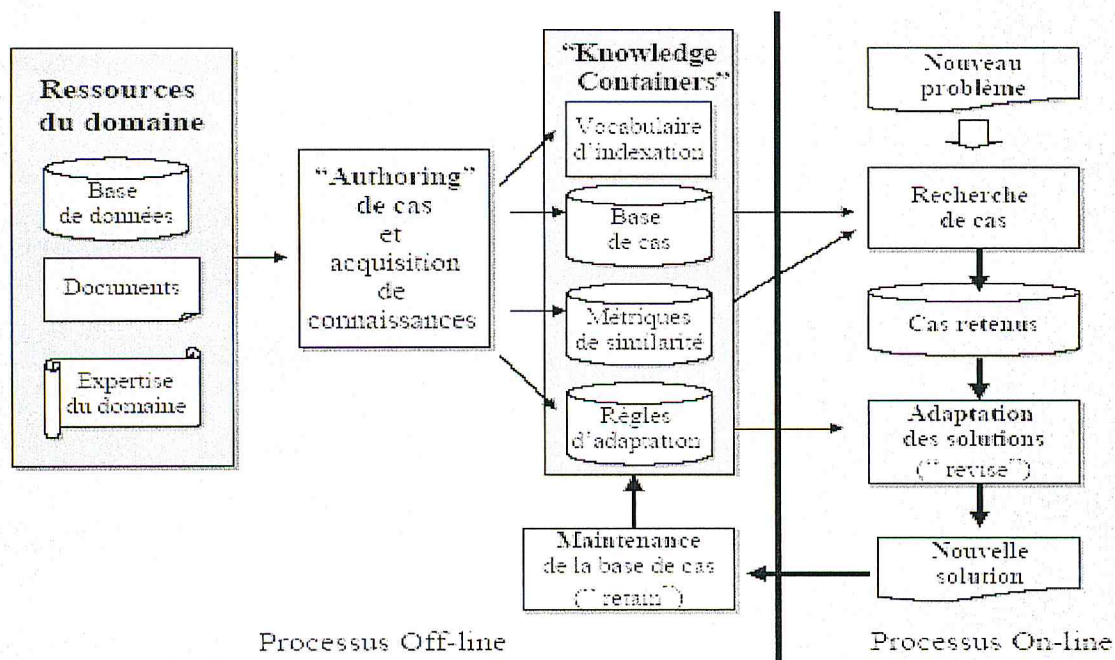


Figure 2.2. Modèle générique d'un système RàPC [52]

### 3. Model de RàPC

#### 3.1. Model simple

Commençons tout d'abord par le modèle le plus simpliste : l'organisation linéaire. Cette organisation n'est pas utilisée pour gérer l'ensemble de la mémoire des cas. Cependant elle peut être implicitement combinée à d'autres modèles plus complexes au niveau de petits sous-ensembles de cas.

Il est possible d'organiser la mémoire sous la forme d'un arbre de décision : chaque nœud correspond à une question sur l'un des index et les fils correspondent aux différentes réponses. Pour être le plus efficace possible l'arbre doit poser les questions dans le bon ordre et être le moins profond possible. Cet arbre doit être construit dynamiquement. La meilleure méthode pour le construire est d'utiliser le data mining.

Un autre modèle consiste à construire la mémoire sous la forme d'une hiérarchie de prototypes. Un prototype permet de décrire des conditions sur des caractéristiques des cas. Tous les cas vérifiant ces conditions sont associés à ce prototype. Les prototypes sont organisés dans une hiérarchie d'héritage. On peut ainsi spécifier des prototypes généraux desquels héritent des prototypes plus spécifiques. En combinant les arbres de décision à cette hiérarchie de prototypes, on obtient une structure intéressante. Les prototypes « terminaux » ne stockent alors plus leurs cas dans une liste mais dans un arbre de décision. La hiérarchie de prototype représente la connaissance a priori du système et les arbres de décision générés dynamiquement permettent une structure assez flexible.

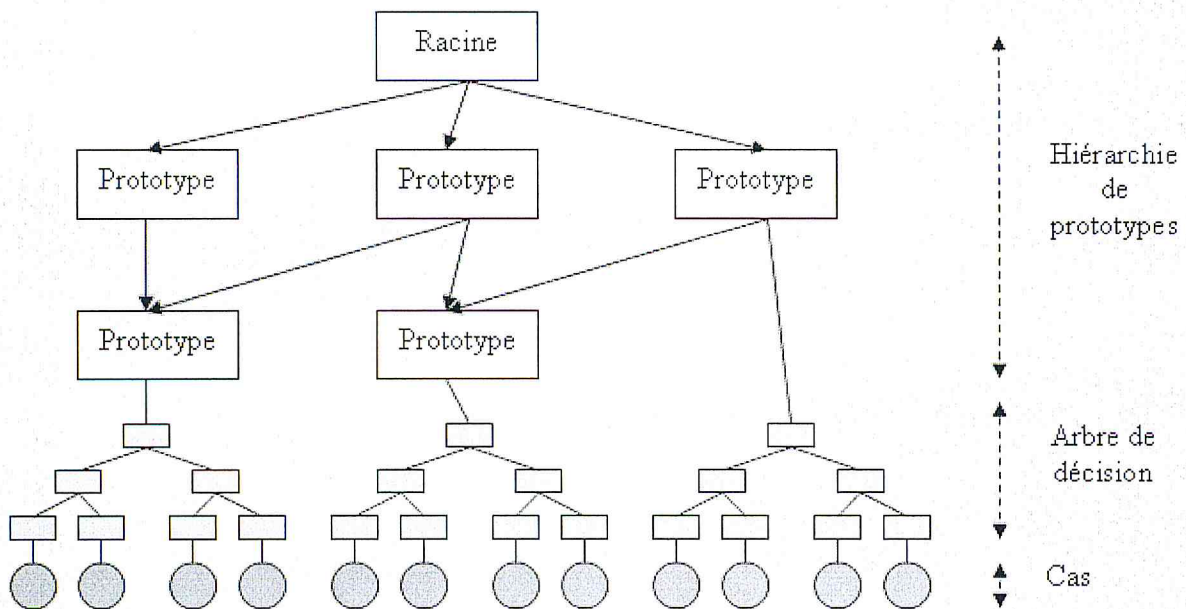


Figure 2.3 : Model simple [15]



### 3.2. Modèle à mémoire dynamique

Le modèle à mémoire dynamique a été introduit par Roger Schank et Janet Kolodner. Dans ce modèle, les cas sont stockés dans une structure hiérarchique appelée épisode généralisé. On parle aussi de MOP pour Memory Organisation Packets. Les différents cas ayant des propriétés similaires sont regroupés dans une structure plus générale, un épisode généralisé. Ils contiennent trois types d'objets :

- \* Les normes : Les caractéristiques communes à chacun des cas indexés sous l'épisode généralisé.
- \* Les index : Les éléments discriminant les cas contenus dans l'épisode généralisé. Un index possède deux champs : son nom et sa valeur. Il peut pointer vers un autre épisode ou simplement vers un cas.
- \* Les cas : La connaissance du système. On y accède donc par l'intermédiaire d'un index.

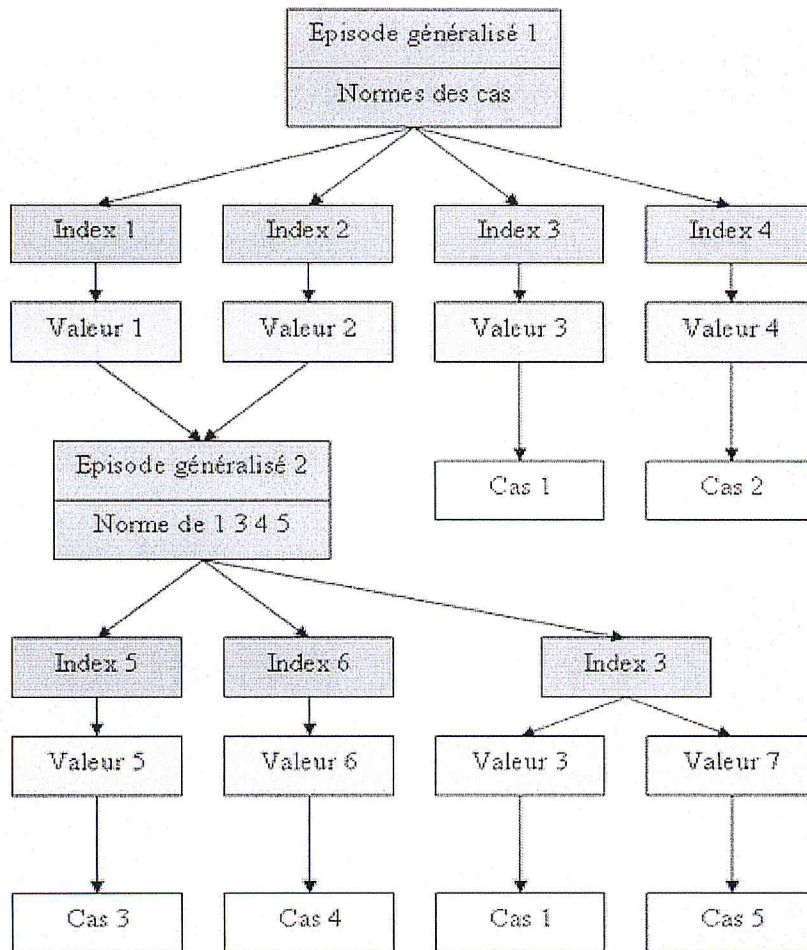


Figure 2.4 : Modèle à mémoire dynamique [15]

Le schéma donne une idée du modèle à mémoire dynamique. Il possède une structure proche d'un arbre. On retrouve bien les trois types d'objets énoncés, à la différence près qu'une distinction est faite entre les index et les valeurs. On peut remarquer aussi qu'il est possible d'atteindre certains cas de différentes manières. Ce modèle est donc redondant.

La recherche des cas similaires s'effectue à partir du nœud racine. On va chercher l'épisode généralisé possédant le plus de caractéristiques en commun avec le problème courant. Ensuite on parcourt les index, qui représentent les caractéristiques absentes de la norme de l'épisode généralisé sur lequel on travaille. Le couple index-valeur sélectionné est celui qui est le plus similaire avec le problème. À

partir de celui-ci, soit on arrive à un autre épisode généralisé, dans ce cas, on recommence le processus, soit on obtient un cas similaire au problème posé.

La procédure d'ajout de nouveaux cas fonctionne d'une manière proche à la recherche de cas similaires. En effet le parcours du graphe est identique. Lorsque l'on a trouvé l'épisode généralisé ayant le plus de normes en commun avec le cas courant, on effectue l'ajout. Pour cela, il faut générer un couple index-valeur distinguant le nouveau cas aux autres fils de l'épisode généralisé. S'il existe déjà un cas possédant le même couple, on crée un nouvel épisode généralisé contenant ces deux cas.

On obtient donc un réseau discriminant à l'aide des index qui permettent de retrouver les cas. Les épisodes généralisés sont principalement des structures d'indexation. Les normes permettent de représenter une connaissance générale des cas sous-jacents alors que les couples index-valeur définissent les spécificités. Cependant, ce processus d'indexation peut mener à une croissance exponentielle du nombre d'index par rapport au nombre de cas. On adjoint donc généralement certaines limites dans le choix des index même si cela entraîne une baisse de performances.

### 3.3. Modèle à base de catégories

Ce modèle est une alternative au modèle précédent. Ici, un cas est aussi appelé exemple. L'idée directrice est que la réalité devrait être définie de manière extensive par des cas. Les caractéristiques décrites généralement par un nom et une valeur, possèdent un niveau d'importance fonction de l'adhésion d'un cas à une catégorie.

Dans ce modèle, la base de cas est un réseau de catégories et de cas. Les index sont des liens qui peuvent être de trois sortes :

- \* De rappel : reliant une caractéristique à une catégorie ou un cas.
- \* D'exemple : reliant une catégorie aux cas auxquels elle est associée.
- \* De différence : reliant deux cas ne différant que d'un nombre restreint de caractéristiques.

Le schéma ci-dessous illustre les différents types de liens disponibles. Cependant il ne représente qu'une seule catégorie. Il faut donc ajouter que les exemples peuvent appartenir à plusieurs catégories.

La recherche des cas similaires consiste à retrouver la catégorie qui possède les caractéristiques les plus proches du nouveau problème. Lorsqu'elle est trouvée, on retourne les cas les plus prototypiques.

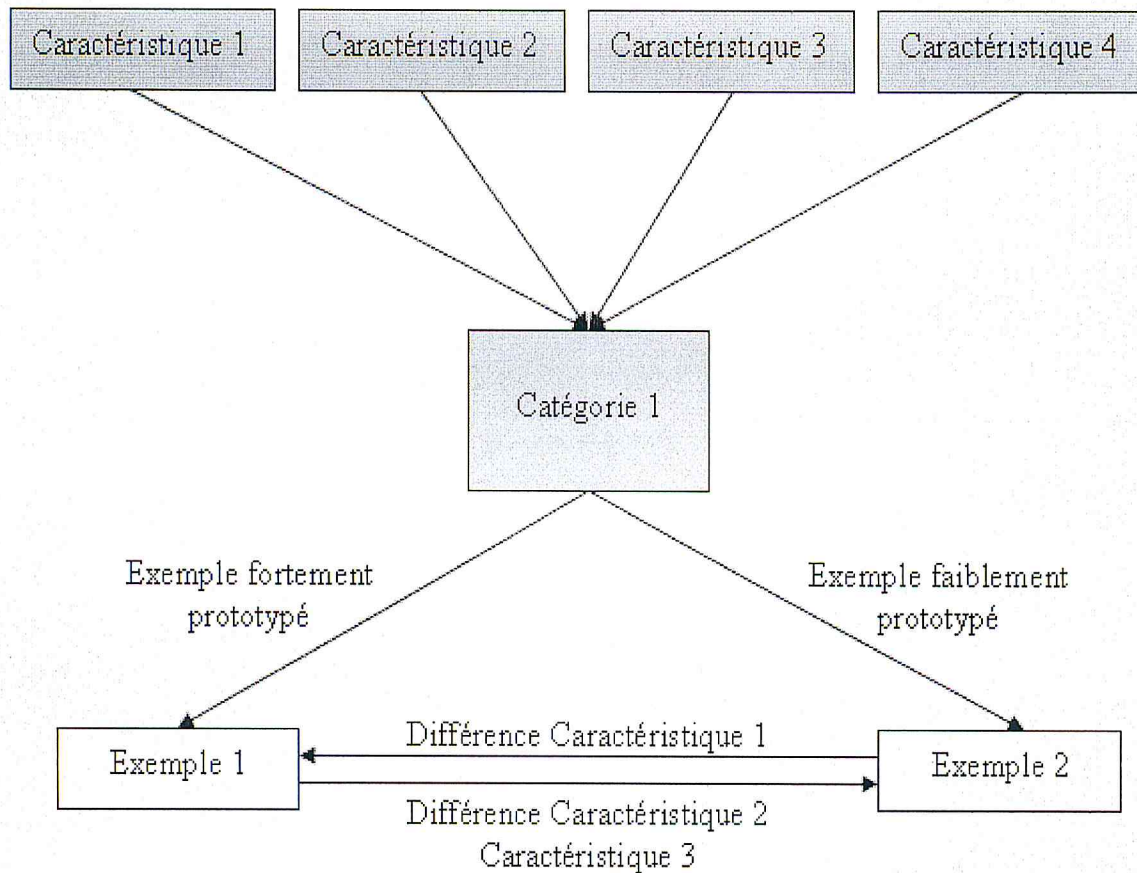


Figure 2.5 : Modèle à base de catégories [15]

#### 4. Cycle de RàPC

Ce sont Aamodt & Plaza qui ont proposé en 1994 un cycle de raisonnement qui a été repris par l'ensemble de la communauté comme une référence. Nous en présentons ici une version légèrement complétée sur la phase d'élaboration d'un cas qui se révèle être un aspect d'ingénierie de la connaissance important.

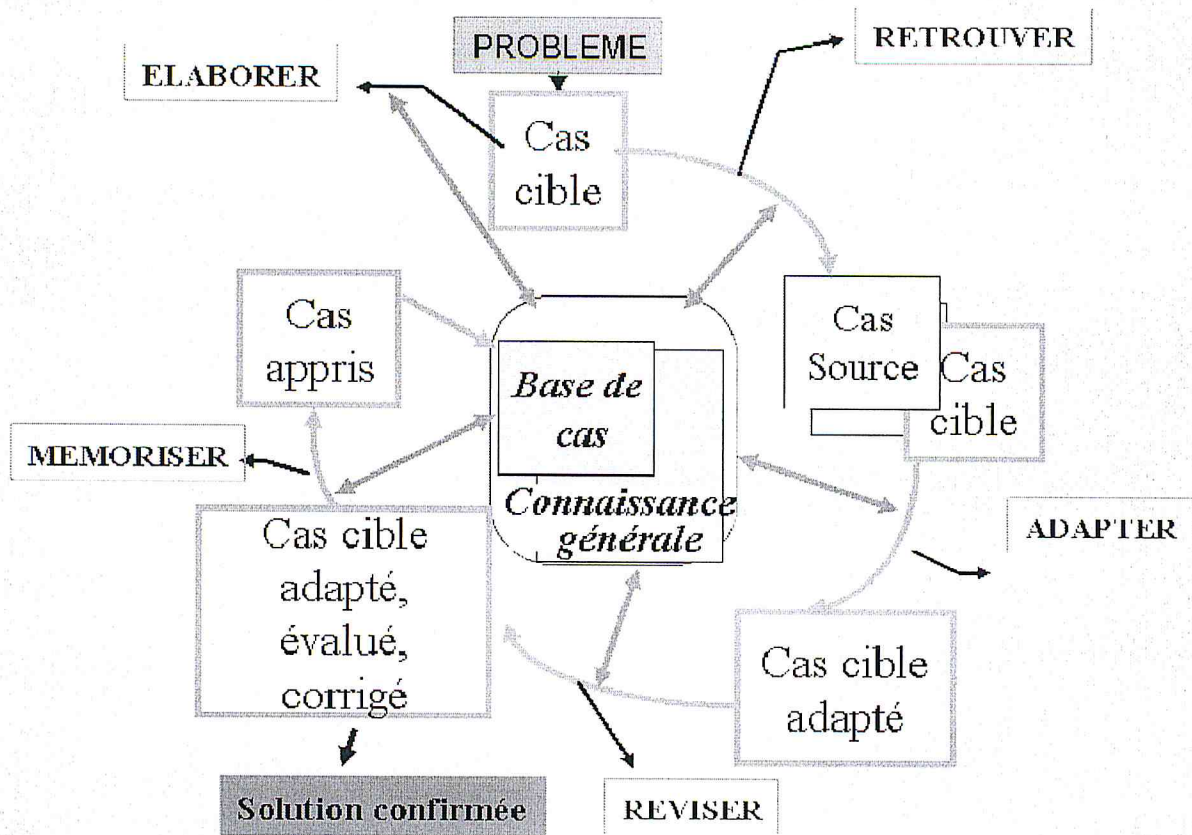


Figure 2-6 : Le cycle d'étapes du Raisonnement à Partir de Cas [15]

#### 4.1. Elaborer

L'élaboration d'un nouveau cas consiste à faciliter la description du problème pour permettre la recherche d'un cas dont la solution sera la plus facilement adaptable. L'adaptabilité d'un cas se mesure à "l'effort" d'adaptation qui sera nécessaire pour que la solution passée (du cas source) puisse servir de base à la solution courante (du cas cible).

La méthode générale consiste à compléter ou filtrer la description d'un problème en se fondant sur la connaissance du domaine pour inférer tout ce qu'il est possible à partir d'une description éventuellement incomplète, et pondérer (mettre des poids) les descripteurs en fonction des dépendances identifiées entre les descripteurs problème cible et les descripteurs de la solution recherchée.

C'est ainsi que l'on va utiliser l'ontologie du domaine pour compléter les descripteurs importants et d'une manière générale, l'on va considérer l'ensemble des règles définies sur le domaine.

#### 4.2. Retrouver

L'étape de recherche de cas sources similaires est naturellement essentielle dans le cycle. Nous rappelons que le cas source qui sera choisi sera normalement le cas ayant la description de

## Chapitre 2 : Raisonnement à partir de cas

problème la plus proche possible de la description du problème cible dans la classe de solution la plus représentée (voir la session 2 de ce cours RàPC pour une illustration graphique de ce principe).

En pratique, il est nécessaire de définir une mesure de similarité qui tiendra compte de l'influence d'une variation d'une valeur de descripteur problème sur une variation de valeur d'un descripteur solution. Intuitivement, on comprend bien qu'il faudra donner un "poids" important aux descripteurs problèmes qui ont beaucoup d'influence sur la solution et une valeur faible aux descripteurs ayant une faible influence sur la solution

Il existe de nombreuses mesures de similarités dans la littérature adaptées à la comparaison d'épisodes de résolution de problème tenant compte de spécificités des descriptions : mesures faisant intervenir le temps, l'espace, des structures complexes, des plans, des séries, etc... Il n'est pas question ici d'en faire le tour, mais il est important de se rappeler qu'il ne faut pas hésiter à étudier cette mesure qui fait partie intégrante de la connaissance nécessaire pour mettre en place un système fondé sur le RàPC.

La distance entre deux descriptions de problèmes est donc constituée de la somme pondérée des distances des attributs décrivant les problèmes. Ce sont les poids  $p_i$  qui portent la connaissance sur l'importance de l'influence du descripteur  $d_i$  sur la solution.

La phase de recherche consiste donc à utiliser ces poids dans la mesure de similarité pour choisir d'une part la classe de solution prédite pour ce cas et d'autre part quel est le cas source de cette classe de solution qui devrait être le plus facile à adapter.

L'algorithme le plus habituellement utilisé pour faire cette recherche est celui des ***K plus proches voisins***.

Remarque importante : il est tout à fait habituel de considérer qu'il y a autant de classes de solutions que de solutions différentes dans la base de cas. Dans ce cas,  $K=1$  et seul le seuil de similarité permet de considérer un cas comme éligible.

Cet algorithme est décrit ci-dessous :

Début

{Initialiser la valeur de T // T est le seuil (Threshold) de similarité minimum pour être considéré comme voisin de ce cas cible

// en toute théorie, il pourrait y avoir autant de seuils différents que de cas cibles

Initialiser la valeur de K // K représente le nombre de cas voisins nécessaire et suffisant pour qu'une classe de solution soit éligible comme candidate pour la solution cible

Initialiser liste\_cas\_voisins // la liste est vide au départ

Tant\_que (il reste des cas à comparer)

{Comparer la description du cas cible avec la description du cas source suivant de la base de cas;

Si (la similarité cas\_cible/cas\_source > T) Alors (Insérer le cas source dans liste\_cas\_voisins);

Passer au cas suivant;

} Fin\_Tant\_que

Choisir la classe solution éligible la plus représentée dans liste\_cas\_voisin;

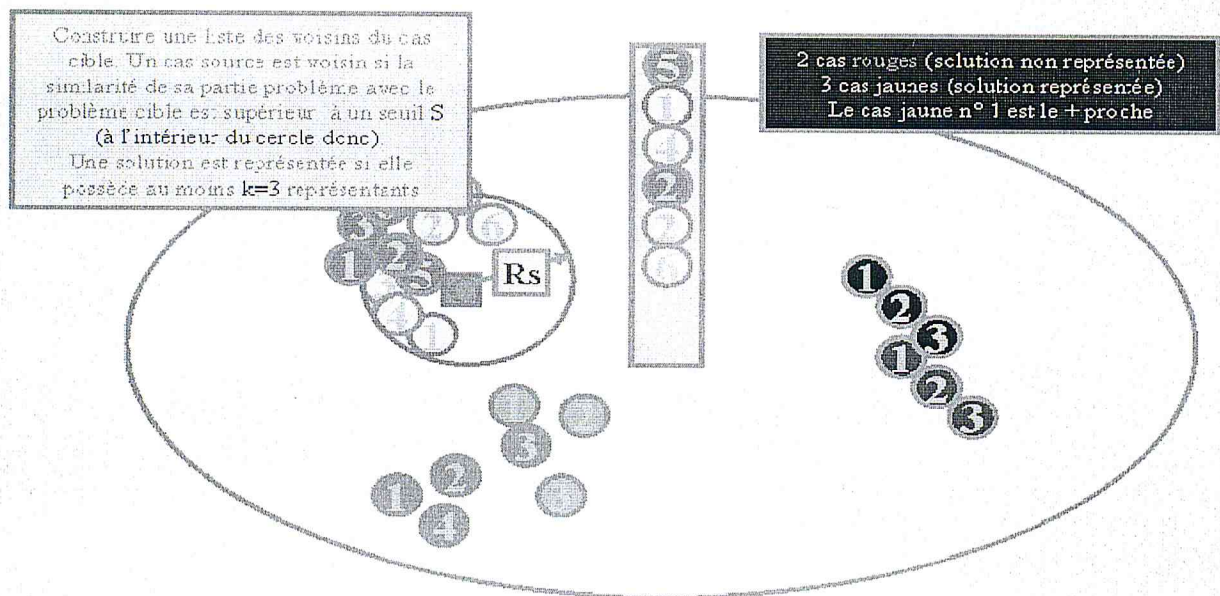
Si (aucune classe de solution éligible) Alors (conclure au manque d'expérience pour ce type de cas cible); FIN;

Choisir le cas source le plus similaire de la classe solution élue;

}

Fin

Plusieurs situations typiques sont illustrées par les figures suivantes (figures 2.7, 2.8 et 2.9). Le seuil  $S$  est illustré par le rayon  $R_s$  de la boule autour de la description du problème cible.



**Figure 2.7 :** Scénario de déroulement de l'algorithme KPPV sélectionnant le cas 1 (de classe solution "jaune") pour le cas cible C. Seule la solution jaune était éligible (représentation > 3).

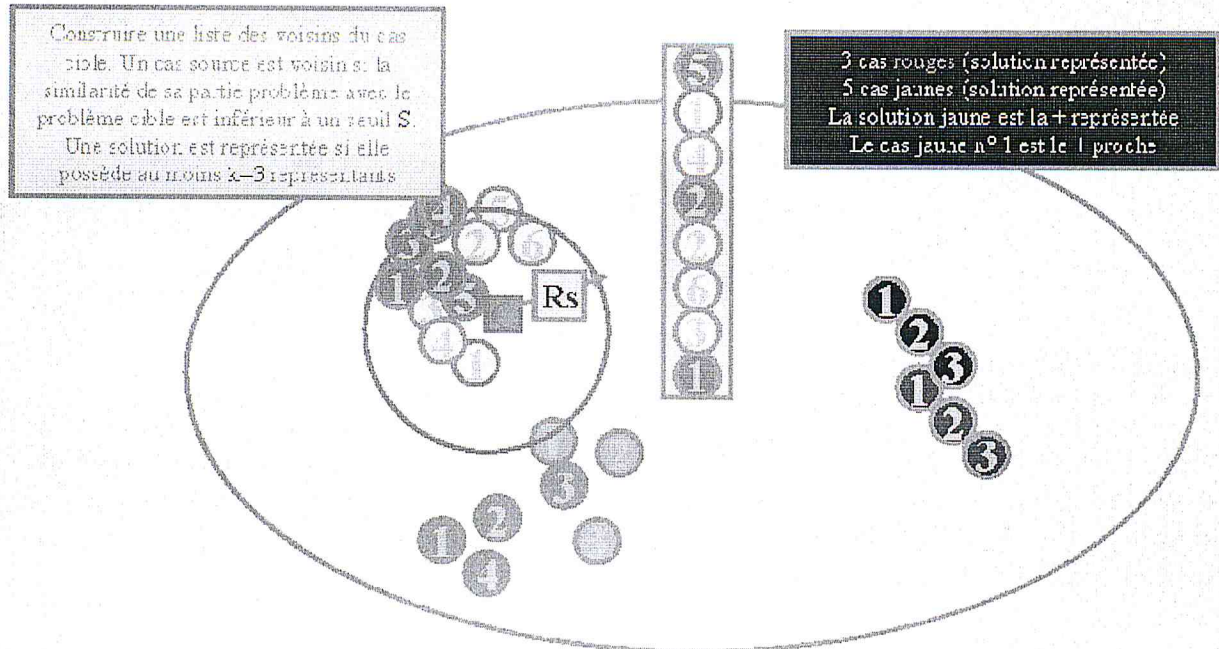


Figure 2.8 : Même scénario qu'en 2-7, mais avec un seuil plus bas ( $R_s$  plus grand) la solution "rouge" devient éligible. Elle n'est pas la plus représentée. Le cas 1 (jaune) est encore sélectionné.

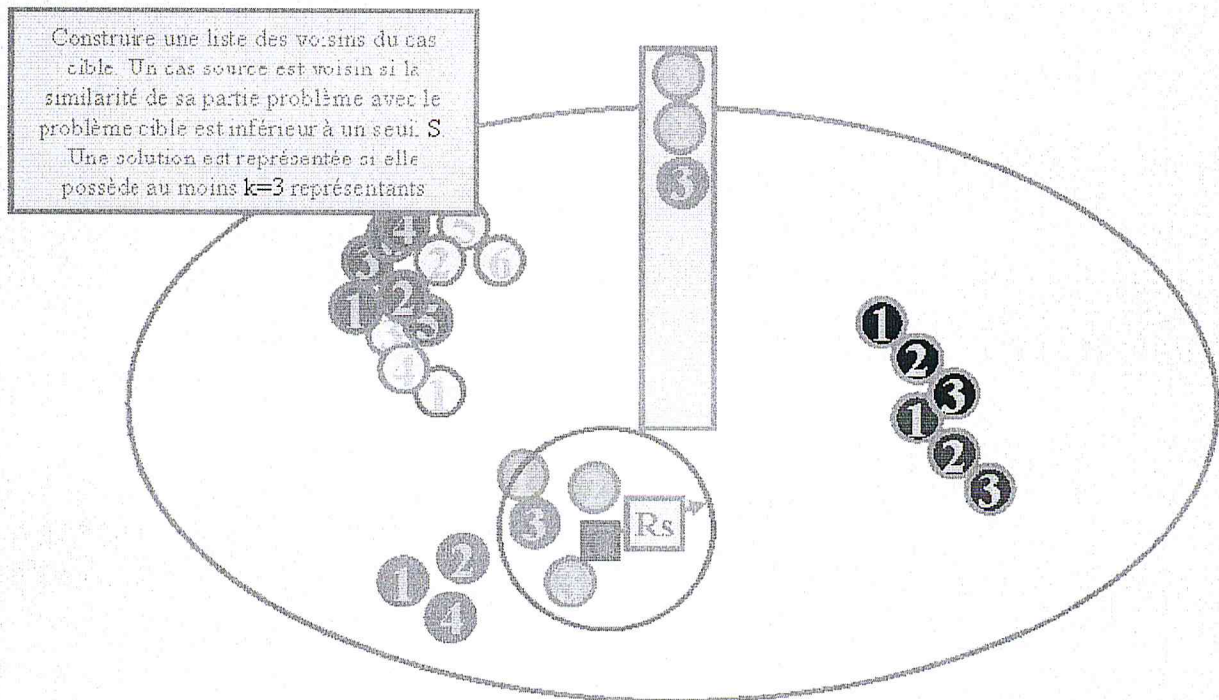
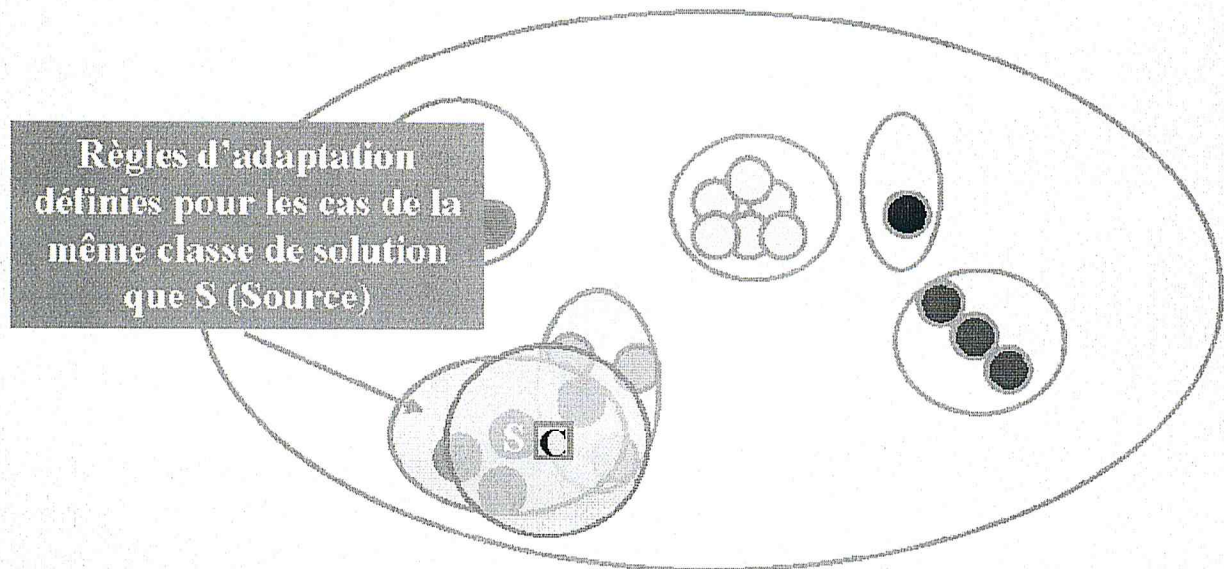


Figure 2.9 : Dans ce scénario, aucune solution n'est éligible. Aucune expérience ne peut être rationnellement remémorée. Ce cas cible devra être résolu par une autre méthode.

### 4.3. Adapter

L'adaptation termine "l'inférence analogique" en calculant quelle pourrait être la solution au problème du cas cible inspirée de la solution du cas source le plus similaire.

La figure 2.10 rappelle qu'une zone d'appartenance à une classe de solution correspond à une zone de règles d'adaptation communes lorsqu'un cas source doit être adapté pour un cas cible



considéré comme similaire en termes de solution potentielle.

**Figure 2.10** : Les règles d'adaptation dépendent de la zone de regroupement des cas d'une même classe de solution. Ce regroupement est précisément réalisé sur la base des connaissances d'adaptation.

Les règles d'adaptation de la solution s'expriment en fonction des écarts relevés entre les descriptions des problèmes cibles et source.

Le schéma suivant (figure 2.11) illustre les connaissances et l'inférence analogique ainsi menée~:



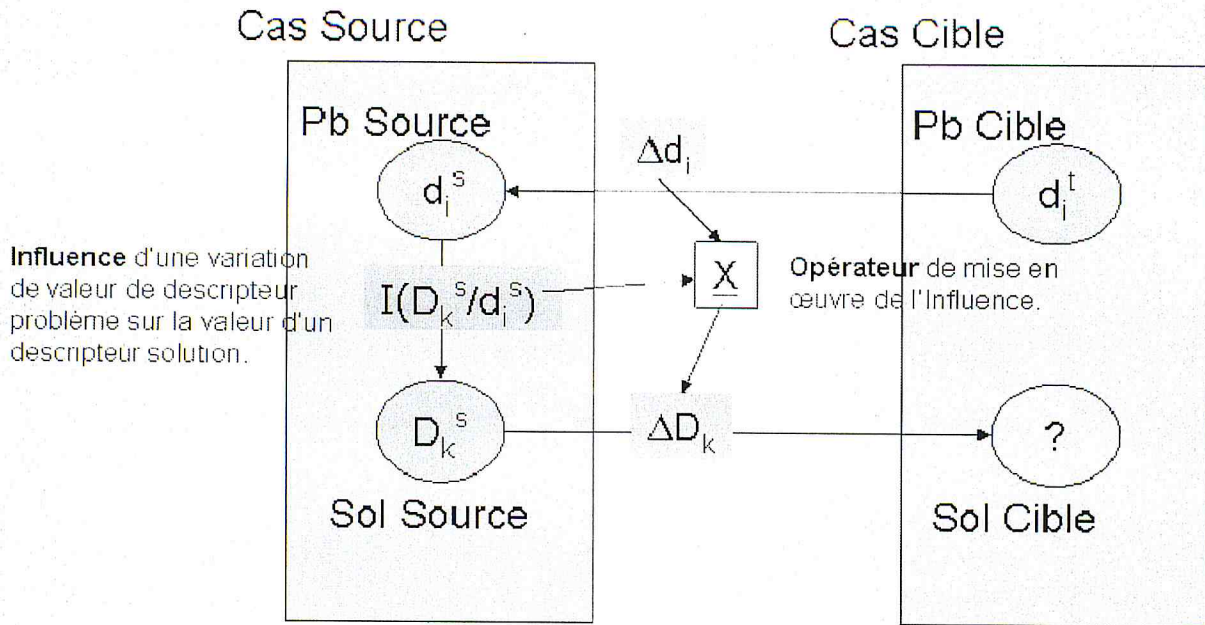


Figure 2.11 : Schéma général des connaissances et opérateurs d'adaptation

Dans le cas d'une variation simple, le calcul d'un descripteur solution cible à partir de descripteurs solution et selon l'influence des écarts sur les descripteurs problème correspondent aux formules suivantes :

### Cas d'une variation locale simple

$$D_k^t = D_k^s + \Delta D_k$$

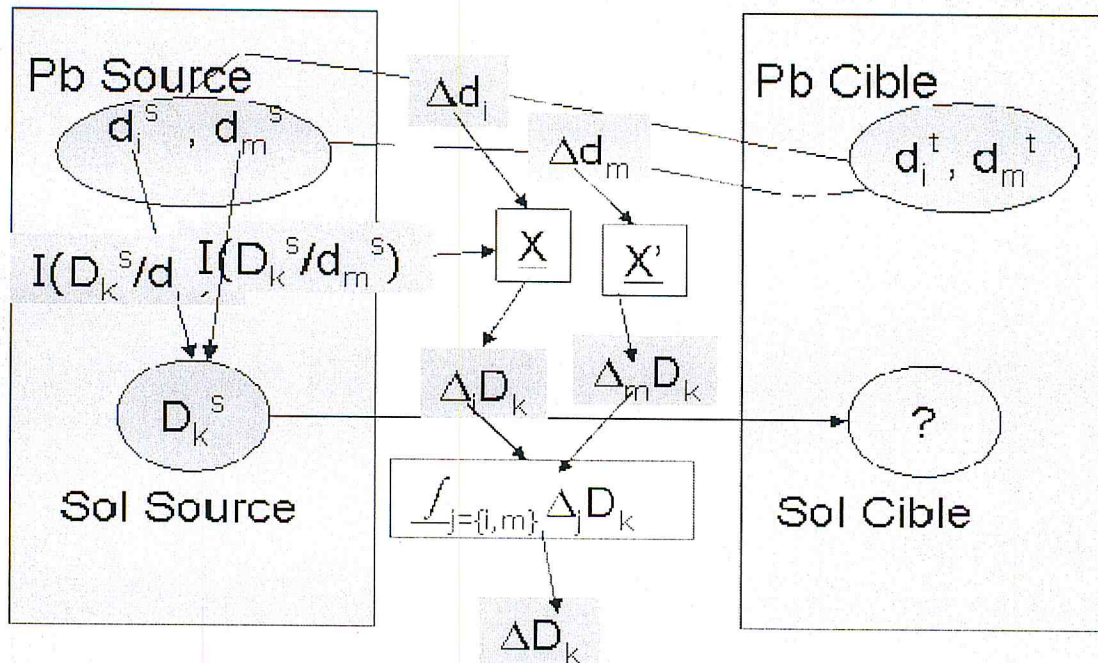
$$D_k^t = D_k^s + I(D_k^s/d_i^s) \times \Delta d_i$$

Cette formule signifie simplement qu'il faudra calculer un écart à appliquer au descripteur solution qui dépendra directement de l'influence d'un écart élémentaire entre les descripteurs cible et source correspondant sur la valeur du descripteur source, cette influence dépendant naturellement de l'importance de l'écart.

Bien entendu, un descripteur solution ne dépend pas en général d'un seul descripteur problème et le schéma plus général est alors le suivant :

Cas Source

Cas Cible



**Cas d'une variation globale**

$$D_k^t = D_k^s \pm \Delta D_k$$

$$D_k^t = D_k^s \pm \sum_{j=\{i,m\}} (I(D_k^s/d_i^s) \underline{X} \Delta d_j)$$

Figure 2.12 : Schéma général d'adaptation d'un cas source en fonction des influences pb/sol et des opérateurs d'adaptation associés avec

- $\Delta d_i$  = écart entre les valeurs de descripteurs problème selon une fonction d'appariement spécifique. Le résultat de la fonction d'appariement doit donner cet écart comme information de départ.
- $I(D_k^s/d_i^s)$  = Influence de la valeur de  $d_i^s$  sur la valeur de  $D_k^s$ . Ce peut être une simple proportion... mais n'importe quoi d'autre qui présente un « rapport » entre la valeur d'un descripteur problème avec un descripteur solution.
- $\underline{X}$  = opérateur de mise en oeuvre de l'Influence en fonction de l'écart observé sur le problème. Cet opérateur peut-être vu comme « mutlipliant » l'effet de l'influence par l'écart constaté.
- $\sum_{j=\{i,m\}}$  = opérateur de « sommation » des effets des différentes Influences appliquées aux écarts de problème constaté (pour un descripteur solution donné)
- $\oplus$  = opérateur « d'ajout » des effets des écarts problème sur le descripteur solution source pour fournir une valeur candidate pour le descripteur solution cible

Pour faciliter les formules d'adaptation, il est courant de tenter de passer du symbolique au numérique pour exprimer des différences. Par exemple pour exprimer une différence entre mauvais et très bon, on aurait pu exprimer une échelle d'état entre 1 et 10 et rendre notre règle plus simple à exprimer.

A l'issue de l'adaptation, une solution hypothétique est donc proposée. Sa valeur n'est pas garantie puisqu'il s'agit d'un raisonnement analogique sur l'expérience et bien entendu les connaissances utilisées pour décrire ces connaissances d'adaptation peuvent avoir à être révisées.

Une originalité importante du RàPC est que le cycle de raisonnement prévoit explicitement cette révision.

### 4.4. Réviser

La révision consiste à continuer éventuellement l'élaboration de la solution cible en cas de besoin. Pour réviser, on peut :

- essayer la solution dans le monde "réel". Dans l'illustration des véhicules d'occasion, on essaye de vendre son véhicule au prix résultant de l'adaptation.
- faire une introspection dans la base de cas en utilisant cette fois-ci l'ensemble des descripteurs problème + solution pour vérifier que les cas similaires ont bien donné toute satisfaction. Par exemple, on vérifierait qu'il y a bien eu des véhicules similaires vendus au prix calculé.
- utiliser une autre méthode d'évaluation de la solution (simulateur, système expert classique, etc...)

Dans tous les cas, en cas de divergence entre ce qui a été construit comme solution et ce qu'il aurait fallu faire, il est important :

- de corriger, de façon à ce que le cas soit un cas avec une solution correcte ;
- éventuellement, tracer l'écart et l'expliquer sous forme d'exception (ce qui revient à particulariser une forme d'adaptation à une zone plus petite pour la classe de solutions)

Ce cas, avec les explications de révision, devient une source d'apprentissage importante pour faire évoluer les connaissances mobilisées par le raisonnement. Il n'existe pas encore de méthode standardisée pour rendre compte de la tâche de révision dans le RàPC.

### 4.5. Mémoriser

La mémorisation du nouveau cas résolu est l'occasion d'enrichir la base de connaissances :

- naturellement, un cas supplémentaire résolu augmente l'expérience du système,
- en cas de révision, la connaissance générale peut être modifiée et en particulier les connaissances duales liées aux tâches "retrouver" et "adapter".
  - la mesure de similarité peut être raffinée pour éviter de sélectionner une classe de solution erronée ;
  - les connaissances d'influences peuvent être raffinées pour piloter l'adaptation. Ces connaissances d'influence sont duales des connaissances de similarité ; elles sont directement liées aux poids utilisés pour pondérer la mesure globale de similarité.
  - des dépendances nouvelles peuvent être découvertes, etc...
- il peut être utile de garder la "trace" de l'ensemble du cycle avec le détail des corrections faites. Même si on n'a pas encore pu mettre à jour les connaissances du système, cette trace

pourra être utilisée pour considérer ce cas comme un modèle pour "corriger" en s'inspirant de cette correction une nouvelle adaptation qui se ferait avec le même type de similarité (on découvrirait ce cas résolu et corrigé au moment de la révision).

### 5. Connaissances dans un système RàPC

Richter [RIC 98] définit les systèmes de RàPC comme des systèmes à base de connaissances (SBC). Ces systèmes exploitent quatre catégories de connaissances ("knowledge containers") distinctes à savoir :

- Vocabulaire d'indexation : un ensemble d'attributs ou de traits ("features") qui caractérisent la description de problèmes et de solutions du domaine. Ces attributs sont utilisés pour construire la base de cas et jouent un rôle important lors de la phase de recherche.
- base de cas : l'ensemble des expériences structurées qui seront exploitées par les phases de recherche, d'adaptation et de maintenance.
- mesures de similarité : des fonctions pour évaluer la similarité entre deux ou plusieurs cas. Ces mesures sont définies en fonction des traits et sont utilisées pour la recherche dans la base de cas.
- Connaissances d'adaptation : des heuristiques du domaine, habituellement sous forme de règle traite permettant de modifier les solutions et d'évaluer leur applicabilité à de nouvelles situations.

### 6. Les principaux systèmes RàPC en diagnostic médical

En médecine, le RàPC a été utilisé à de nombreuses reprises pour la conception d'outils d'aide à la décision. En effet, la médecine est un domaine à fort potentiel pour ce qui est de la prise de décision, les cas illustrant des exemples de prises de décisions médicales sont assez répandus. Il existe un lien intense entre la médecine et le raisonnement à partir de cas [27]. Ce dernier étant un paradigme bien adapté pour le développement d'un système d'aide à la prise de décision médicale. Certainement, un des dispositifs intuitivement attrayants des systèmes RàPC dans la médecine est que les concepts du patient et de la maladie se prêtent naturellement à une représentation de cas [38]. De nombreux autres exemples de systèmes de RàPC médicaux sont proposés dans les ouvrages de référence sur le RàPC dans [48] [53]. Dans la suite, nous décrivons quelques systèmes RàPC de diagnostic médical les plus connus.

#### 6.1. Système CASEY

Le système d'aide au diagnostic « CASEY » est réalisé par Koton en 1988 [31]. L'entrée de ce système est une description du nouveau patient et qui peut être des signes normaux, des signes présents et/ou des symptômes. La sortie est une explication causale de la maladie. Cette explication relie les symptômes à l'état interne. CASEY procède au diagnostic des patients en appliquant un appariement basé modèle et une adaptation heuristique. Sa base de cas contient 25 cas tous diagnostiqués par le programme de détection des problèmes cardiaques (HFP pour Heart Failure Program) qui est un raisonneur basé modèle (Fig.1.6).

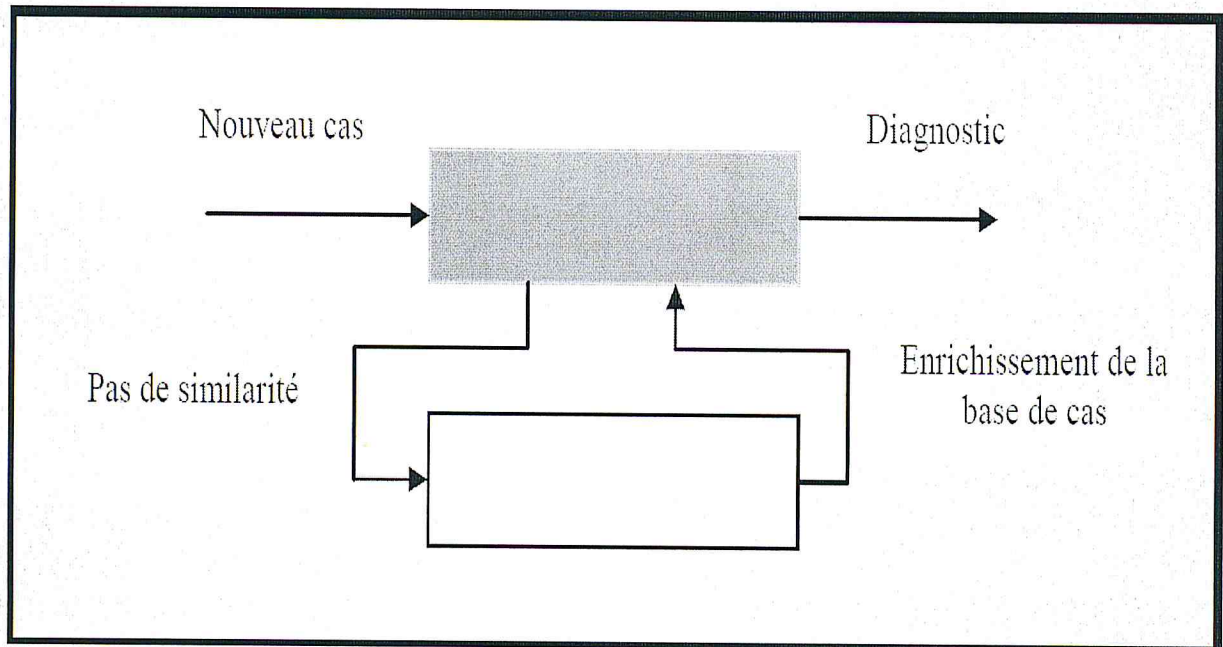


Figure 2.13 : Coopération CASEY/ HFP [53]

Lorsqu'un nouveau patient se présente, CASEY cherche un cas similaire. S'il en trouve, il l'utilise pour le diagnostic du nouveau patient. S'il n'en trouve pas, il soumet le nouveau cas au HFP qui le diagnostique et retourne les résultats à CASEY pour une utilisation future (enrichissement de la base). CASEY utilise un processus à deux étapes pour diagnostiquer un nouveau cas. Il procède à une recherche des cas similaires en utilisant des règles d'évidence pour déterminer parmi les cas s'appariant partiellement quels sont ceux suffisamment similaires pour proposer un diagnostic précis. Puis, il applique des règles de réparation pour adapter l'ancien diagnostic à la nouvelle situation. CASEY utilise, pour son adaptation, un ensemble de règles de réparation basées modèle indépendantes du domaine. Chaque règle de réparation est associée à une ou plusieurs règles d'évidence. CASEY garde la trace de la règle d'évidence appliquée et active la règle de réparation correspondante lors de l'adaptation.

### 6.2. Système PROTOS

PROTOS [22] est un système de RàPC pour l'aide au diagnostic des malentendants. Le cas est représenté par un vecteur de caractéristiques. Il n'y a aucun modèle utilisé et la phase de remémoration est basée sur l'algorithme des k plus proches voisins. Protos ne dispose pas de phase d'adaptation. Si la solution proposée par le système n'est pas satisfaisante alors une phase d'apprentissage s'enclenche. De ce fait, le système commence à rechercher une autre solution sinon il doit accepter une solution proposée par l'utilisateur. Dans ce cas, l'utilisateur est obligé de définir les termes dans le système qui sont inconnus de PROTOS, en décrivant leur relation avec les termes existants. Cette phase permet de construire la connaissance générale du système. Si le problème a été résolu tout de suite, le système renforce (par pondération) les caractéristiques ayant permis la résolution. Parallèlement, si le problème n'est résolu qu'en seconde étape, PROTOS affaiblit les caractéristiques fautive. L'originalité de PROTOS tient au fait que son apprentissage repose sur l'utilisateur.

### 6.3. Système IDEM

IDEM (Images et Diagnostic par l'Exemple en Médecine) est un système développé en France par le service d'informatique médicale de l'Hôpital Broussais à Paris. Ce système est développé par Bozec [24] pour l'aide au diagnostic en imagerie médicale dans le domaine de l'anatomie pathologique mammaire. Les trois fonctionnalités principales du système sont : □ Se familiariser avec le système. □ Naviguer dans un ATLAS de cas de pathologie tumorale mammaire. □ Interroger la base de cas. Les cas sont des exemples d'interprétations d'images macroscopiques et microscopiques. Pour modéliser ces cas, les auteurs se sont basés sur des comptes-rendus d'experts, et en ont obtenu une représentation composite. Un cas est alors composé d'une arborescence de «zones», elles-mêmes composées d'un ensemble de couples d'attribut : valeur>. La mesure de similarité préalablement définie, opère à trois niveaux: au niveau de l'attribut, au niveau de la zone et enfin au niveau du cas. Au niveau de la zone et du cas, la mesure de similarité intègre une ressemblance de surface et une ressemblance de structure. La mesure globale de ressemblance est calculée comme la moyenne arithmétique d'une similarité de surface et d'une similarité de structure

## 7. Traitement d'image et le RàPC

Le RàPC standard est généralement réalisé à des données tabulaires. Les recherches actuelles sur le RàPC vise à élargir le champ d'application de la technologie en l'appliquant à d'autres formes de données comme des images, son, vidéo, etc. Il y a deux questions principales qui doivent être examinées lors de l'application du RàPC à des données non-standard. La première question quelle est la meilleure façon de représenter la variable d'entrée afin de faciliter le RàPC ? La seconde question quelle est la nature du mécanisme de vérification de similitude à appliquer ? Les deux questions sont étroitement liées [25].

Le RàPC et le traitement d'image peuvent être intégrés dans des domaines où le raisonnement et la prise de décision repose largement sur des informations visuelles ou spatiales dérivées à partir d'images. Des exemples de ces domaines sont la médecine, où l'information provenant d'images IRM, d'images angiographiques rétiniennes, etc. Ainsi les systèmes d'informations géographiques, où l'information sur l'image satellitaire joue un rôle crucial. Un autre champ d'application est la biologie moléculaire, où la compréhension de la géométrie à trois dimensions d'une structure moléculaire est souvent indispensable pour résolution des problèmes.

## 8. Application du RàPC dans l'imagerie médicale

Le [28], cite huit travaux qui combinent le RàPC et l'imagerie médicale, ces travaux sont illustrés dans le tableau 2.1 selon l'utilisation de la phase adaptation, phase recherche, la segmentation d'image et l'interprétation d'image.

Author, year	Case adaptation	Case retrieval	Image segment	Image interpret
Alexandrini et al., 2003	No	Yes	No	No
Balaa et al., 2003	No	Yes	No	No
Berger, 1992	Yes	Yes	No	Yes
Grimnes and Aamodt, 1996	Yes	Yes	Yes	Yes
Galushka et al., 2005	Yes	Yes	Yes	Yes

## Chapitre 2 : Raisonnement à partir de cas

<b>Golobardes et al., 2002</b>	No	Yes	Yes	Yes
<b>Haddad et al., 1997</b>	Yes	Yes	Yes	Yes
<b>Jurisica, and Glasgow, 2003</b>	No	Yes	Yes	Yes
<b>Lobozek et al., 1998</b>	No	Yes	No	No
<b>Perner et al., 1999</b>	No	Yes	Yes	Yes
<b>Perner et al., 2003</b>	No	No	No	Yes
<b>Wilson et al., 2006</b>	No	Yes	No	No
<b>Yearwood and Pham. 2000</b>	No	Yes	No	No

**Tableau 2.1** : Adaptation de cas, recherche de cas, segmentation d'image et image interprétation [28]

Dans le tableau 2.2, les mêmes travaux sont illustrés selon les caractéristiques utilisées couleur, forme, contour, texture.

<b>Author, year</b>	<b>Case adaptation</b>	<b>Case retrieval</b>	<b>Image segment</b>	<b>Image interpret</b>
<b>Alexandrini et al., 2003</b>	No	No	No	No
<b>Balaa et al., 2003</b>	No	No	No	No
<b>Berger, 1992</b>	No	Yes	No	No
<b>Grimnes and Aamodt, 1996</b>	No	No	No	No
<b>Galushka et al., 2005</b>	Yes	No	No	Yes
<b>Golobardes et al., 2002</b>	No	No	No	No
<b>Haddad et al., 1997</b>	No	No	No	No
<b>Jurisica, and Glasgow, 2003</b>	No	Yes	Yes	No
<b>Lobozek et al., 1998</b>	No	No	No	No
<b>Perner et al., 1999</b>	Yes	No	No	No
<b>Perner et al., 2003</b>	Yes	Yes	No	Yes
<b>Wilson et al., 2006</b>	No	No	Yes	No
<b>Yearwood and Pham. 2000</b>	No	Yes	No	No

**Tableau 2.2** : Couleur, forme, contour, texture [28]

## **9. Conclusion**

Le raisonnement à partir de cas est sans conteste, le processus le plus proche du raisonnement humain dans sa capacité de résoudre les problèmes de la vie courante. Ainsi comme le raisonnement humain, le raisonnement à partir de cas suit la même démarche, l'identification du problème à résoudre (étape d'élaboration), la recherche d'une ou plusieurs situations similaires déjà résolues (remémoration), l'adaptation des solutions de ces situations, la révision de la solution retenue pour l'améliorer et enfin la mémorisation de cette situation pour les problèmes futurs.

Toutes ces étapes supposent l'existence d'une base de cas bien organisée servant aux différentes étapes du processus. L'exploitation des expériences capitalisées de cette base de cas dans l'analyse d'image peut apporter une aide importante pour le diagnostic automatique d'imagerie médicale.



# Chapitre 3

### 1. Introduction

Un diagnostic médical est le résultat du raisonnement d'un médecin, décision très souvent prise à partir d'informations incertaines et/ou incomplètes. De nombreuses techniques d'intelligence artificielle ont été appliquées pour essayer de modéliser ce raisonnement. Citons, par exemple, des systèmes à base de règles comme MYCIN et Internist-1/QMR (Quick Medical Reference) [32] [46].

En amont de ce raisonnement, il faut aussi être capable de modéliser ces informations incertaines et/ou incomplètes. Certaines approches ont utilisé des formalismes comme la logique floue ou les fonctions de croyance de Dempster-Shafer. Une autre consiste à se placer dans le cadre de la théorie des probabilités, ce qui nous amène tout naturellement aux réseaux bayésiens proposés par Pearl dans les années 80 [39], retrouvés parfois sous le nom de systèmes experts probabilistes.

Les réseaux bayésiens constituent une technique d'acquisition, de représentation et de manipulation de connaissance. Toute application mettant en œuvre des connaissances peut relever de l'utilisation des réseaux bayésiens, qu'il s'agisse de formaliser la connaissance d'experts, d'extraire la connaissance contenue dans des bases de données, ou d'utiliser le plus rationnellement possible l'un ou l'autre type de connaissance. Les réseaux bayésiens sont utilisés pour leur capacité d'effectuer des inférences dans un contexte d'incertitude, en quelque sorte comme alternative aux systèmes experts. Comme ils sont utilisés pour leurs algorithmes d'apprentissage et d'inférence, comme alternative aux autres méthodes de modélisation quantitative, en les considérant comme des modèles de régression.

L'utilisation des réseaux bayésiens pose un certain nombre de questions méthodologiques [33] :

- . Comment choisir la structure du réseau bayésien ?
- . Comment estimer les probabilités du réseau ?
- . Comment prendre en compte les données incomplètes ou les variables latentes ?
- . Comment faire de l'inférence, exemple: calculer la probabilité de telle ou telle maladie sachant certains symptômes ?

Dans ce chapitre nous allons définir les réseaux bayésiens. Nous présentons les principales définitions, les notions d'indépendance conditionnelle, les principaux théorèmes et les étapes de la construction des réseaux bayésiens. Ensuite, nous élaborons les deux algorithmes d'inférences exactes à savoir : l'algorithme Pearl et l'algorithme de Jensen Lauritzen Olesen (JLO). Ce chapitre donne une idée détaillée sur les réseaux bayésiens, il aborde les points pertinents pour comprendre cette méthode.

### 2. Définition

Un réseau bayésien est un système représentant la connaissance et permettant de calculer des probabilités conditionnelles apportant des solutions à différentes sortes de problématiques. La structure de ce type de réseau est simple : un graphe dans lequel les nœuds représentent des variables aléatoires, et les arcs (le graphe est donc orienté) reliant ces dernières sont rattachées à des probabilités conditionnelles. Notons que le graphe est acyclique : il ne contient pas de boucle. Les arcs représentent des relations entre variables qui sont soit déterministes, soit probabilistes. Ainsi, l'observation d'une ou plusieurs causes n'entraîne pas systématiquement l'effet ou les effets qui en dépendent, mais modifie seulement la probabilité de les observer. L'intérêt particulier des réseaux bayésiens est de tenir compte simultanément de connaissances a priori d'experts (dans le graphe) et de l'expérience contenue dans les données.

Les domaines d'utilisation principaux sont: diagnostic (médical et industriel), analyse de risques, détection de spam, datamining, détection de fraudes, exploitation du retour d'expérience, modélisation et simulation de systèmes complexes, détection d'intrusions, TextMining, analyse de Bio Puces, analyse de trajectoires de santé. En bref, un réseau bayésien est un modèle probabiliste graphique permettant d'acquérir, de capitaliser et d'exploiter des connaissances, né du besoin de créer des systèmes experts à base de probabilités [35].

### 3. Définition Formelle

Un réseau bayésien peut être formellement défini 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 probabilisé fini  $(W, p)$ .
- un ensemble de variables aléatoires associées aux nœuds du graphe et définies sur  $[\Omega, p]$  tel que :

$$p(V_1, V_2, \dots, V_n) = \prod_{i=1}^n p(V_i | C(V_i)) \quad (1)$$

Avec  $C(V_i)$  l'ensemble des parents de  $V_i$  dans le graphe [BOU 05].

### 4. Représentation graphique de la causalité

Un réseau bayésien (réseau probabiliste ou Bayésien Network) est un modèle représentant des connaissances incertaines sur un phénomène complexe, et permettant, à partir des données, un véritable raisonnement. Un réseau bayésien a pour objectif d'acquérir, de représenter et d'utiliser la connaissance. Il est constitué de deux composantes [25]:

- un graphe causal, orienté, acyclique, dont les nœuds sont des variables d'intérêt du domaine, les arcs des relations de dépendance entre ces variables. L'ensemble des nœuds et des arcs forme ce que l'on appelle la structure du réseau bayésien. C'est la représentation qualitative de la connaissance.
- un ensemble de distributions locales de probabilités qui sont les paramètres du réseau. Pour chaque nœud on dispose d'une table de probabilités  $P(\text{variable} | \text{parents}(\text{variable}))$  qui représente la distribution locale de probabilité. Il faut remarquer que l'état de chaque nœud ne dépend que de l'état de ses parents. Il s'agit de la représentation quantitative de la connaissance.

On peut décrire un réseau bayésien comme un système expert probabiliste. Dans un réseau bayésien, un arc de  $A$  vers  $B$  peut être interprété par '  $A$  cause  $B$ ', les cycles ne sont pas autorisés, et le graphe est un graphe acyclique orienté. De plus un nœud est conditionnellement indépendant de ses non-descendants sachant ses parents.

### 4.1. Exemple

Nous allons commencer par l'exemple suivant [39] [26] :

« Ce matin-là le temps est clair et sec, M.X sort de sa maison. Il s'aperçoit que la pelouse de son jardin est humide. Il se demande s'il a plu la nuit, ou s'il a simplement oublié de débrancher son arroseur automatique. Il jette un coup d'œil à la pelouse de son voisin, et s'aperçoit qu'elle est également humide. Il en déduit alors qu'il a plu, et il décide de partir au travail sans vérifier son arroseur automatique ».

La représentation graphique du modèle causal utilisé est dans la figure (Fig3.1). Cette figure représente un réseau bayésien simple contenant quatre variables binaires, et on peut écrire aussi :

$$P(A, B, C, D) = P(A) \cdot P(B) \cdot P(C|A, B) \cdot P(D|B) \quad (2)$$

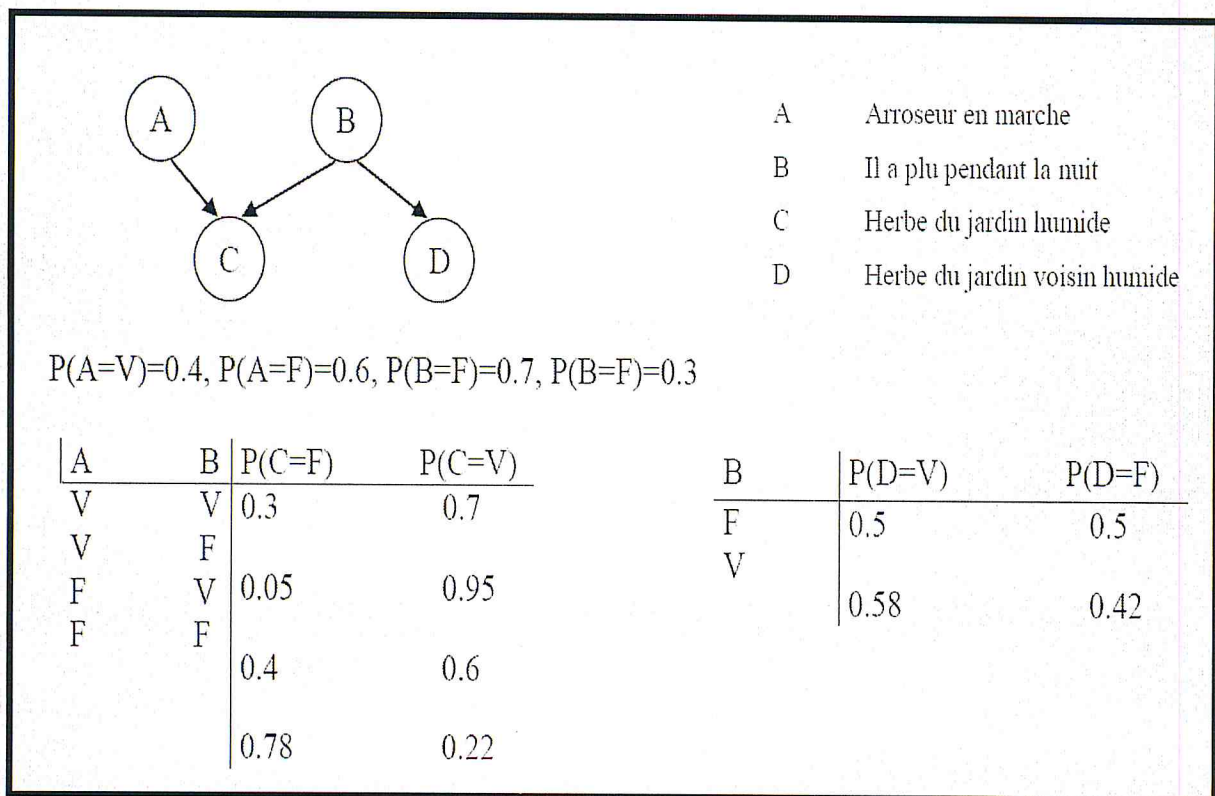


Figure3.1. Représentation graphique du modèle [39]

### 4.2. Circulation de l'information

La représentation graphique la plus intuitive de l'influence d'un événement est de relier la cause à l'effet par une flèche orientée. S'il existe une relation causale de A vers B, toute information sur A peut modifier la connaissance sur B, et réciproquement, toute information sur B peut modifier la connaissance sur A [34].

En présence d'un graphe plus complexe, il est essentiel de conserver à l'esprit que l'information ne circule pas seulement dans le sens des flèches (Fig.3.2).

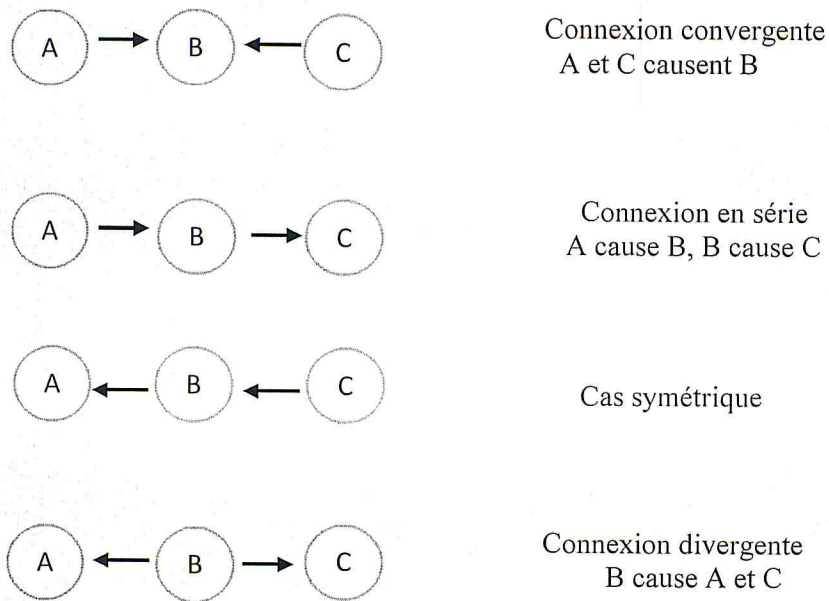


Figure 3.2. Modes de connexion [34]

Dans un graphe on peut rencontrer plusieurs modes de connexion entre les nœuds comme le montre la (Fig.3.3). Nous allons étudier comment l'information circule au sein d'un graphe causal. Dans l'exemple précédent, on voit que l'information a circulé uniquement dans le sens

Effet  $\rightarrow$  cause, par exemple la connaissance de D (herbe du jardin humide) renforce la croyance de la cause B (il a plu). De plus cet exemple nous montre que l'information peut suivre des chemins à première vue contre-intuitifs lorsqu'elle se propage dans un réseau de causalités.

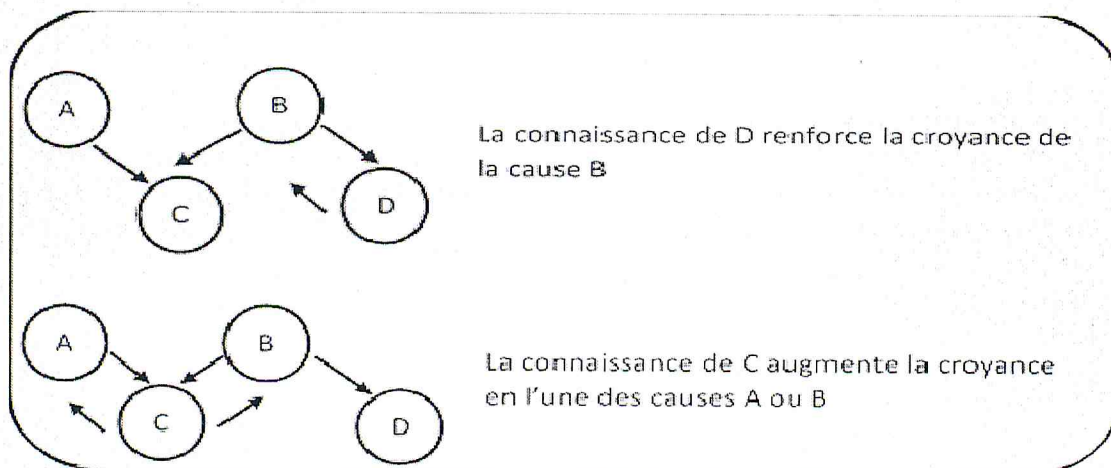


Figure 3.3. Circulation de l'information dans le graphe de la Fig.3.1 [26]

### 4.3. Définitions et propriétés

Dans ce paragraphe, nous présentons les définitions relatives aux réseaux bayésiens [26]:

#### Définition 1

Soient

- un graphe acyclique orienté  $G=(V,A)$ , où  $V$  est l'ensemble des nœuds de  $G$  et  $A$  l'ensemble des arcs de  $G$ .

- un espace probabilisé fini  $(\Omega, Z, P)$ ,  $n$  variables aléatoires  $(X_i)$ .

$(G, P)$  est un réseau bayésien si et seulement si

- il existe une bijection entre les nœuds du graphe  $G$  et les variables  $(X_i)$ .
- la propriété suivante (appelée propriété de factorisation), est vérifiée

$$p(X_1, X_2, \dots, X_n) = \prod_{1 \leq i \leq n} p(X_i | C(X_i)) \quad (3)$$

Où  $C(X_i)$  est l'ensemble des causes (parents) de  $X_i$  dans le graphe  $G$ .

#### 4.3.1. Indépendance conditionnelle

##### Définition 2

Soient deux variables aléatoires  $X$  et  $Y$ . on dit que  $X$  et  $Y$  sont indépendantes conditionnellement à  $Z$  et on note  $X \perp Y | Z$ , si l'une des propriétés équivalentes suivantes est vérifiée :

$$P(X|Z, Y) = P(X|Z) \quad (4)$$

$$P(X, Y|Z) = P(X|Z).P(Y|Z) \quad (5)$$

##### Propriété 1

Soient  $X, Y, Z, W$  des variables aléatoires, on a les propriétés suivantes :

$$\text{Symétrie } X \perp Y | Z \Leftrightarrow Y \perp X | Z \quad (6)$$

$$\text{Décomposition } X \perp Y \cup W | Z \Rightarrow X \perp Y | Z \quad (7)$$

$$\text{Union faible } X \perp (Y \cup W) | Z \Rightarrow X \perp W | (Z \cup Y) \quad (8)$$

$$\text{Contraction } X \perp Y | Z \wedge X \perp W | (Z \cup Y) \Rightarrow X \perp (Y \cup W) | Z \quad (9)$$

### 4.3.2. D-séparation

#### Définition 3

Soient  $(X, Y, Z)$  des nœuds du graphe  $G = (V, A)$ . On dit que  $X$  et  $Y$  sont d-séparés par  $Z$  si pour tout chemin entre  $X$  et  $Y$ , l'une au moins des deux conditions suivantes est vérifiée :

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

$X$  est d-séparé de  $Y$  par  $Z$ , et noté par  $(X|Z|Y)$ .

#### Définition 4

Soient  $(X, Y)$  deux nœuds du graphe  $G = (V, A)$ , et soit  $Z$  un ensemble de nœuds de ce graphe.

Soit  $S$  un chemin entre les nœuds  $X$  et  $Y$ . On dit que  $S$  est d-séparé par  $Z$  si au moins l'une des deux conditions suivantes est vérifiée :

- Le chemin converge en un nœud  $w$ , tel que  $w \sqsubseteq Z$  et  $\square Z \sqsubseteq \square Z$ ,  $w \sqsubseteq C^+(Z)$  ou  $C^+(Z)$  l'ensemble des ascendants de  $Z$ .
- Le chemin passe par un nœud  $Z \sqsubseteq Z$ , et il est soit divergent, soit en série en ce nœud.

#### Définition 5

Soient  $(X, Y)$  deux nœuds du graphe  $G = (V, A)$ , et soit  $Z$  un ensemble de nœuds de ce graphe.

On dit que  $X$  et  $Y$  sont d-séparés par  $Z$  si tous les chemins entre  $X$  et  $Y$  sont d-séparés par  $Z$ .

#### Définition 6

Soient  $(X, Y, Z)$  trois ensembles des nœuds du graphe  $G = (V, A)$ . On dit que  $X$  et  $Y$  sont d-séparés par  $Z$ , si tous les éléments de  $X$  sont d-séparés par  $Z$  de tous les éléments de  $Y$ .

#### Remarque

Deux variables  $X, Y$  sont d-séparées s'il existe une variable  $Z$  telle que tout chemin entre  $X, Y$  passe par  $Z$ .

Dans le graphe de la figure (Fig.3.4),  $A$  et  $D$  sont d-séparés par  $B$  car le chemin orienté allant de  $A$  vers  $D$  passe par  $B$ . Et de même le nœud  $D$  d-sépare les nœuds  $ABC$  des nœuds  $EFG$

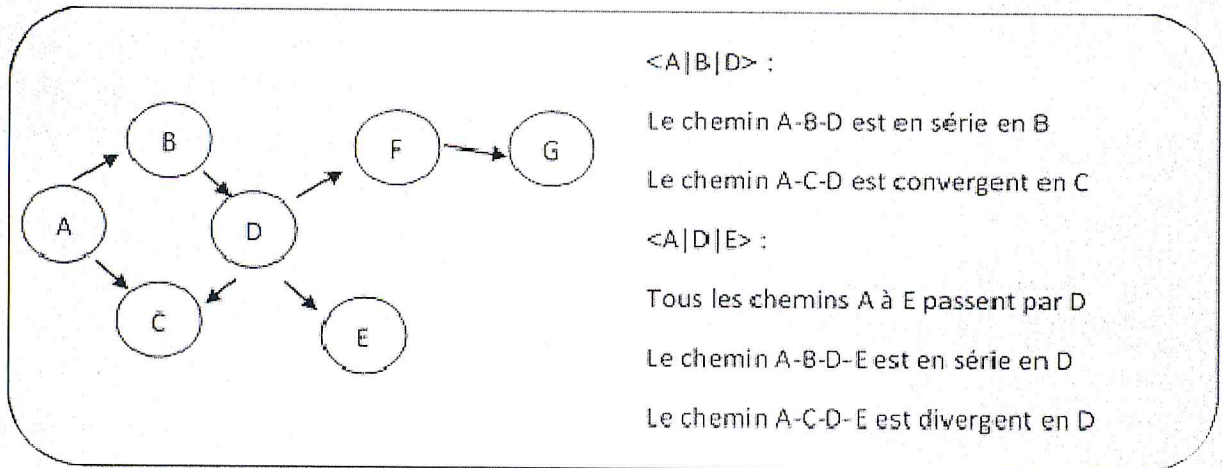


Figure.3.4. Exemple de D-séparation [26]

Avec la représentation graphique de la causalité on peut connaître la direction de circulation de connaissances dans le graphe mais on ne peut pas connaître la quantité de cette circulation de connaissances. Alors, il faut une représentation probabiliste associée avec le graphe. Avec une relation causale :  $A \Rightarrow B$  on peut représenter la quantité de cette relation par la probabilité conditionnelle :  $p(B|A)$ .

**Définition 7**

A et B sont indépendantes conditionnellement à C ssi :

- Lorsque l'état de C est connu toute connaissance sur B n'altère pas A
- $P(A|B, C) = P(A|C)$

-Le réseau bayésien permet de représenter graphiquement les indépendances conditionnelles. Trois types de relations possibles entre A, B et C à savoir :

Connexion en série :

- A et B sont dépendants
- A et B sont indépendantes conditionnellement à C ssi :
  - \*\*  $P(C)$  est connue { $P(A)$  n'intervient pas le calcul de  $P(B)$ }
  - \*\*  $P(S8|S5, S2) = P(S8|S4) = P(S8|Parents(S8))$ .

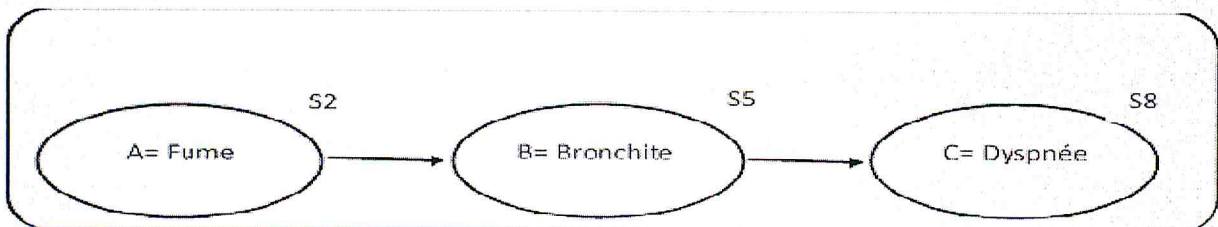


Figure.3.5. Connexion en série [40]



### Connexion divergente

- A et B sont dépendants.
- A et B sont indépendantes conditionnellement à C ssi :
  - \*\*  $P(C)$  est connue  $\{P(A)$  n'intervient pas le calcul de  $P(B)\}$
  - \*\*  $P(S5|S2, S4) = P(S5|S2) = P(S4|Parents(S4))$ .

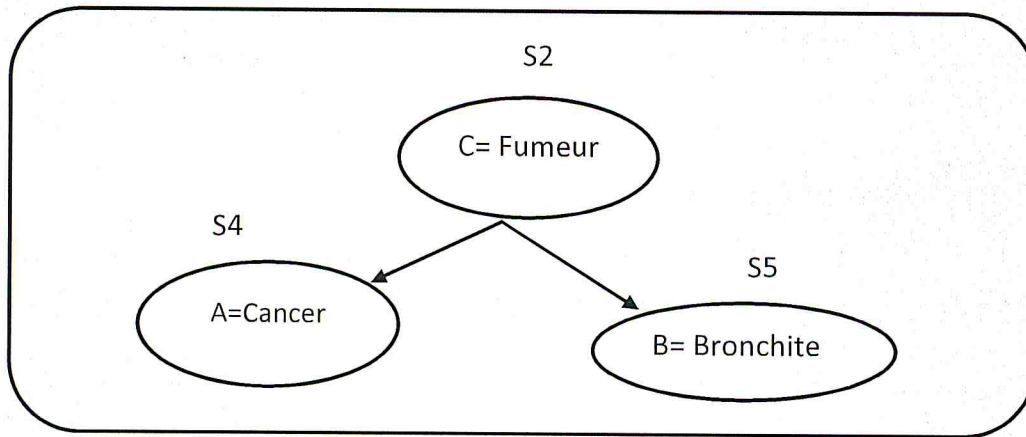


Figure.3.6. Connexion divergente [40]

### Connexion convergente

- A et B sont indépendants.
- A et B sont indépendantes conditionnellement à C ssi :
  - \*\*  $P(C)$  est connue  $\{P(A)$  n'intervient pas le calcul de  $P(B)\}$
  - \*\*  $P(S8|S6, S5) = P(S8|Parents(S8))$ .

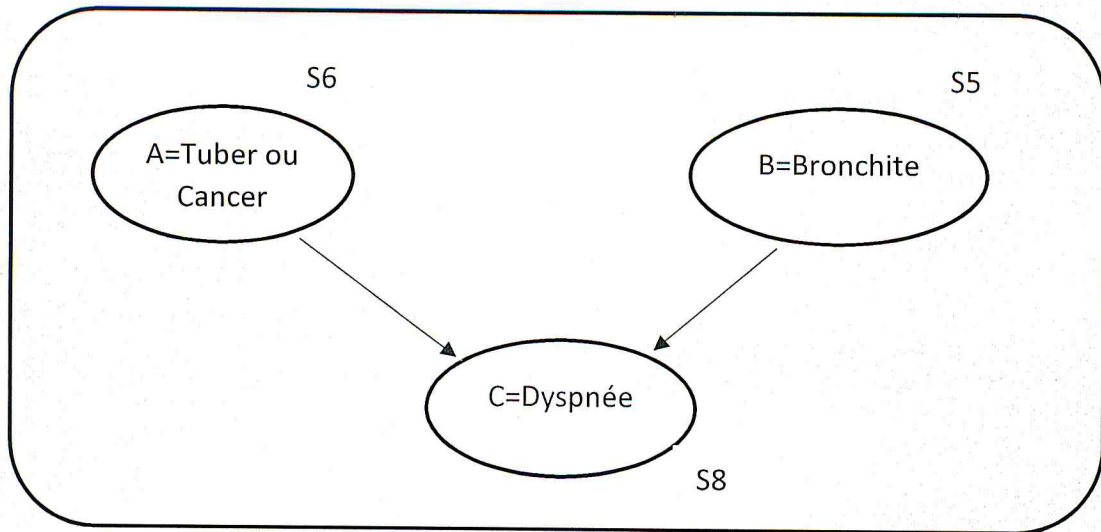


Figure3.7. Connexion convergence [40]

### 5. Formule de Bayes

Le théorème de Bayes est utilisé dans l'inférence statistique pour mettre à jour ou actualiser les estimations d'une probabilité ou d'un paramètre quelconque, à partir des observations et des lois de probabilité de ces observations. Il y a une version discrète et une version continue du théorème [26].

- L'école bayésienne utilise les probabilités comme moyen de traduire numériquement un degré de connaissance (la théorie mathématique des probabilités n'oblige en effet nullement à associer celles-ci à des fréquences, qui n'en représentent qu'une application particulière résultant de la loi des grands nombres). Dans cette optique, le théorème de Bayes peut s'appliquer à toute proposition, quelle que soit la nature des variables et indépendamment de toute considération ontologique.
- L'école fréquentiste utilise les propriétés de long terme de la loi des observations et ne considère pas de loi sur les paramètres, inconnus mais fixés

En théorie des probabilités, le théorème de Bayes énonce des probabilités conditionnelles : étant donné deux événements A et B, le théorème de Bayes permet de déterminer la probabilité de A sachant B, si l'on connaît les probabilités :

- de A ;
- de B ;
- de B sachant A.

Ce théorème élémentaire (originellement nommé « de probabilité des causes ») a des applications considérables. Pour aboutir au théorème de Bayes, on part d'une des définitions de la probabilité conditionnelle [26] :

$$p(A|B)p(B) = p(A \cap B) = p(B|A)p(A) \quad (10)$$

En notant  $p(A \cap B)$  la probabilité que A et B aient tous les deux lieu. En divisant de part et d'autre par  $P(B)$ , on obtient :

$$p(A|B) = \frac{p(B|A)p(A)}{P(B)} \quad (11)$$

Soit le théorème de Bayes. Chaque terme du théorème de Bayes a une dénomination usuelle. Le terme  $P(A)$  est la probabilité a priori de A. Elle est « antérieure » au sens qu'elle précède toute information sur B.  $P(A)$  est aussi appelée la probabilité marginale de A. Le terme  $P(A|B)$  est appelée la probabilité a posteriori de A sachant B (ou encore de A sous condition B). Elle est « postérieure », au sens qu'elle dépend directement de B. Le terme  $P(B|A)$ , pour un B connu, est appelée la fonction de vraisemblance de A. De même, le terme  $P(B)$  est appelé la probabilité marginale ou a priori de B.

### 5.1. Autres écritures du théorème de Bayes

On améliore parfois le théorème de Bayes en remarquant que [26] :

$$p(B) = p(A \cap B) + p(A^c \cap B) = p(B|A)p(A) + p(B|A^c)p(A^c) \quad (12)$$

Afin de réécrire le théorème ainsi :

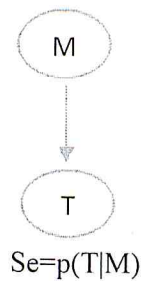
$$P(A|B) = \frac{p(B|A)p(A)}{p(B|A)p(A) + p(B|A^c)p(A^c)} \quad (13)$$

Où  $AC$  est le complémentaire de A. Plus généralement, si  $\{A_i\}$  est une partition de l'ensemble des possibles,

$$p(A_i|B) = \frac{p(B|A_i)p(A_i)}{\sum_j p(B|A_j)p(A_j)} \quad (14)$$

### 5.2. Exemple d'application de la formule de Bayes

L'application de la formule de Bayes est le cas le plus simple d'inférence bayésienne. Prenons l'exemple d'un test médical. Notons T la variable donnant le résultat du test (positif ou négatif) et M la variable représentant l'état de la personne (malade ou non malade). Le test est caractérisé par sa sensibilité (Se) et sa spécificité (Sp). La sensibilité nous donne la probabilité que le test soit positif sachant que la personne est malade [42]:



**Figure 3.8.** Représentation graphique de la relation entre la variable M (malade ou non malade) et la variable T (test positif ou négatif) dans le cas d'un test médical [54]

La spécificité à l'inverse nous donne la probabilité que le test soit négatif lorsque la personne n'est pas malade :

$$Sp = p(T|M) \quad (15)$$

La relation entre M et T est représentée graphiquement par la figure (Fig. 3.8) :

La règle de Bayes nous permet de calculer la probabilité que la personne soit malade sachant que le test est positif.

$$\begin{aligned} p(M|T) &= p(T|M) \frac{p(M)}{p(T)} \\ &= p(T|M) \frac{p(M)}{p(T|M)p(M) + p(T|M)(1-p(M))} \quad (16) \\ &= \frac{Se p(M)}{Se p(M) + (1-Sp)(1-p(M))} \end{aligned}$$

La probabilité  $p(M)$  est la probabilité a priori que la personne soit malade. Elle est par exemple donnée par la prévalence de la maladie dans la population à laquelle appartient la personne.

Dans [54] les auteurs montrent l'importance et l'intérêt de la probabilité a priori dans le raisonnement bayésien au travers de l'exemple du diagnostic d'une sérologie positive pour le VIH, d'une part, chez une femme de 75 ans sans antécédents et, d'autre part, chez un toxicomane de 27 ans. Le test a les caractéristiques suivantes :  $Se = 0.97$  et  $Sp = 0.98$ .

– Dans le cas d'une la femme âgée la prévalence de la maladie est  $P(M) = 1/500000$ , le calcul donne dans ce cas,  $P(M|T) = 0.00097$ .

– Dans le cas d'un toxicomane avec une prévalence  $P(M) = 1/10$ , nous obtenons  $P(M|T) = 0.982$ .

L'inférence bayésienne nous permet ici de fusionner l'information donnée par les caractéristiques du test avec celle donnée par la prévalence de la maladie. L'observation « test positif » est prise en compte pour calculer a posteriori, c'est à dire après observation du résultat du test, la probabilité que la

personne soit malade. Cet exemple nous apprend que dans le cas de la femme âgée la spécificité du test n'est pas suffisante pour conclure à la présence de la maladie lorsque le test est positif.

### 6. Construction des réseaux bayésiens

La construction d'un réseau bayésien s'effectue en trois étapes essentielles, qui sont présentées sur la figure (Fig.3.9) ci-après. Chacune des trois étapes peut impliquer un recueil d'expertise, au moyen de questionnaires écrits, d'entretiens individuels ou encore de séances de brainstorming. Préconiser, dans un cadre général, l'une ou l'autre de ces approches serait pour le moins hasardeux [37].

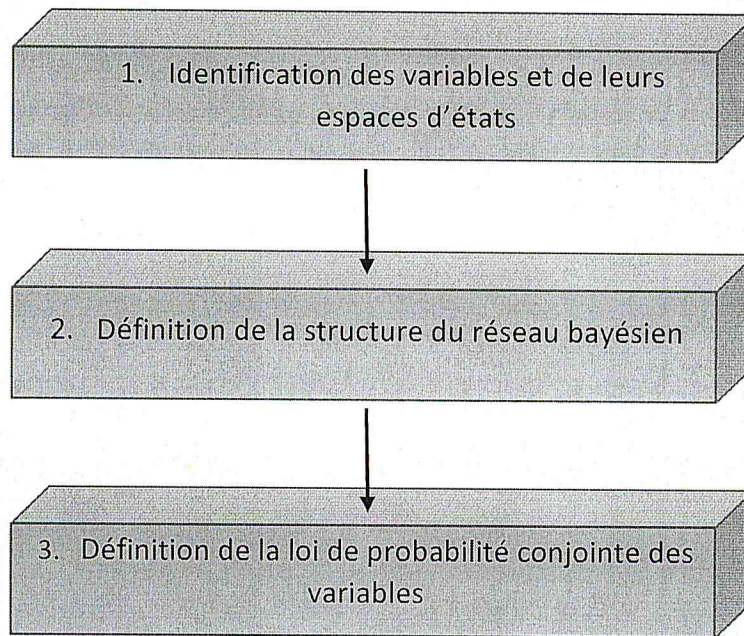


Figure 3.9. Etapes de construction d'un réseau bayésien [37]

#### 6.1. Identification des variables et de leurs espaces d'états

La première étape de construction du réseau bayésien est la seule pour laquelle l'intervention humaine est absolument indispensable. Il s'agit de déterminer l'ensemble des variables  $X_i$ , catégorielles ou numériques, qui caractérisent le système. Comme dans tout travail de modélisation, un compromis entre la précision de la représentation et la maniabilité du modèle doit être trouvé, au moyen d'une discussion entre les experts et le modélisateur.

Lorsque les variables sont identifiées, il est ensuite nécessaire de préciser l'espace d'états de chaque variable  $X_i$ , c'est-à-dire l'ensemble de ses valeurs possibles.

La majorité des logiciels de réseaux bayésiens ne traite que des modèles à variables discrètes, ayant un nombre fini de valeurs possibles. Si tel est le cas, il est impératif de discrétiser les plages de variation

des variables continues. Cette limitation est parfois gênante en pratique, car des discrétisations trop fines peuvent conduire à des tables de probabilités de grande taille, de nature à saturer la mémoire de l'ordinateur [37].

### 6.2. Définition de la structure du réseau bayésien

La deuxième étape consiste à identifier les liens entre variables, c'est-à-dire à répondre à la question : pour quels couples  $(i, j)$  la variable  $X_i$  influence-t-elle la variable  $X_j$  ? Dans la plupart des applications, cette étape s'effectue par l'interrogation d'experts. Dans ce cas, des itérations sont souvent nécessaires pour aboutir à une description consensuelle des interactions entre les variables  $X_i$ . L'expérience montre cependant que la représentation graphique du réseau bayésien dans cette étape est un support de dialogue extrêmement précieux.

Un réseau bayésien ne doit pas comporter de circuit orienté ou boucle. Cependant, le nombre et la complexité des dépendances identifiées par les experts laissent parfois supposer que la modélisation par un graphe sans circuit est impossible. Il est alors important de garder à l'esprit que, quelles que soient les dépendances stochastiques entre des variables aléatoires discrètes, il existe toujours une représentation par réseau bayésien de leur loi conjointe. Ce résultat théorique est fondamental et montre bien la puissance de modélisation des réseaux bayésiens [37].

Lorsque l'on dispose d'une quantité suffisante de données de retour d'expérience concernant les variables  $X_i$ , la structure du réseau bayésien peut également être apprise automatiquement par le réseau bayésien, à condition bien sûr que le logiciel utilisé soit doté de la fonctionnalité adéquate [37].

### 6.3. Loi de probabilité conjointe des variables

La dernière étape de construction du réseau bayésien consiste à renseigner les tables de probabilités associées aux différentes variables. Dans un premier temps, la connaissance des experts concernant les lois de probabilité des variables est intégrée au modèle. Concrètement, deux cas se présentent selon la position d'une variable  $X_i$  dans le réseau bayésien :

- La variable  $X_i$  n'a pas de variable parente: les experts doivent préciser la loi de probabilité marginale de  $X_i$ .
- La variable  $X_i$  possède des variables parentes: les experts doivent exprimer la dépendance de  $X_i$  en fonction des variables parentes, soit au moyen de probabilités conditionnelles, soit par une équation déterministe (que le logiciel convertira ensuite en probabilités).

Le recueil de lois de probabilités auprès d'experts est une étape délicate du processus de construction du réseau bayésien. Typiquement, les experts se montrent réticents à chiffrer la plausibilité d'un événement qu'ils n'ont jamais observé.

Cependant, une discussion approfondie avec les experts, aboutissant parfois à une reformulation plus précise des variables, permet dans de nombreux cas l'obtention d'appréciations qualitatives. Ainsi, lorsqu'un événement est clairement défini, les experts sont généralement mieux à même d'exprimer si celui-ci est probable, peu probable, hautement improbable, etc...

Le cas d'absence totale d'information concernant la loi de probabilité d'une variable  $X_i$  peut être rencontré. La solution pragmatique consiste alors à affecter à  $X_i$  une loi de probabilité arbitraire, par

exemple une loi uniforme. Lorsque la construction du réseau bayésien est achevée, l'étude de la sensibilité du modèle à cette loi permet de décider ou non de consacrer davantage de moyens à l'étude de la variable  $X_i$ .

La quasi-totalité des logiciels commerciaux de réseaux bayésiens permet l'apprentissage automatique des tables de probabilités à partir de données. Par conséquent, dans un second temps, les éventuelles observations des  $X_i$  peuvent être incorporées au modèle, afin d'affiner les probabilités introduites par les experts [37].

## 7. Propagation de messages (Algorithme Pearl)

Cette première méthode de propagation de messages [39] n'est valable que pour les réseaux bayésiens ayant comme graphe une arborescence. Cette particularité permet de déduire qu'il existe une unique chaîne entre deux sommets de ce graphe.

Soit  $O$  l'ensemble des variables observées. A partir d'un sommet  $X$ , nous pouvons définir deux ensembles de sommets  $P_X$  et  $E_X$  tels que  $P_X$  regroupe les sommets de  $O$  dont la chaîne vers  $X$  passe par un parent de  $X$  et  $E_X$  regroupe les sommets de  $O$  dont la chaîne vers  $X$  passe par un enfant de  $X$  et  $O = P_X \cup E_X$ . Il est possible de montrer que :

$$p(X|O = o) \propto \lambda(X) \cdot \pi(X) \quad (17)$$

Avec  $\lambda(X) \propto p(E_X|X)$  et  $\pi(X) \propto p(X|P_X)$

Il s'agit ensuite de calculer des messages  $\lambda$  et  $\pi$  pour chaque variable selon un ordre topologique :

### - Les messages $\lambda$

Pour chaque enfant  $Y$  de  $X$  :

$$\lambda_y(X = x) = \sum_{y \in D_y} p(Y = y|X = x) \cdot \lambda(Y = y) \quad (18)$$

$$\lambda(X = x) = \begin{cases} \delta_{x_0}^x & \text{si } X \text{ est instancié à } x_0 \\ 1 & \text{si } X \text{ est une feuille} \\ \prod_{Y \in \text{enfants}(X)} \lambda_Y(X = x) & \text{sinon} \end{cases}$$

Où le symbole de Kronecker  $\delta_i^j$  a la signification suivante :  $\begin{cases} 1 \text{ si } i = j \\ 0 \text{ si } i \neq j \end{cases}$

### - Les messages $\pi$

Pour l'unique parent  $Z$  de  $X$  :

$$\pi_x(Z = z) = \pi(Z = z) \prod_{U \in \text{enfants}(Z) \setminus \{X\}} \lambda_U(Z = z) \quad (19)$$

$$\pi(X = x) = \begin{cases} \delta_{x0} & \text{si } X \text{ est instancité à } x_0 (X \in O) \\ p(X) & \text{si } X \text{ est une racine} \\ \sum_{z \in Dz} p(X = x | Z = z) \cdot \pi_X(Z = z) & \text{sinon} \end{cases}$$

Dans ce cas sa complexité en temps est  $O(nd^2)$  et en espace  $O(nd)$  avec  $n=|X|, d = \max_{x_i \in X} D_{x_i}$  cette propagation peut également s'effectuer dans les polyarbres<sup>4</sup>.

## 8. Domaines d'application des réseaux Bayésiens

### ➤ Santé

Les premières applications des réseaux bayésiens ont été développées dans le domaine du diagnostic médical. Les réseaux bayésiens sont particulièrement adaptés à ce domaine parce qu'ils offrent la possibilité d'intégrer des sources de connaissances hétérogènes (expertise humaine et données statistiques), et surtout parce que leur capacité à traiter des requêtes complexes (explication la plus probable, action la plus appropriée) peuvent constituer une aide véritable et interactive pour le praticien. Le système Pathfinder, développé au début des années 1990 a été conçu pour fournir une assistance au diagnostic histopathologique, c'est-à-dire basé sur l'analyse des biopsies. Il est aujourd'hui intégré au produit Intellipath, qui couvre un domaine d'une trentaine de types de pathologies [37].

### ➤ Industrie

Dans le domaine industriel, les réseaux bayésiens présentent également certains avantages par rapport aux autres techniques d'intelligence artificielle par exemple la société danoise Hugin, considérée comme l'un des pionniers dans le développement des réseaux bayésiens. Hugin a développé pour le compte de Lockheed Martin le système de contrôle d'un véhicule sous-marin autonome. Ce système évalue en permanence les capacités du véhicule à réagir à certains types d'événements [37].

### ➤ Défense

La fusion de données est particulièrement un domaine d'application privilégié des réseaux bayésiens, grâce à leur capacité à prendre en compte des données incomplètes ou incertaines et guider la recherche ou la vérification de ces informations. La société Mitre a développé un système de défense tactique embarqué pour les navires de guerre de la marine américaine décide des ripostes à adopter. Ce système analyse les informations permet en particulier de gérer les menaces multiples, qui peuvent générer des conflits sur l'affectation des armes [37].

### ➤ Banque/finance

Les applications dans le domaine de la banque et de la finance sont encore rares, ou du moins ne sont pas publiées. Mais cette technologie présente un potentiel très important pour un certain nombre d'applications relevant de ce domaine, comme l'analyse financière, le scoring, l'évaluation du risque ou la détection de fraudes.



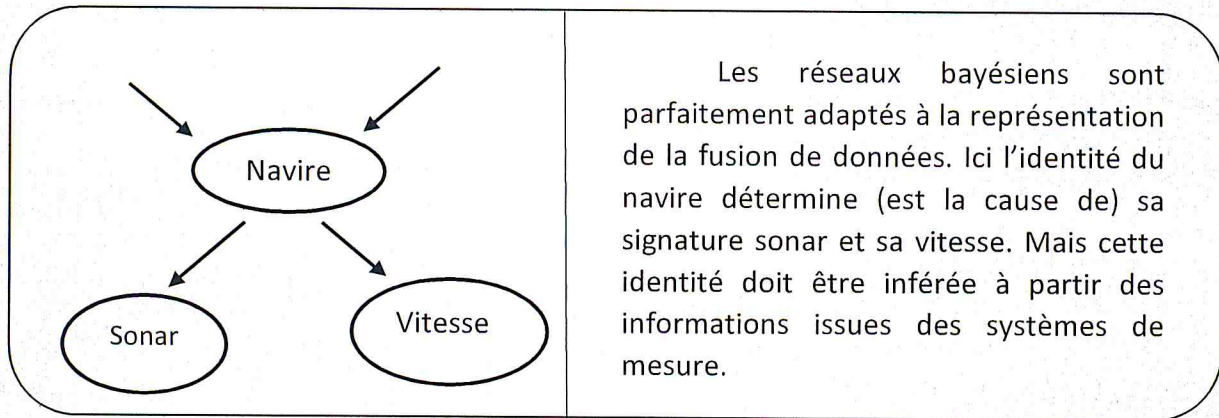


Figure.3.10. Principe de la fusion de données par réseau bayésien [37]

Dans les années 1980 avec des systèmes experts, des applications comme l'analyse financière, le scoring ou la détection de fraudes ont été progressivement considérées comme relevant du domaine de la modélisation quantitative, et donc abordées par des techniques comme les réseaux neuronaux ou les arbres de décision, techniques quantitatives qui se révèlent incapables de prendre en compte par elles-mêmes la révision des modèles.

L'exemple de l'autorisation des transactions sur cartes bancaires est assez significatif. L'un des premiers systèmes experts développés dans ce domaine fut l'Authorize Assistant d'American Express, au début des années 80 [37].

### 9. Réseau bayésien et l'incertitude médicale

L'action médicale repose sur la capacité de raisonnement du médecin et son aptitude à prendre des décisions alors que les informations utilisées sont potentiellement entachées d'incertitude. Cette incertitude est d'origine multiple : possibilité d'erreur dans les données, ambiguïté de la représentation de l'information, incertitude sur les relations entre les diverses informations. Une première approche à la représentation de la connaissance dans le contexte d'incertitude a utilisé la théorie des probabilités. Ainsi, plusieurs études ont montré que les systèmes d'aide à la décision élaborés sous le modèle probabiliste pouvaient faire aussi bien, voire même mieux que le médecin [40].

Cependant, bien qu'utile pour codifier l'incertitude dans un problème de décision, l'approche purement probabiliste comporte des limites. Tout d'abord, l'élaboration, la représentation et la manipulation de telles bases de connaissances sont souvent nécessaire de postuler des hypothèses insuffisantes pour obtenir un réel impact sur la pratique médicale.

Enfin les alternatives à une décision ou les préférences de l'utilisation ne peuvent être prises en compte. Les systèmes basés sur des règles, apparus dans les années 1970, circonvenaient la difficulté

du calcul probabiliste en utilisant d'autres paramètres pour représenter l'incertitude : facteur de certitude, calcul de Dempster-Shafer ou logique floue. Plus récemment, le développement de la technologie informatique a permis de reconsidérer le formalisme probabiliste. Ceci est particulièrement le cas du domaine de l'analyse de décision, méthodologie basée sur la théorie des probabilités mais qui permet de représenter explicitement les problèmes de décision et les préférences de l'utilisateur. Les réseaux bayésien ou diagramme d'influence (Belief network) constituent un des modèles de représentation des connaissances utilisables en analyse de décision [40].

Un réseau bayésien, également appelé diagramme d'influence ou réseau causal est un graphe dirigé acyclique dans lequel les nœuds représentent les variables et les arcs précisent les dépendances probabilistes entre les variables. Il permet d'afficher graphiquement les variables d'un problème de décision et les relations ou influences entre variables, qui peuvent mener à des décisions complexes. Les réseaux bayésien centrent l'attention de la décision exclusivement sur les composants du problème en relation avec la tâche de décision présente. Exclure l'information non pertinente du diagramme d'influence facilite le travail de la décision et lui permet de gagner du temps puisqu'il existe moins de variables à interpréter [28].

### 10. Avantage des réseaux bayésiens

Le point clé dans l'utilisation de réseaux bayésiens (modèles graphiques) est dans le fait que l'analyse et la manipulation de modèles à plusieurs variables entraînant des relations de dépendance peuvent être considérablement facilitées en explorant les relations entre le modèle de probabilité et la représentation graphique. Les principaux avantages s'inscrivent dans la description du modèle et dans l'efficacité du calcul [44] [45]:

#### ➤ Description du modèle

Les graphes sont une façon naturelle de représenter l'information sous une forme compacte que l'humain peut saisir, comprendre et utiliser. En particulier, la structure d'un modèle graphique montre de manière simple et claire les dépendances conditionnelles dans le modèle de probabilité correspondant, permettant l'évaluation et la révision du modèle. De plus, le fait que le modèle graphique nécessite l'encodage explicite et la confrontation des hypothèses, peut être extrêmement utile dans la construction du modèle.

#### ➤ Efficacité du calcul

Les modèles graphiques sont une base solide pour la spécification d'algorithmes d'inférence efficaces pour effectuer des mesures requises dans le modèle de probabilité. L'avantage du calcul dans les modèles graphiques réside dans le fait que ces algorithmes d'inférence peuvent être spécifiés automatiquement dès lors que la structure initiale du graphe est déterminée. Il est à noter que le cadre des modèles graphiques ne permet pas automatiquement d'éviter l'explosion combinatoire des paramètres qui peut résulter de la construction de modèles plus réalistes. Il permet plutôt d'identifier une procédure d'inférence efficace de manière automatique si la structure du modèle le permet.

Du point de vue des applications, l'avantage des réseaux bayésiens est de pouvoir être utilisé aussi bien à partir d'une connaissance explicite que d'une connaissance implicite, ou de toute combinaison des deux, parmi leurs buts [42]:

- Intégrer la notion d'incertitude dans les systèmes experts, vu que la construction de celui-ci nécessite presque toujours la prise en compte de l'incertitude dans le raisonnement.

- Leur capacité d'effectuer des inférences dans un contexte d'incertitude, en quelque sorte comme alternative aux systèmes experts.
- La combinaison des connaissances à priori et les connaissances empiriques (expérimentales), ce qui permettra l'enrichissement par l'expérience de la base de connaissances à priori par la connaissance empirique contenue dans les données.
- Quantifier la notion de « plus probable ».
- Nouveaux algorithmes d'apprentissage.
- Analyser d'autres algorithmes ne manipulant pas explicitement des probabilités.
- Ils permettent de rendre quantitatifs les raisonnements sur les causalités que l'on peut faire à l'intérieur du graphe.

### 11. Conclusion

Les réseaux bayésiens font partie de la famille des modèles graphiques. Ils regroupent au sein d'un même formalisme la théorie des graphes et celle des probabilités afin de fournir des outils efficaces autant qu'intuitifs pour représenter une distribution de probabilités jointe sur un ensemble de variables aléatoires. Ce formalisme très puissant permet une représentation intuitive de la connaissance sur un domaine d'application donné et facilite la mise en place de modèles performants et clairs.

La représentation de la connaissance se base sur la description, par des graphes, des relations de causalité existant entre des variables décrivant le domaine d'étude. A chaque variable est associée une distribution de probabilités locale quantifiant la relation causale.

Dans ce chapitre nous avons couvert les principaux aspects concernant les réseaux bayésiens. Nous avons vu que ces modèles graphiques sont des outils bien adaptés à la description des problèmes de décisions dans l'incertain, en permettant notamment de tenir compte des connaissances du problème à traité.

# Chapter 4

### 1. Introduction

Le microanévrisme est le premier signe ophtalmoscopique de la rétinopathie diabétique. Le nombre de microanévrismes augmente progressivement et cette augmentation est un bon indice de sa progression dans les stades initiaux. Plusieurs études ont montré que le nombre initial de microanévrismes et leur taux d'augmentation étaient de bons facteurs prédictifs de l'aggravation de la rétinopathie diabétique [19].

Dans ce chapitre nous proposons une méthode pour le développement d'un système d'aide au diagnostic de microanévrisme nous utilisons deux méthodes le RàPC et les Réseaux Bayésiens pour la classification en exploitant des connaissances déjà acquises dans des expériences passées. Le choix du cas similaire se base sur l'estimation d'un degré de flou.

Nous avons implémenté les deux méthodes (RàPC et RB) en utilisant l'environnement MATLAB

### 2. Description de système

Le système proposé comprend deux principales étapes : l'étape recueil d'information qui permet de trouver les microanévrismes et calcul la probabilité de la déduction, la deuxième étape et celle de recherche le cas le plus probable.

Le système trouve des microanévrismes à partir d'une image normale (en couleur). Dans ce système les microanévrismes détecté sont sélectionné à partir des plusieurs détection sur l'image et seront considéré comme des vrais microanévrismes, le système produit beaucoup de résultat pour obtenir la probabilité de détection de microanévrismes.

On utilisant le résonnement à partir de cas on pourra retrouver les cas de bas et choisir le cas le plus probable on utilisant les réseaux bayésiens naïf

Le terme « cas » représente une détection de microanévrismes, les descripteurs de problème sont la probabilité et le nombre de la détection de vrai microanévrisme, la solution représenté sur le cas le plus probable, la variance de la probabilité  $\pm 0,02$ .

Le schéma suivant représente l'architecture du système

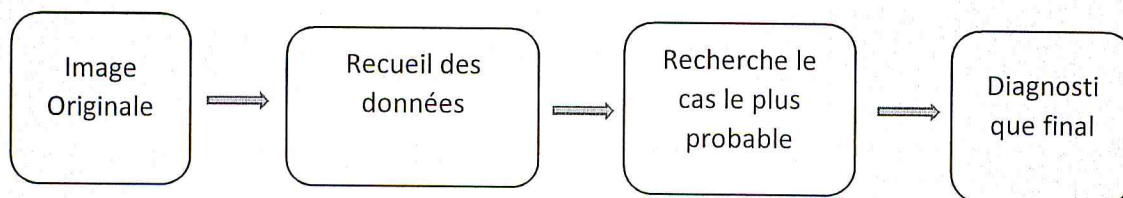


Figure (4.1) :L'architecture de système

#### 2.1.phase recueil de donnée

Dans cette phase nous allons décrire la méthode adoptée pour la détection des microanévrismes. Le résultat final de cette étape est un ensemble des microanévrismes détecté pour un nombre de caractéristiques calculé.

Cette phase comprend deux étapes:

- une première étape correspondant à un prétraitement afin de diminuer le bruit et améliorer l'image.
- une deuxième étape correspondant à la segmentation d'image pour éliminer les structures qui ne correspondent pas à la structure d'un microanévrisme, afin d'extraire les probabilités.

Les étapes de la méthode sont illustrées dans la figure :

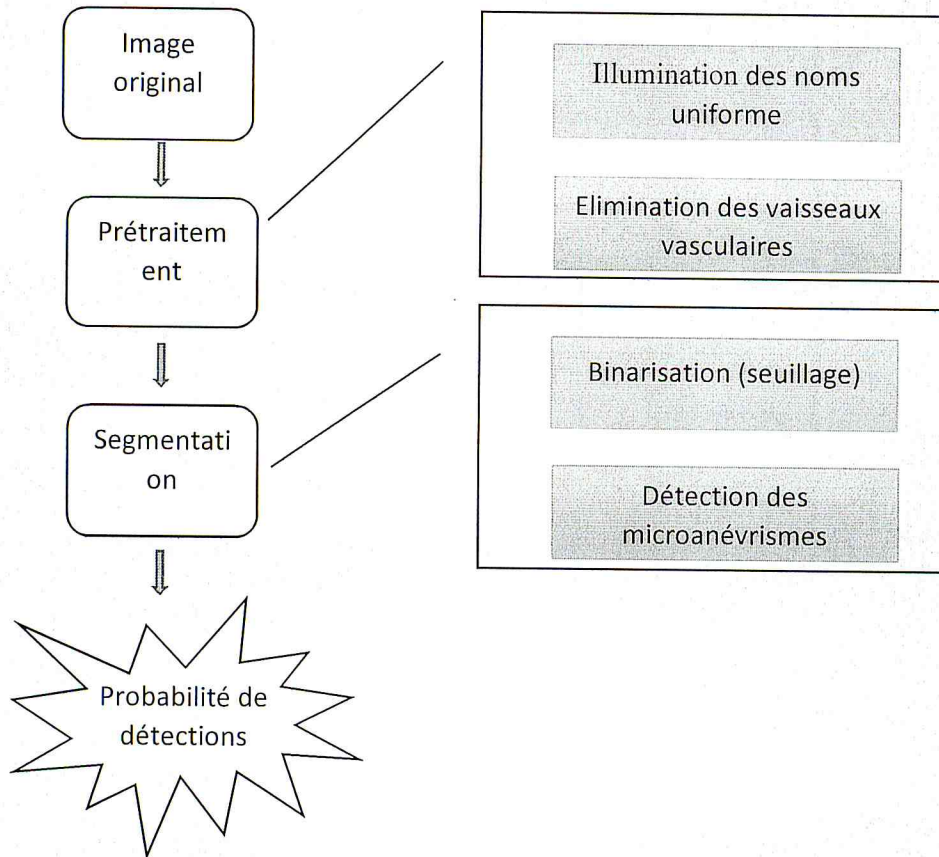


Figure (4.2) : La procédure de recueil de données

### 2.1.1. Prétraitement

Le prétraitement a pour but l'amélioration des images rétinienne acquises, par la correction de l'illumination non uniforme dans l'image et par élimination du réseau vasculaire.

#### 2.1.1.1. Correction de l'illumination non uniforme

La correction d'illumination non uniforme a beaucoup d'importance pour l'analyse automatique des images rétinienne, car les structures de la rétine ne peuvent pas être caractérisées par leur couleur ou par leur niveau de gris. La non-uniformité de l'éclairage de la rétine est due aux méthodes d'illumination et à la géométrie de l'œil.

Pour corriger l'illumination non uniforme un filtre moyen est appliqué à l'image au niveau de gris. L'image corrigée est obtenue en soustrayant l'image moyenne de l'image originale. On utilise les fonctions définies sur MATLAB pour multiplier la contracte

### 2.1.1.2. L'élimination du réseau vasculaire

Dans une image rétinienne le réseau vasculaire et les microanévrismes sont des éléments clairs, pour distinguer entre le réseau vasculaire et les microanévrismes, une transformation chapeau haut de forme associée au supremum d'ouvertures est utilisée. Les microanévrismes correspondent à des pics dans une image à illumination non uniforme. Le meilleur critère pour les détecter est leur taille « leur diamètre ». Le réseau vasculaire est une structure fine et linéaire de forme allongée qui contient un élément structurant linéaire dans une direction. Le supremum d'ouvertures avec des éléments structurants linéaires efface les pics et préserve des structures fines suffisamment linéaires pour contenir l'élément structurant. Le supremum d'ouvertures est décrit par la formule suivante :

$$I_s = \sup_{i=1, dn} (\gamma Bi(I_c)) \quad (4.2)$$

Où :  $I_s$  : est le supremum d'ouvertures

$I_c$  : est l'image corrigée

$B$  : est l'élément structurant linéaire

$dn$  : nombre de directions

Après l'élimination des pics par le supremum d'ouvertures qui sont les éléments recherchés, un chapeau haut de forme « top hat » est appliqué, dont sa formule est la suivante :

$$I_{\text{tophat}} = I_c - I_s \quad (4.3)$$

Le chapeau haut de forme permet de préserver que des éléments clairs de taille inférieure à la taille de l'élément structurant, pour cela la taille de l'élément structurant choisie doit être assez importante pour favoriser les microanévrismes.

### 2.1.2. La segmentation d'image

#### 2.1.2.1. La binarisation (Seuillage)

Les régions candidates sont représentées par des objets sur une image binaire. L'image binaire est obtenue par l'application d'un seuil sur l'image à teintes de gris. Le seuil correspond au contraste minimal d'une région pour être considérée comme un microanévrisme.

Un mauvais choix du seuil peut entraîner un nombre de détection extrêmement élevé. La détection est une région considérée comme un microanévrisme potentiel et que son contraste dépasse la valeur seuil, cette valeur est la même pour toutes les images. Le choix du seuil est très important pour distinguer entre les points détectés susceptibles d'être des microanévrismes et ceux qui ne le sont pas. Le seuil est un compromis entre la sensibilité et la spécificité. La sensibilité est le pourcentage des microanévrismes qui ont été correctement détectés, alors que la spécificité est pourcentage des objets trouvés par l'algorithme sans qu'ils ne correspondent à des vrais microanévrismes.

#### 2.1.2.2. Détection des microanévrismes

Après l'application de seuillage sur l'image corrigée les régions détectées correspondent à des microanévrismes et les fins du réseau vasculaire qui ne sont pas éliminées.

Afin d'éliminer les non-microanévrismes on applique une fonction de détection des formes et on sélectionne les formes structurées comme des microanévrismes en choisissant leur diamètre et leur forme (disque) après on les colore en vert.

On calcule la probabilité de détection de microanévrisme par la fonction suivante

$$P_m = D_{vm}/D_t$$

$P_m$  : probabilité de détection de microanévrismes

$D_t$  : nombre de détection total sur l'image

$D_{vm}$  : nombre de détection de vrais microanévrismes

### 2.2. recherche de cas le plus probable

Dans cette phase nous allons décrire la méthode adoptée pour la recherche de cas. Le résultat final de cette étape est les cas les plus probables.

Cette phase comprend deux étapes:

- La première étape est la construction d'un réseau bayésien naïf à partir d'une base des cas.
- La deuxième étape est la recherche des cas les plus probables dans le réseau

Après il vient l'affichage final de l'image avec les détections des microanévrismes et la structure des cas les plus probables, la figure suivante montre le processus de recherche :

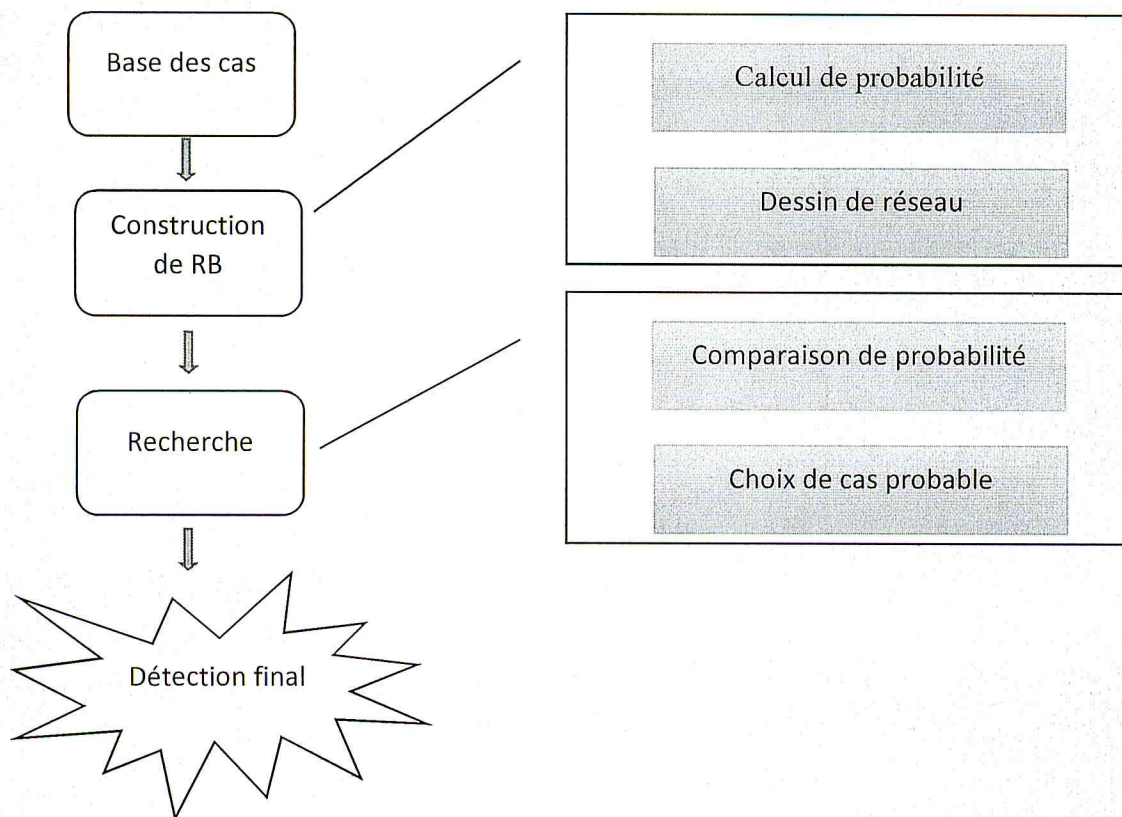


Figure (4.3) : Processus de recherche



### 3. Implémentation

#### 3.1. Environnement de Programmation

##### 3.1.1. MATLAB

MATLAB est un langage de programmation basé essentiellement sur le calcul matriciel, avec des fonctionnalités mathématiques et graphique étendues. Le choix de langage est justifié par le nombre de boites à outils (toolbox) dédié au traitement d'image qui sont très utiles. En utilisant MATLAB, la résolution des problèmes de calcul scientifique est plus rapide qu'avec les langages des programmations traditionnelles, tel que C++ et JAVA

MATLAB dispose de boite à outil supplémentaire qui élargissant l'environnement MATLAB pour résoudre des catégories particulières de problème dans plusieurs domaines d'application, incluant le traitement de signal, la conception de système de contrôle, les tests est les mesures, la modélisation est l'analyse financière, ainsi que la biologie informatique

MATLAB offre un certain nombre de fonctionnalité mathématique pour l'algèbre linéaire, les statistiques, l'analyse des fourriers, le filtrage, l'optimisation et l'intégration numérique.

Des fonctionnalités graphiques 2D et 3D pour la visualisation des données, des outils pour la construction d'interfaces graphique personnalisées, ainsi de fonctions pour l'intégration d'algorithmes développés en langage MATLAB, dans des applications et langage externes.

#### 3.2. Présentation de l'application détecte\_micro

##### 3.2.1. Fenêtre principale



Figure (4.4) : Fenêtre principale

### 3.2.2. Recueil d'information

Cette fenêtre est l'implémentation de la phase recueil d'information

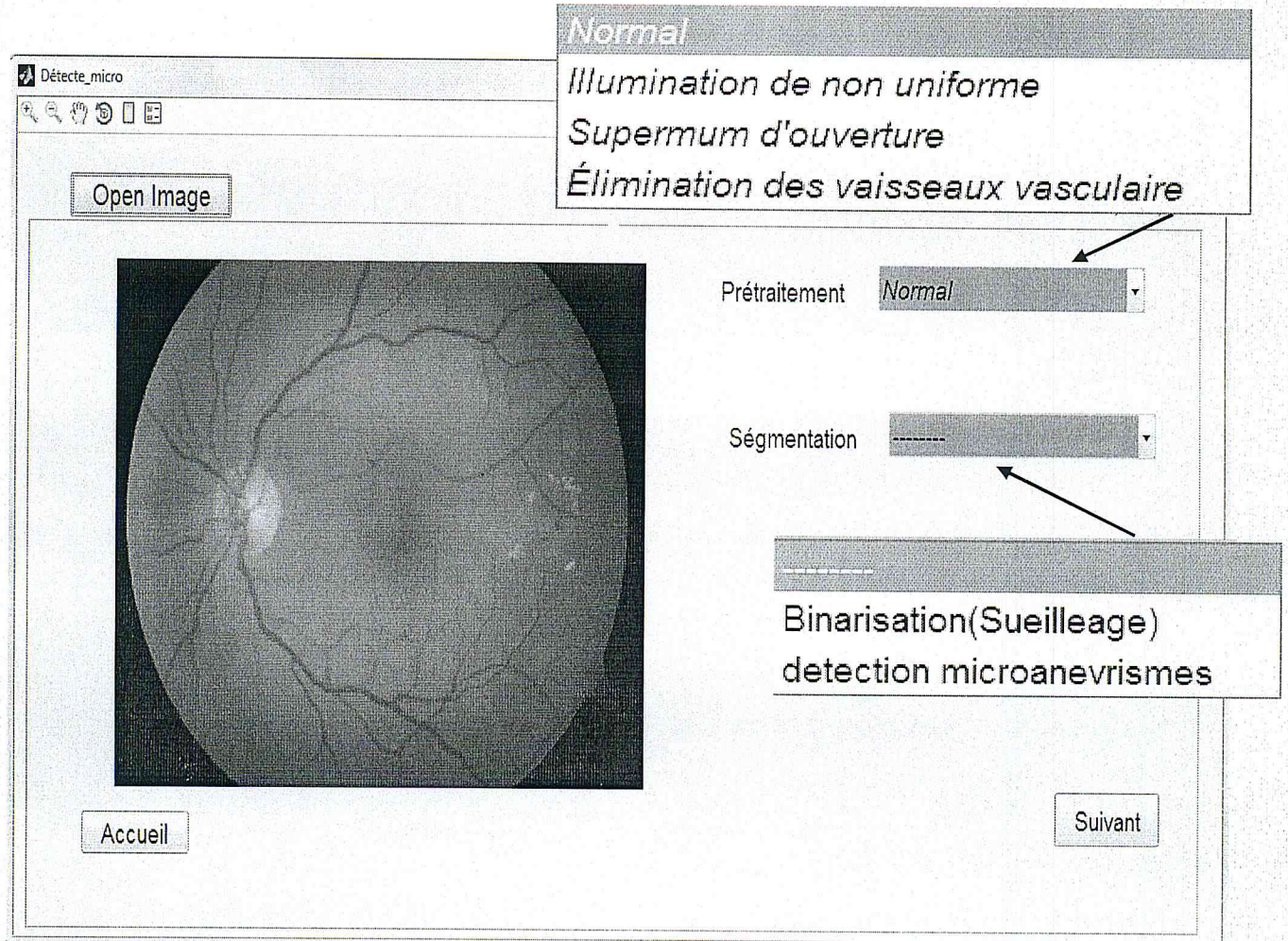


Figure (4.5) : Fenêtre Recueil d'information

Elle dispose d'un bouton « Open image », « recueil », « suivant », et des menus déroulant « prétraitement », « segmentation »

Le bouton Open image : à son exécution, une boîte de sélection apparait permettant à l'utilisateur de choisir le nom du fichier image à ouvrir.

Le bouton recueil : pour revenir à la Fenêtre principal

Le menu prétraitement : il contient trois choix

Illumination de non uniforme : a l'exécution un traitement de illumination de non uniforme s'applique sur l'image et l'affichera a la place de l'original

Supremum d'ouverture : à son exécution un traitement de supremum s'appliquera sur l'image amélioré avec illumination de non uniforme et l'affichera a sa place

## Chapitre 4 : détection des microanévrismes

Elimination des vaisseaux vasculaire : à son exécution on soustraire l'image améliorer de l'image supremum, et affichera sur le panneau.

Le menu segmentation : il contient deux choix

La binarisation (seuillage) : à son exécution on binarise l'image déjà prétraité avec une valeur de seuil

Détection microanévrismes : on détecte les vrais microanévrismes sur l'image binarisé.

Le bouton suivant : à son exécution un calcul de probabilité et nombre de microanévrismes sera déterminer après il va à la phase suivante

### 3.2.3. Fenêtre recherche de cas

A l'exécution de la Fenêtre la base de cas s'affichera

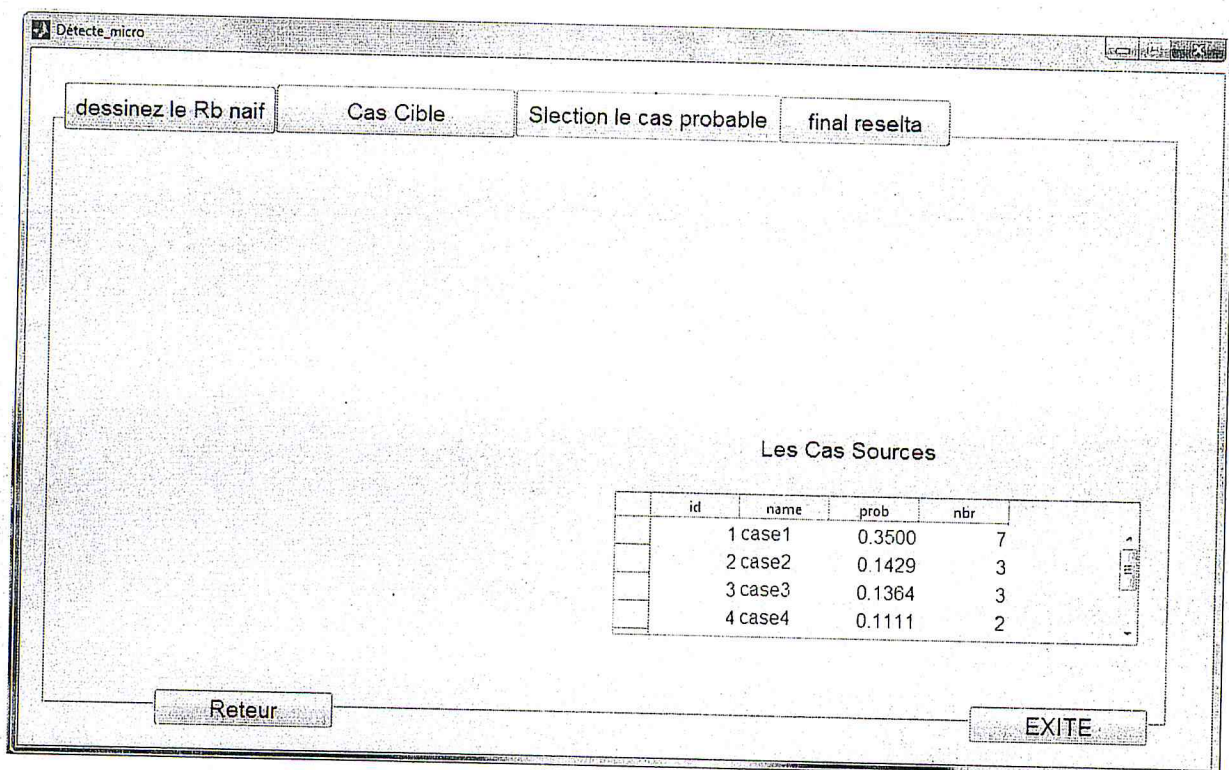


Figure (4.6) : Recherche d'un cas

Les cas sont structuré par un « id », « name », « prob », « nbr » :

Id : il est le numéro de cas

Name : c'est le nom attribue au cas.

Prob : c'est la probabilité de détection

Nbr : c'est le nombre des microanévrismes détecté.

## Chapitre 4 : détection des microanévrismes

Elle dispose des boutons « dessiner le Rb naïf », « Cas Cible », « Sélection le cas probable », « final résultat », « Retour », « EXIT »

Le bouton Retour : à son exécution on retour à la Fenêtre recueil d'information.

Le bouton EXIT : pour quitter l'application.

Le bouton dessiner le Rb naïf : à son exécution la construction et l'affichage de RB a partir des cas sources va être fait.

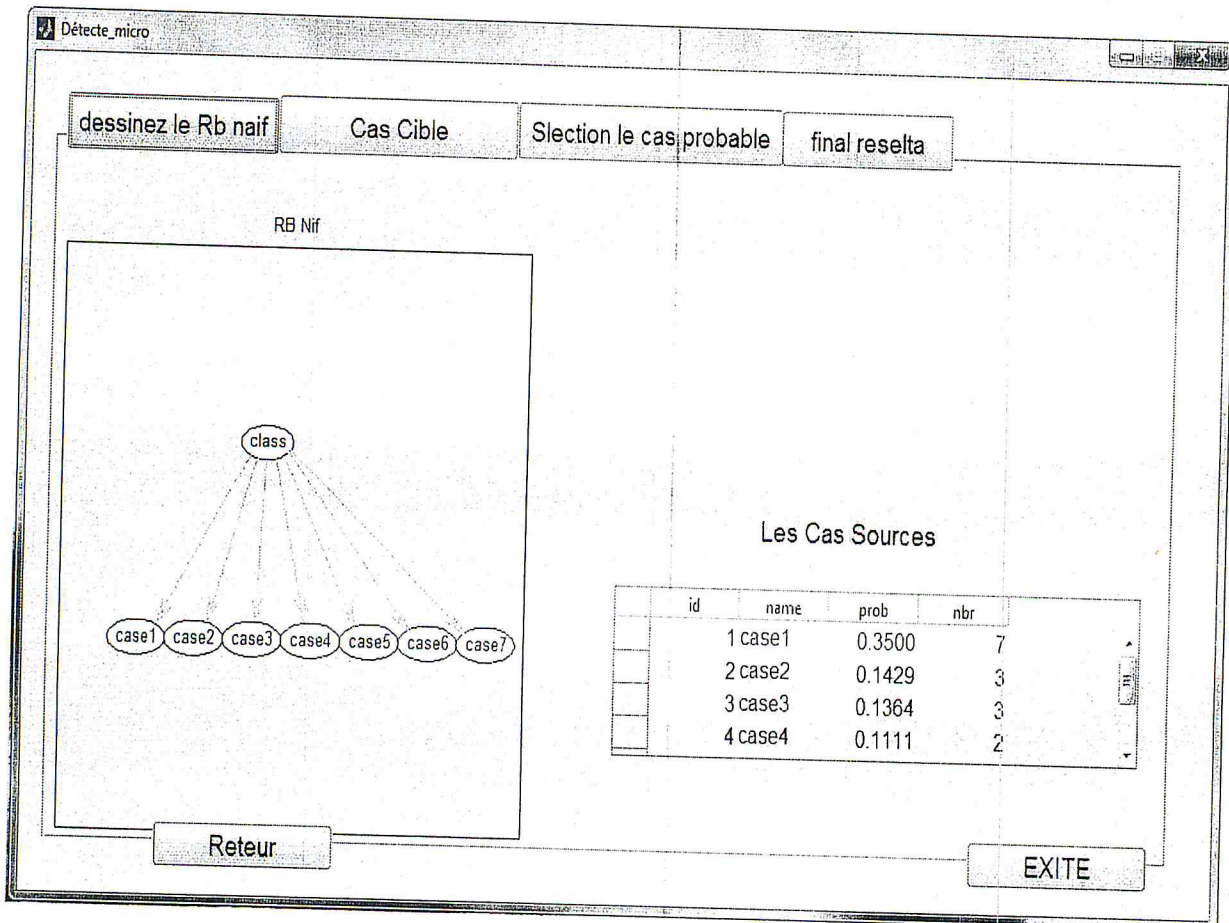


Figure (4.7) : Fenêtre construction de RB

Le bouton cas cible : à son exécution il affichera la structure de cas cible.

Le bouton Sélection le cas probable : à son exécution on retrouve les cas les plus probables comme montre la figure suivante :

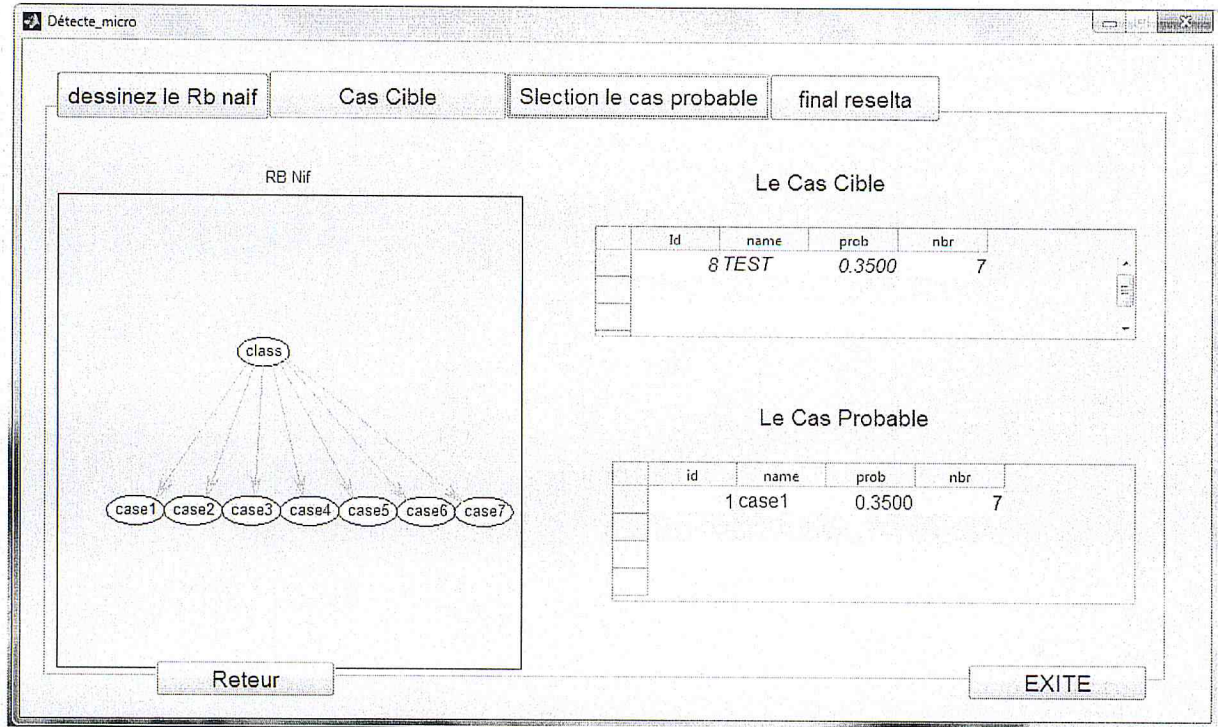


Figure (4.8) : Sélection le cas probable

Le bouton Final résultat : il affiche le résultat final et l'image cible

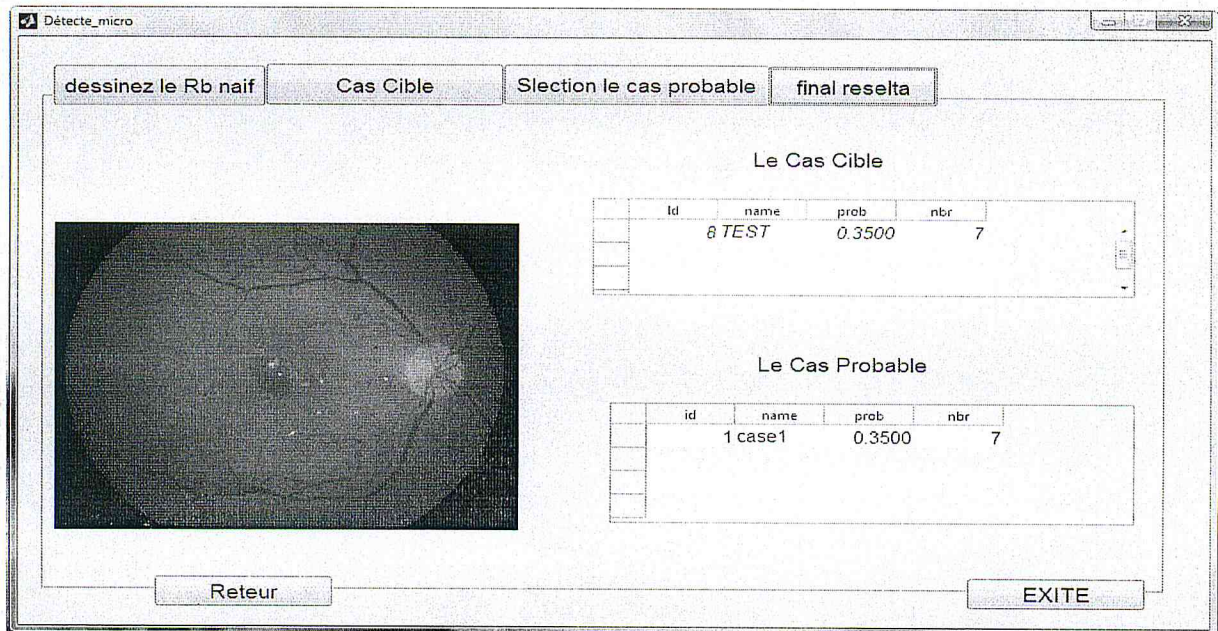


Figure (4.9) : Final résultat

### 3.3. Application et interprétation des résultats

Les différentes étapes nécessaires pour la détection automatique des microanévrismes sont présentées ci-dessous:

#### 3.3.1. Phase recueil des informations

- 1) Il faut tout d'abord charger l'image en couleur de la rétine, qui passera par la suite par plusieurs traitements : le prétraitement et la segmentation
- 2) Une correction d'illumination est ensuite réalisée. Les régions candidates recherchées se caractérisent par leurs petites tailles et les distinguer du bruit est très difficile. Pour corriger l'illumination non uniforme, nous avons utilisé un filtre moyen avec une taille de fenêtre qui permet de favoriser les microanévrismes par rapport aux autres structures dans l'image ayant une taille plus grandes. L'image corrigée est obtenue par la soustraction de l'image filtrée de l'image originale.

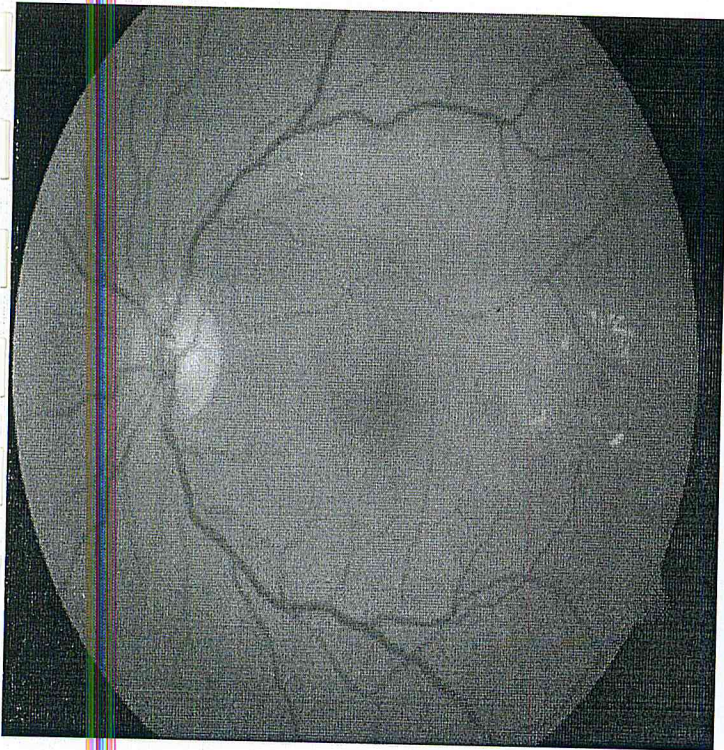


Image Original

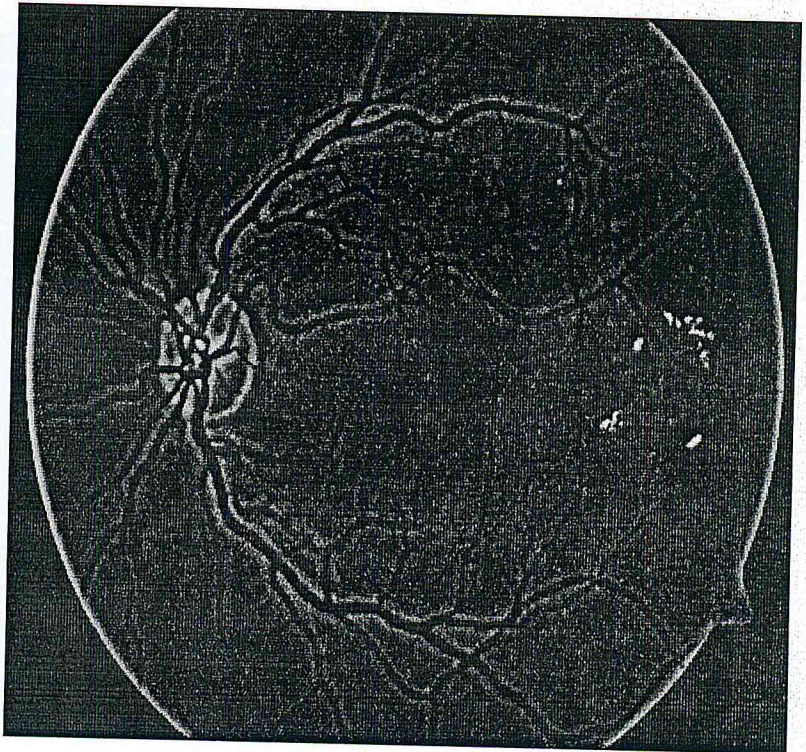


Image corrigé

Figure (4.10) : L'image après correction d'illumination non uniforme

- 3) Le supremum d'ouvertures est appliqué en considérant un élément structurant linéaire. Pour que ce filtrage morphologique soit assez discriminant, la taille de l'élément structurant utilisé doit être

relativement élevée. Après plusieurs tests, nous avons obtenu de bons résultats avec une taille égale à 17 dans 12 directions différentes en type linière

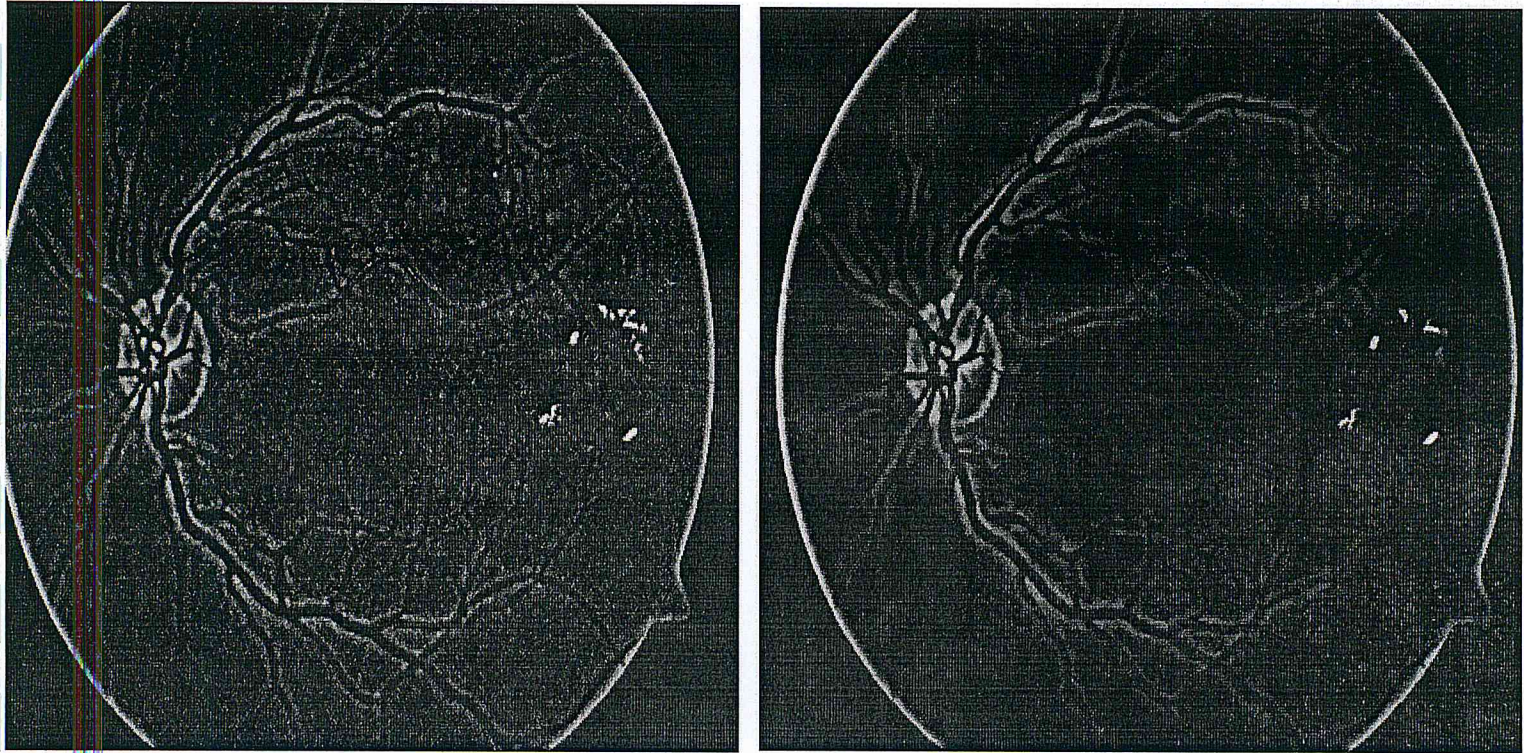


Image corrigé

L'image obtenue après l'application de supremum d'ouvertures

Figure (4.11) : L'image après le supremum

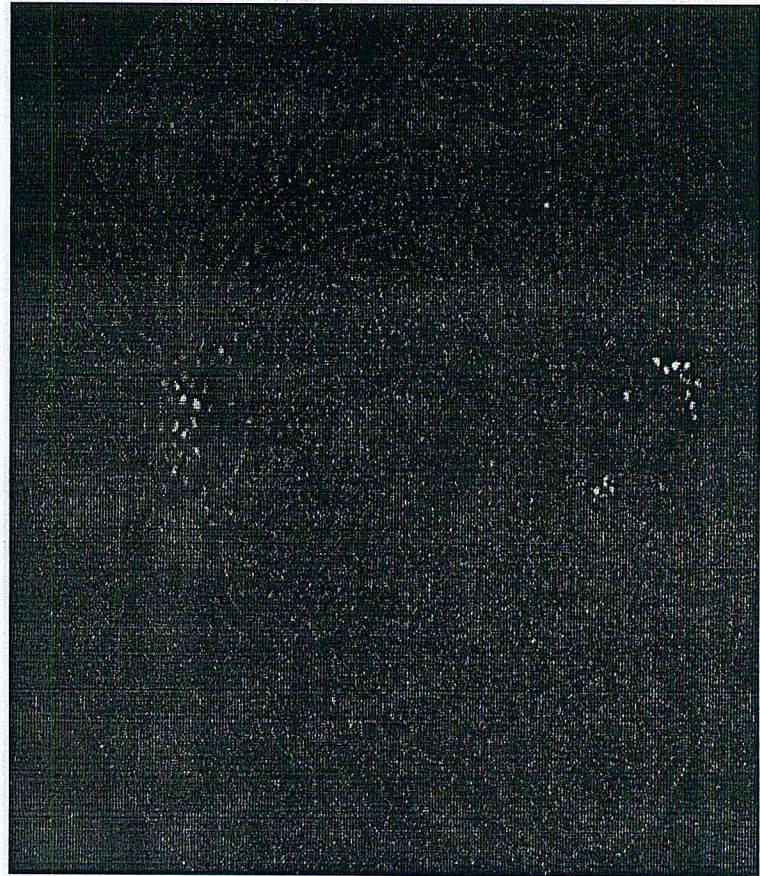
#### 4) Elimination des vaisseaux vasculaire :

Un chapeau haut de forme est appliqué sur l'image améliorée pour éliminer le vaisseau vasculaire. Le chapeau haut de forme comme illustre la figure suivante permet la distinction entre les vaisseaux vasculaire à l'aide d'un élément structurant.

L'application de chapeau haut de forme pour éliminer les structure de forme linière en soustraire l'image obtenu par le supremum d'ouverture a l'image corrigé



L'image obtenue après l'application de supremum d'ouvertures



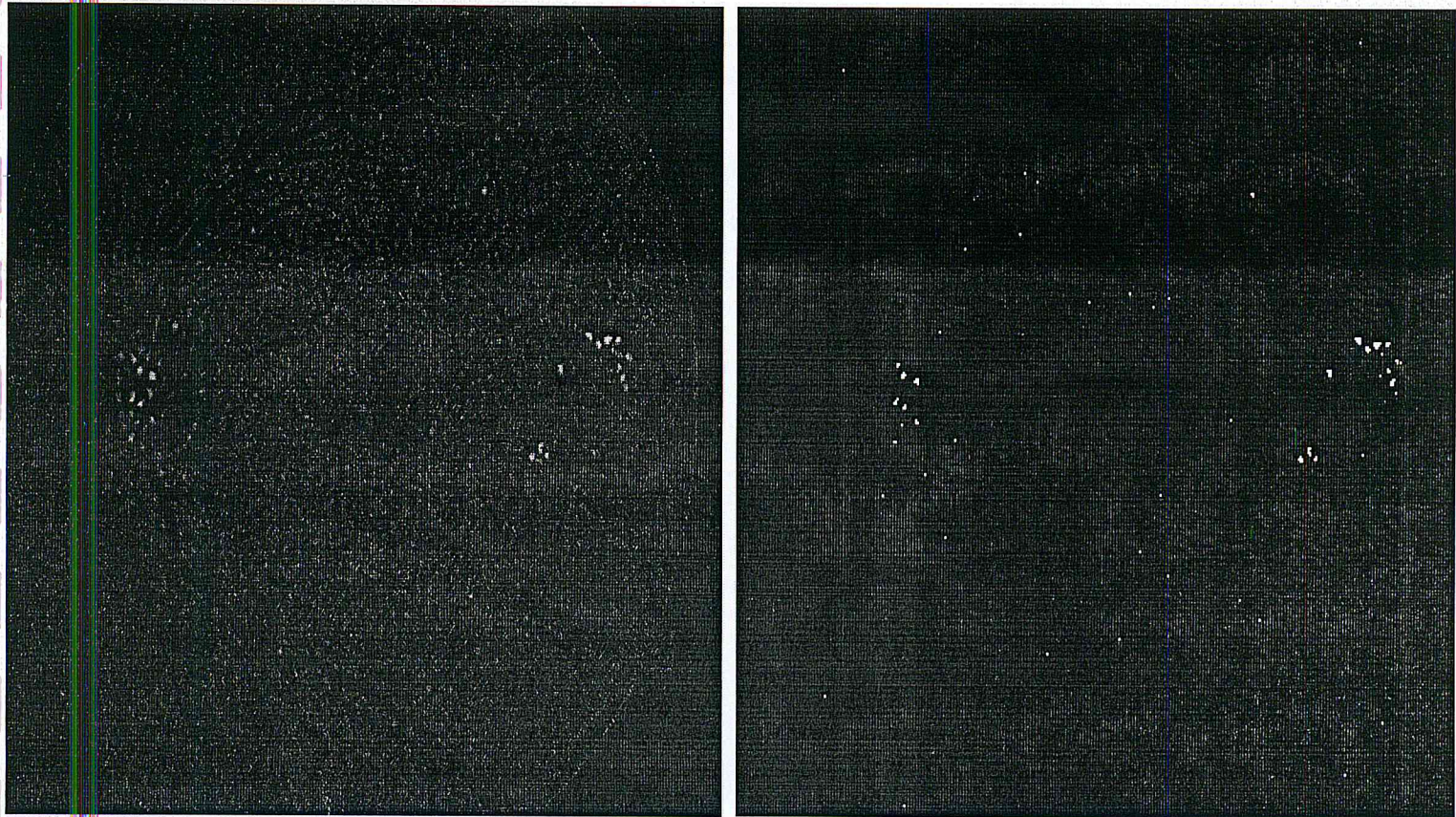
L'image obtenue après l'élimination

**Figure (4.12) :** L'image après L'image après l'application de chapeau haut de forme

5) Binarisation de l'image : cette étape repose sur le choix d'un seuil pour la détection des objets recherchés. Binariser une image revient à segmenter l'image en deux classes : le fond et l'objet. Un seuil mal choisi peut générer un grand nombre de faux positifs ou des vrais positifs peuvent être ratés

Le choix de seuil est très important pour avoir une image qui nous aide à détecter les microanévrismes on a choisis le seuil en calculant la construct moyen de l'image plus une valeur calculer par le nombre de pixel de l'image et la construct moyen.





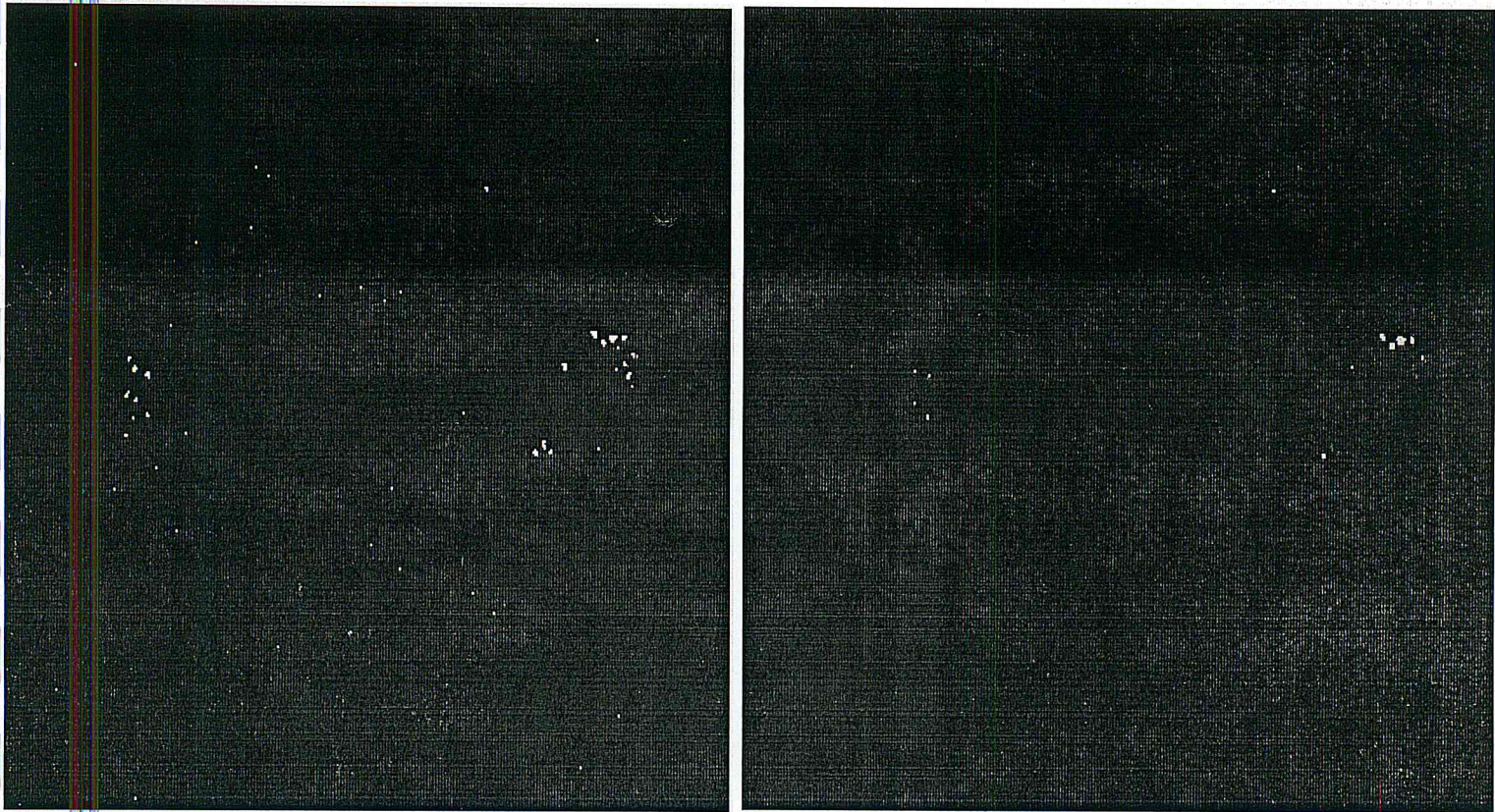
L'image obtenue après l'élimination  
des vaisseaux vasculaires

Binarisation

**Figure (4.13)** : L'image obtenue après la binarisation

6) détection des microanévrismes : cette étape repose à sélectionner les vrais microanévrismes par leur structure (disque) est leur diamètre.

Le choix de mètre une détection vrai ou faux microanévrisme est une étape importante donc on a choisit un diamètre pour désigner les vrais microanévrismes en sélectionnent que les structures circulaire (disque), on peut avoir des faux microanévrismes désigné comme vrais, pour cela en a éliminé tous les microanévrismes qui situé en d'hors de la fin des vaisseaux vasculaires.



Binarisation

b) obtenu après détection microanévrismes

**Figure (4.14) : Détection des microanévrismes**

On voit sur l'image b de type des éléments vert et bleu, les éléments verts sont les vrais et les bleus sont suspects comme faux microanévrisme

7) diagnostique final : dans cette phase on superpose les microanévrismes détecté sur l'image originale pour obtenir une image avec des microanévrismes détectés

La super position de l'image obtenu par la détection des microanévrismes on super positionné les microanévrismes sur l'image original pour bien confirmer la fiabilité de détection

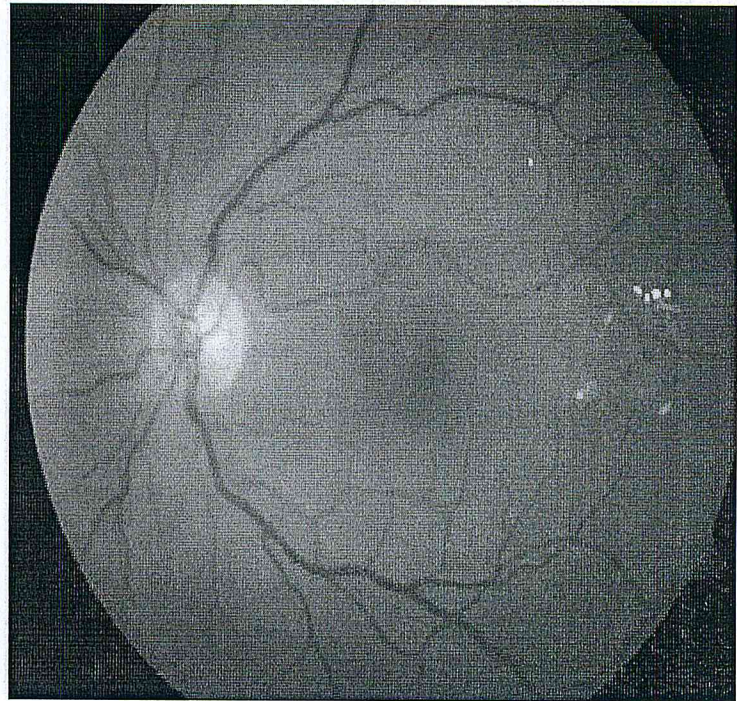
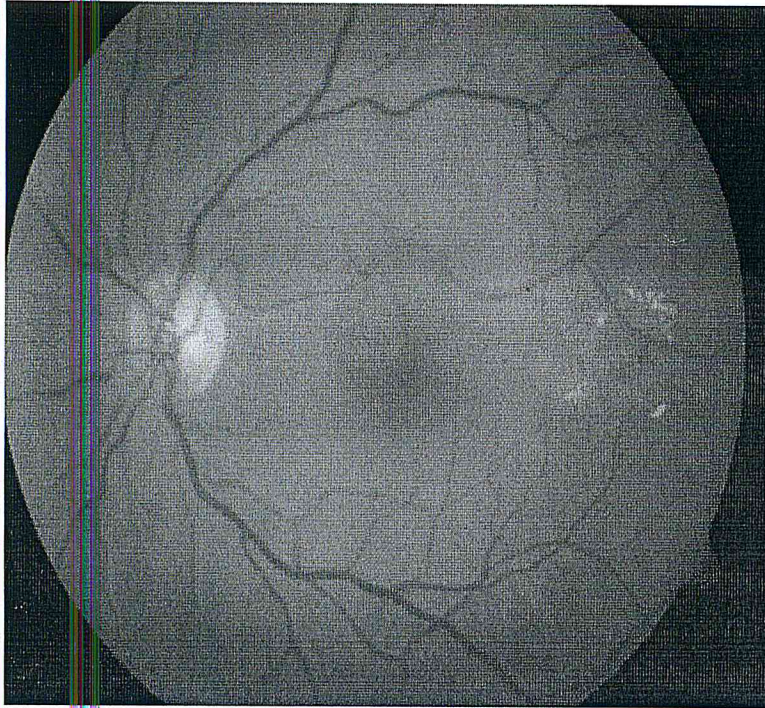


Image original

Image final

**Figure (4.15) :** Après détection des microanévrismes

## 4. Conclusion

Dans ce chapitre nous avons présenté le système développé, Détecte\_micro, que nous avons réalisé dans le but de la détection des microanévrismes sur des images rétinienne pour l'aide a diagnostique médicale. Nous avons décrit l'approche adoptée pour l'extraction d'informations pertinentes à partir de l'image rétinienne- Après avoir détecté les régions d'intérêt, nous avons réalisé une étape de classification des cas avec des réseaux bayésiens (RB), et la recherche de cas le plus probable dans le RB, qui correspond à un vrai microanévrisme.

# Conclusions generales

## Conclusion Générale

---

Nous avons abordé dans ce travail, l'analyse d'images rétinienne par le raisonnement à partir de cas et les réseaux bayésiens pour l'aide au diagnostic médical. Le raisonnement à partir de cas (RàPC) est un outil qui permet de capitaliser les connaissances acquises dans une base de cas, afin d'émuler le comportement de l'ophtalmologiste. Le RàPC évite ainsi la perte des connaissances, en préservant toutes les expériences pertinentes nécessaires pour la prise de décision. Les réseaux bayésiens (RB) qui nous permis d'élaborer la classification des cas pour enrichir l'incertitude médical en utilisant dans le RàPC on a combiner deux grandes méthode pour aide a le diagnostic Médical ,les RB sont des bons classificateurs pour la base de cas

Les images médicales sont principalement destinées à l'observation humaine. Actuellement, elles sont analysées automatiquement par des systèmes intelligents assurant la surveillance et le suivi des maladies. L'analyse d'images rétinienne par le RàPC doit traiter le problème de représentation d'informations extraites, dans une structure appropriée, à partir de l'image.

Nous avons proposé et réalisé une approche pour la détection automatique des microanévrismes. La première étape est l'étape d'accueil d'information. Elle consiste tout d'abord à améliorer l'image rétinienne. Altérée par les conditions d'acquisition, et qui se caractérise par un faible contraste. Un ensemble de traitements morphologiques permet par la suite l'élimination des vaisseaux vasculaires. L'application du supremum d'ouvertures, a permis l'isolation de candidats susceptibles de représenter des microanévrismes. Ensuite, la segmentation de l'image a été réalisée par l'application binarisation (seuillage) un traitement de détection est appliqué afin de bien délimiter les détections obtenu Finalement, les probabilités sont calculées. La deuxième étape de notre système, est la recherche des cas probable dans la base des cas déjà stokes. La base de cas contient un ensemble de cas sources (ensemble d'apprentissage). Chaque cas source correspond à un candidat.

Plusieurs perspectives peuvent être envisagées dans le prolongement de ces travaux de thèse.

-D'autre amélioration pourrons être apporté à notre travaille comme l'analyse des images rétinienne pour la détection et la classification d'autre structures pathologique comme les exsudats, les hémorragies, etc.

-Détecer d'autres maladies liées à la RD tels que les néo vaisseaux et les nodules cotonneux et d'autres pathologies qui affectent la rétine à savoir : le glaucome, la cataracte.

## *Conclusion Générale*

---

- Un système de classification de la RD doit être mis en œuvre afin de faciliter aux ophtalmologues l'identification du stade de la RD.

- Comme les patients sont censés avoir un examen régulier de la rétine, il est important de développer une procédure de détection de changement progressif dans la rétine au fil du temps qui pourrait être intéressant pour le dépistage de la RD dans le long terme.

## Bibliographie :

- [1] Le corps humain : Etude, structure et fonction, Brooker de boeck, 2 édition 2001
- [2] Site de RETINA France, association française de patients, Website: <http://www.retina.fr/ret50-anatomie.html> visité : mai 2014
- [3] The Bioinformatics Organisation: La perception des couleurs par l'œil: chapitre 1.L'œil. Website : <http://www.bioinformatics.org/oeil-couleur/dossier/anatomie.html> visité : mai 2014
- [4] BACHA Siham, Samah BENDALI Analyse d'image rétinienne pour l'aide au diagnostic médical par résonnement à Base de Cas, Université de Saad Dahleb de Blida (Bibliothèque Pav 14, réf : MA-004-49-2), 2007
- [5] Goupe Quantal, Website: <http://www.quantel-medical.fr/pathologies-retinopathie.php> visité : aout 2014
- [6] Diabète Rétinopathie Website : [http://santecheznous.com/channel\\_condition\\_info\\_details.asp?disease\\_id=196&channel\\_id=143&relation\\_id=57197#Symptoms](http://santecheznous.com/channel_condition_info_details.asp?disease_id=196&channel_id=143&relation_id=57197#Symptoms) visité : aout 2014
- [7] Organisation mondiale de la santé. Diabète sucré. Website : <http://www.who.int/mediacentre/factsheets/fs138/fr/index.html> visité : mai 2014
- [8] Organisation mondiale de la santé. Diabète sucré. Website : <http://www.who.int/mediacentre/factsheets/fs312> visité : mai 2014
- [9] Université Médicale Virtuelle Francophone site web : <http://umvf.univ-nantes.fr/ophtalmologie/enseignement/ophtalmo16/site/html/cours.pdf> visité : mai 2014
- [10] F. Boudierba, A. yahia, Analyse topologique des images mammographiques, mémoire d'ingénieur, Université Saad Dahleb Blida, 2009.
- [11] M. Nicols Passat, Contribution à la segmentation des réseaux vasculaires cérébraux obtenus en IRM. Intégration de connaissance anatomique pour le guidage d'outils de morphologie mathématique. Thèse de doctorat, L'université Louis Pasteur de Strasbourg - Ecole MSII, 2005.
- [13] Segmentation par approche contours, Website : <http://titaude30.wordpress.com/2009/03/19/segmentation-par-approche-contours/> visité : aout 2014
- [14] T. Schouten, Segmentation contour based, Website: <http://www.es.ru.nl/~ths/rt2/col/h7/7contourENG.html> visité : aout 2014

**[15] Document internet : raisonnement par cas, Wikipédia fondation.Inc.website :**

[http://fr.wikipedia.org/wiki/Raisonnement par cas](http://fr.wikipedia.org/wiki/Raisonnement_par_cas) visité : aout 2014

**[16] Document internet web site :**

<http://www.linternaute.com/dictionnaire/fr/definition/raisonner-par-analogie/> visité : aout 2014

**[17] Document internet website :**

[http://liris.cnrs.fr/amille/enseignements/Master PRO/TIA/RAPC/igc rapc Folder/rapc session3 cycle raisonnement.html](http://liris.cnrs.fr/amille/enseignements/Master_PRO/TIA/RAPC/igc_rapc_Folder/rapc_session3_cycle_raisonnement.html) visité : aout 2014

**[18]** Philipp-Folguet, S. et Lekkat, M. "Recherche d'images à partir d'une requête partielle utilisant la disposition des régions". Rapport de recherche, Equipes Traitement des Images et du signal, ETIS, octobre 2003.

**[19]** Siddiqi, K., Shokoufandeh, A., Dickinson, S.J., et Zucker, S.W. "Shock graphs and shape matching". Dans Sixth International Conference on Computer Vision, Bombay, Inde, 1998.

**[20]** Boughorbel, S., Boujemaa, N. et Vertan, C. "Description de la répartition spatiale de la couleur pour l'Indexation d'Images". RFIA'2002, vol. 2 pp. 407-414

**[21]** Kolodner J., Case-Based Reasoning. Morgan Kaufmann, San Mateo, Publishers, Inc, UCA, 1993.

**[22]** Bareiss E. R., PROTON: A Unified Approach to Concept Representation, Classification and Learning. PhD Thesis, University of Texas, 1988.

**[23]** Boucher A., Réseaux Bayésiens. Rapport du travail d'Intérêt personnel Encadré, Semestre II, pp. 1-29, 2005.

**[24]** Bozec C., Jaulent C.M., Erie Z. et Partrice D., Accès Internet à IDEM images et diagnostics par l'exemple en médecine. Santé et Réseaux Informatiques, Vol 10, 1998.

**[25]** Brossier J.M., Une introduction aux modèles graphiques. Laboratoire des images et des signaux, Février 2005.

**[26]** Finn V.J., An introduction to Bayesian Networks. Eyrolles, Première Edition, 1996.



[27] Heinisch R. H., Weber R., Martin A. et Barcia R. M., Industrial Engineering and Cardiology Service of University Hospital. Florianopolis, SC, Brazil, 1999.

[28] Hénaut A., Corvol P. et Degoulet P., Nouvelle méthode de traitement de l'information en médecine, 1992.

[29] Gebhardt F., Voß A., Gräther W. et Schmidt-Belz B., Reasoning With Complex Cases. Kluwer academic publishers, 1997.

[30] Kolodner J., Workshop on case-based Reasoning. Editor DARPA 88, Clearwater, Morgan Kaufmann, San Mateo, Florida, 1988.

[31] Konton P., Reasoning about Evidence in Causal Explanations. In proceedings of the workshop on Case-Based reasoning (DARPA), pp. 260-270, 1998.

[32] King L. S., What is a diagnosis?. JAMA: The Journal of the American Medical Association, Vol 202(8), pp. 714-717, Novembre 1967.

[33] Leray P. et Olivier F., Réseaux Bayésiens pour la Classification : Méthodologie et Illustration dans le cadre du Diagnostic Médical. RIA, 15/2002, Réseaux Bayésien, pp. 1-25, 2002.

[34] Leray PH. et Gallinari P., Une architecture neuro-bayésienne pour le traitement spatio-temporel d'alarmes, Application au diagnostic dans le réseau téléphonique. In Journées Nationales sur les Modèles de Raisonnement, JNMR'99, pp. 134-145, 1999.

[35] Mille A., Les Réseaux Bayésiens A la recherche de la vérité. Cours Cognition et connaissance, Université Claude Bernard Lyon 1, pp. 1-9, 2007.

[36] Minsky M., A framework for representing knowledge. In the psychology of Computer Vision, Editeur P.H. (Ed.) Winston, pp. 211-279, New York, McGraw Hill, 1975.

[37] Naim P., Wuillemin P., Leray H., Pourret P.O., et Becker A., Réseaux Bayésiens. Groupe Eyrolles, ISBN 2-212- 11371, 2004.

[38] Nilson M. et Sollenborn M., *Advancements and Trends in Medical Case Resoning: An Overview of Systems and System Development*. 2004.

[39] Pearl J., *Probabilistic reasoning intelligent systems: networks of plausible inference*. Morgan kaufman, Palo Alto, 1988.

[40] Pewsner D., Bleuer J., Bucher P. et Battaglie H.C., *Sur la voie de l'intuition : Théorème de Bayes et Diagnostic en médecine générale*, pp. 41-45, 2001.

[41] Richter M.M., *Introduction In Case-Based Reasoning Technology: From Foundations to Applications*. Edited by M. Lena, B. Bartsch-Sporl, H. D. Burkhard, et S.Wess. Springer-Verlag, Berlin, pp. 1-15, 1998.

[42] Rose C., *Modélisation stochastique pour le raisonnement médical et ses applications à la télémédecine*. Thèse de doctorat, l'université Henri Poincaré, Nancy 1, Mai 2011

[43] Schmidt R. et Vorobieva O., *Combining Statistics and Case-Based Reasoning for Medical*. Research C.L. Mumford and L.C. Jain (Eds.): *Computational Intelligence, ISRL 1*, pp. 673-696, Springer-Verlag, Berlin Heidelberg, 2009.

[44] Smyth P., *Belief networks, hidden Markov models, and Markov random fields a unifying view*. *Pattern Recognition Letters*, N° 18, pp. 1261-1268, 1997.

[45] Souafi S., *Contribution à la reconnaissance des structures des documents écrits: Approche probabiliste*. Thèse de doctorat, l'institut national des sciences appliquées de Lyon, 21 Septembre 2002.

[46] Turner R., *Organizing and using schematic knowledge for medical diagnosis*. In: Kolodner, J. (Ed.): *Proceedings of the Workshop on Case-Based Reasoning*, pp. 435-446, Florida, 1988.

[47] Veloso M., Carbonell J., Pérez A., Borrajo D., Fink E. et Blythe J., *Integrating Planning and Learning: The PRODIGY Architecture*. *Journal of Experimental and Theoretical Artificial Intelligence*, Vol 7(1), pp. 81-120, 1995.

[48] Watson I., Applying Case-Based Reasoning: Techniques for Enterprise Systems. Morgan Kaufmann Publishers Inc, 1997.

[49] Schank R.C., Dynamic Memory: A Theory of Reminding and Learning in Computers and People. Cambridge University Press, New York, NY, 1982.

[50] Richter M.M., Wess S., Althoff K.D. et Maurer F., First European Workshop on Case-Based Reasoning, University of Kaiserslautern, Germany, Lecture Notes in Artificial Intelligence, Vol 837, Springer Verlag, Berlin, 1993.

[51] HajSaid A., Distances sémantiques pour la comparaison des connaissances objets dans le cadre du raisonnement à partir de cas. Mémoire de DEA Informatique Théorique et Applications, Université de le Havre, pp. 25-40, 2004.

[52] Afouba N., Kerbrat S. et Labarang Z., Rapport du projet d'intelligence artificielle : Le raisonnement à partir des cas : Définitions et principes de fonctionnement, Novembre 2004.

[54] Meyer N., Vinzio S. et Goichot B., La statistique bayésienne : une approche des statistiques adaptée à la clinique. La revue de médecine interne, Vol 30(3), pp. 242-249, 2009.