

## UNE APPROCHE QUALITATIVE SUR LE TRAITEMENT DE L'INCERTAIN: APPLICATION AU SYSTEME EXPERT

Reçu le 16/10/2000 – Accepté le 12/01/2003

### Résumé

Nous présentons un système expert de classification automatique supervisé basé sur une approche symbolique. Ce dernier est constitué de deux sous-systèmes. Le premier sous-système génère automatiquement les règles de production à partir d'un ensemble d'apprentissage; les règles générées sont pondérées par un degré de croyance symbolique qui caractérise leurs classes d'appartenance. Le deuxième est le système d'inférence qui reçoit en entrée la base de règles et l'objet à classer; ensuite, à l'aide du raisonnement classique (le Modus Ponens) il fournit la classe d'appartenance de cet objet avec un certain degré de croyance symbolique. Les méthodes d'évaluation du degré de croyance sont nombreuses, mais souvent entachées d'incertitudes. Pour apprécier les performances de notre approche symbolique, nous avons effectué des tests sur deux images: la cryosection de la cuisse humaine et de la texture.

**Mots clés:** Incertitude qualitative, systèmes experts, probabilité symbolique, représentation des connaissances, logique multivaluée.

### Abstract

We present an expert system of supervised automatic classification based on a symbolic approach. This last is composed of two sub-systems. The first sub-system automatically generates the production rules using training set; the generated rules are accompanied by a symbolic degree of belief which characterizes their classes of memberships. The second is the inference system, which receives in entry the base of rules and the object to classify. Using classical reasoning (Modus Ponens), the inference system provides the membership class of this object with a certain symbolic degree of belief. Methods to evaluate the degree of belief are numerous but they are often tarnished with uncertainty. To appreciate the performances of our symbolic approach, we have made tests on two images: the cryosection of the human thigh and texture.

**Keywords:** Qualitative uncertainty, expert systems, symbolic probability, knowledge representation, multivalued logic.

H. SERIDI<sup>1, 2</sup>

H. AKDAG<sup>2, 3</sup>

A. MEDDOUR<sup>4</sup>

<sup>1</sup> LAIG, Université 08 Mai 45  
B.P. 401, 24000 Guelma, Algérie

<sup>2</sup> LERI, Université de Reims

B.P. 1035, 51687 Reims

Cedex 02, France

<sup>3</sup> LIP6 Université P. & M. Curie

75687 Paris Cedex 05, France

<sup>4</sup> Université 08 Mai 45

B.P. 401, 24000 Guelma, Algérie

### ملخص

نقدم نظام خيرة لترتيب أو توماتيكي مراقب على أساس مقارنة رمزية. هذا النظام مكون من نظامين جزئيين. النظام الأول يولد ألياً قواعد الإستنتاج ابتداءً من مجموعة تدريبات؛ القواعد المولدة مرفوقة بدرجة الإعتقاد الرمزي التي تميز الأصناف التي تنتمي إليها. الثاني نظام الإستلزام الذي يستقبل عند الدخول مجموعة القواعد الموضوع المطروح للتصنيف، ثم بواسطة الإستدلال الكلاسيكي (موديس بوننس) يعطي الصنف الذي ينتمي إليه هذا الموضوع بدرجة معينة حسب درجة الإعتقاد الرمزي. طرق تقييم درجة الإعتقاد عديدة، لكنها في أغلب الحالات مشوية بإرتيابات. لتقدير الأداءات لمقاربتنا الرمزية، قمنا بتجارب على صورتين: صورة مقطع من فخذ بشري وصورة من نسيج.

**الكلمات المفتاحية:** إرتيابات نوعية، نظام خيرة، احتمالات رمزية، عرض معارف، منطق متعدد القيم.

De nos jours, l'automatisation du traitement des données et/ou la prise de décision devant des problèmes complexes, a conduit l'homme à l'élaboration des systèmes intelligents connus sous le nom de systèmes experts. Ces derniers ont connu une panoplie d'applications dans différents domaines scientifiques. Un aspect qui préoccupe les chercheurs dans le domaine de l'intelligence artificielle est la réalisation de systèmes experts pour la reconnaissance des formes, en l'occurrence la classification d'images.

Les travaux de Borgi [1] sur différentes bases de données réalisés avec les méthodes numériques d'évaluation des degrés de croyances (degrés de probabilités, de possibilités ou bien des coefficients de certitude) ont montré que dans son système SUCRAGE, les taux de la bonne classification sont obtenus par l'approche probabiliste. Une autre étude [1] comparative a été faite aussi avec d'autres méthodes telles que les systèmes d'inférence flou [2-6], les graphes d'induction et les arbres de décision [7-11]. Il a été montré, à travers cette étude, que les taux de bonne classification obtenus dans SUCRAGE sont nettement supérieurs à ceux obtenus avec les SIF, les graphes d'inductions et les arbres de décision. Par contre, le nombre de règles fourni par les SIF, les graphes d'inductions et les arbres de décision est relativement inférieur à celui généré par le générateur de SUCRAGE.

Compte tenu des différents travaux précités et plus particulièrement ceux de Vernazza [12] et Borgi *et al.* [13] qui consistent à générer automatiquement des règles de production par apprentissage supervisé et qui sont destinées à la reconnaissance des formes en imagerie, nous proposons une nouvelle approche symbolique d'évaluation du degré de croyance qui accompagne la classe d'appartenance. Notre contribution est portée sur l'intégration d'une interface homme-machine au niveau du générateur de règles et la T-conorme probabiliste symbolique au niveau du système d'inférence [14 -16].

Le système expert sur lequel s'applique notre approche symbolique est constitué de deux sous-systèmes: le générateur de règles et le système d'inférence. Le traitement au niveau du générateur commence par la mise sous un format particulier des données de l'ensemble d'apprentissage (puisque'il s'agit d'un apprentissage supervisé), ensuite le module de discrétisation assure la partition régulière de chaque attribut de l'ensemble d'apprentissage [13]. Par la méthode de corrélation linéaire, on recherche les éléments corrélés et on recense les composantes connexes maximales qui permettent de générer les règles de production à prémisses multivaluées [12, 13]. Les degrés de croyances qui accompagnent ces règles sont des probabilités numériques résultant d'une appréciation objective fondée sur les hypothèses d'appartenance des prémisses à une classe. A ce stade, nous traduisons ces probabilités numériques en probabilités symboliques grâce à l'interface homme-machine. Rappelons toutefois que les degrés de croyance utilisés dans la littérature, peuvent être des probabilités [17], des coefficients de certitude [18] ou des degrés de possibilités [19].

Le système d'inférence reçoit en entrée cette base de règles ainsi que les nouvelles données discrétisées, mais démunies de leurs classes d'appartenance. En se basant sur le raisonnement exact, "le Modus Ponens classique" et la T-conorme probabiliste symbolique, le système d'inférence fournit à sa sortie des données classées. Le degré avec lequel chaque donnée (vecteur représentatif de l'objet à classer) appartient à une classe est de nature symbolique.

## GENERATEUR AUTOMATIQUE DE REGLES DE PRODUCTION

La phase d'apprentissage consiste à générer automatiquement les règles de la forme:

Si [prémisse] Alors [conclusion]

Une telle règle est notée:

prémisse  $\rightarrow$  conclusion

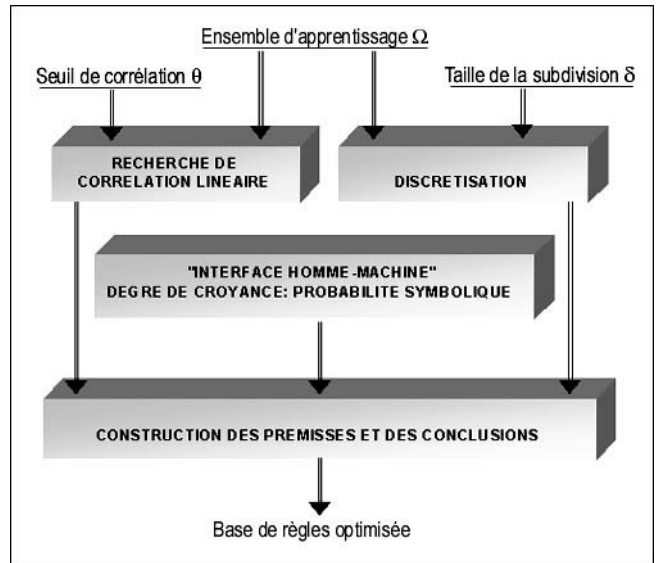
La morphologie des règles générées par notre système est la suivante:

$A_1$  et  $A_2$  et...et  $A_n \rightarrow y, \alpha$

$A_j$  est une condition du type:  $X_j$  est dans  $[a, b]$ ;  $X_j$  est la  $j^{\text{ème}}$  composante du vecteur représentant un objet;  $[a, b]$  est un intervalle issu de la discrétisation des domaines de variation des attributs (ici de l'attribut  $X_j$ );  $y$  est une hypothèse sur l'appartenance à une classe et  $\alpha$  est un degré de croyance représentant l'incertitude de la conclusion.

Les différentes phases qui consistent à générer la base de règles sont représentées par le schéma de la figure 1;

quant à la définition de chaque phase, elle sera aussi donnée dans la suite de ce paragraphe.



**Figure 1:** Schéma représentant le générateur automatique de règles de production :  $\Omega$  est ensemble d'apprentissage,  $\theta$  est le seuil de corrélation et  $\delta$  est la taille de la subdivision.

## L'ensemble d'apprentissage

Les exemples de la base d'apprentissage doivent être mis sous un format particulier, puisque'il s'agit de la classification supervisée. L'ensemble d'apprentissage  $\Omega$  est constitué de  $N$  exemples représentés par des vecteurs numériques de dimension  $p$ , étiquetés par le nom de la classe à laquelle ils appartiennent. Les classes sont représentées par la variable  $Y$  qui prend ses valeurs dans un ensemble  $C$  discret et de cardinal  $C$ . Les classes sont notées  $y_1, y_2, \dots, y_C$ . La représentation d'un exemple  $\pi_i$  de l'ensemble d'apprentissage  $\Omega$  est la suivante:  $\pi_i$  est représenté par  $X_i = \{X_1^i, X_2^i, \dots, X_p^i\}$  de classe  $Y(\pi_i)$ .

## Discrétisation

La méthode de discrétisation utilisée est celle qui a été proposée par Borgi *et al.* [13]; elle consiste à construire une partition régulière. Cette méthode est non supervisée dans la mesure où elle ne tient pas compte de la connaissance des classes. Seule l'information relative à l'attribut est prise en compte, indépendamment de tout autre attribut. Si  $\delta$  est la taille de la subdivision et  $R_i$  le domaine de variation de l'attribut  $X_i$ , alors  $R_i$  est découpé en  $\delta$  sous-intervalles de même amplitude  $L$ . Les sous-intervalles obtenus sont totalement ordonnés et notés:  $rg_0, rg_1, \dots, rg_{(\delta-1)}$ .

## Recherche de corrélation linéaire

La construction des règles de production est fondée sur une méthode polythétique [12] qui consiste à rechercher la corrélation linéaire entre les composantes de tous les vecteurs de l'ensemble d'apprentissage.

La première étape traite le calcul de la matrice de corrélation  $R$  entre les composantes des vecteurs de

l'ensemble d'apprentissage. La matrice  $R=(r_{i,j})_{1 \leq i \leq p, 1 \leq j \leq p}$  est symétrique et de dimension  $p \times p$ ;  $p$  étant le nombre total d'attributs. L'élément  $r_{i,j}$  ( $i, j \in \{1, 2, \dots, p\}$ ) de la matrice  $R$  est le coefficient de corrélation linéaire entre les composantes  $X_i$  et  $X_j$  des vecteurs de l'ensemble d'apprentissage.

La deuxième étape consiste à déduire de la matrice  $R$  une matrice seuillée  $R_{seuil}$ , c'est à dire décider à partir de quel seuil  $\theta$ , deux attributs  $X_i$  et  $X_j$  sont corrélés. La procédure de décision est définie comme suit:

Si  $|r_{i,j}| < \theta$  alors  $X_i$  et  $X_j$  ne sont pas corrélés et  $r_{i,j}$  est mis à 0,

Sinon  $X_i$  et  $X_j$  sont corrélés et  $r_{i,j}$  est mis à 1.

La troisième étape, basée sur la recherche de la partition orthogonale minimale, a pour but de regrouper les composantes corrélées de la matrice seuillée et d'extraire les composantes connexes maximales.

### Construction des prémisses et conclusions des règles

Une fois que nous disposons des composantes connexes, on regroupe les attributs corrélés pour chaque composante connexe dans les mêmes prémisses. Chaque attribut prend ses valeurs de l'ensemble totalement ordonné  $\{rg_0, rg_1, \dots, rg_{(\delta-1)}\}$ . Chaque prémisses construite conduit à la génération de "C" règles; "C" étant le nombre total de classes. Pour chaque prémisses, toutes les conclusions possibles d'appartenance à une classe sont générées. Donc, une règle est une association d'une prémisses à une classe  $y_i$  de  $\{y_1, y_2, \dots, y_C\}$  avec un certain degré de croyance, c'est à dire une certaine incertitude. Lors de la construction des règles, le degré de croyance, qui associe une prémisses à une conclusion, est une probabilité numérique. L'approche utilisée pour estimer la probabilité ( $y_i$ /prémisses), s'apparente à la formule Bayésienne:

$$\alpha = P(y_i / \text{prémisses})$$

Si une prémisses ne contient aucun exemple d'apprentissage, alors la règle correspondante ne sera pas générée.

### Interface homme-machine

Nous introduisons une échelle de degrés afin de représenter l'incertitude [14, 15, 16]. Nos degrés font partie d'une échelle de gradualité symbolique  $L_M$ , composée de  $M$  valeurs:  $L_M = \{\tau_\alpha, \alpha=0, \dots, M-1\}$  munie de la relation d'ordre totale.

Pour des raisons pratiques, nous posons  $M=7$ , et à chaque degré faisant partie de notre échelle, nous associons une valeur linguistique:

$\tau_0 \Leftrightarrow$  Impossible -  $\tau_4 \Leftrightarrow$  Assez possible  
 $\tau_1 \Leftrightarrow$  Très peu possible -  $\tau_3 \Leftrightarrow$  Possible -  $\tau_5 \Leftrightarrow$  Très possible  
 $\tau_2 \Leftrightarrow$  Peu possible -  $\tau_6 \Leftrightarrow$  Certain

A chaque degré de notre échelle de croyance (valeur linguistique) nous associons une valeur de l'échelle des degrés de probabilités symboliques définie par:

$$D_M = \{dg(\tau_\alpha) = \alpha, \alpha=0, \dots, M-1\}$$

où  $dg(\tau_\alpha)$  est le degré de probabilité symbolique.

A chaque mesure de probabilité  $P$ , définie sur l'ensemble  $P(X)$ , ensemble des parties de  $X$ , à valeurs dans

$[0, 1]$ , nous associons une valeur de l'échelle  $D_M$  des degrés de probabilités symboliques telle que:

Si $P=0$	alors	$dg(\tau_0)=0$
Si $P \in ]0, 0.5[$	alors	$dg(\tau_1)=1$
Si $P \in [0.5, 0.7[$	alors	$dg(\tau_2)=2$
Si $P \in [0.7, 0.8[$	alors	$dg(\tau_3)=3$
Si $P \in [0.8, 0.9[$	alors	$dg(\tau_4)=4$
Si $P \in [0.9, 0.95[$	alors	$dg(\tau_5)=5$
Si $P \in [0.95, 1]$	alors	$dg(\tau_6)=6$

La méthodologie de découpage de l'espace des probabilités relève directement de l'expert humain ; cette première expertise consiste à diviser l'espace probabiliste en sous intervalles numériques associés respectivement aux valeurs linguistiques allant de l'impossible au certain. Dans notre approche, l'expert humain formule ses énoncés relatifs aux faits sous forme d'expressions linguistiques pondérées par des mots plutôt que par des nombres, parce que l'usage des mots est un aspect très naturel. La deuxième expertise, plus computationnelle, est aussi nécessaire; elle assure la représentation des valeurs linguistiques par des degrés de probabilités symboliques, par exemple, nous croyons que l'événement dont la probabilité d'apparition est nulle est représenté par le mot impossible et est associé à un degré de probabilité symbolique « 0 ».

Dans ce qui suit, la méthode d'optimisation utilisée est celle qui consiste à éliminer les règles ayant un degré de croyance nul. Ces règles sont neutres et n'ont aucune influence pendant la phase d'inférence.

Considérons l'exemple suivant: Soit un ensemble d'apprentissage représenté par 4 attributs  $\{X_1, X_2, X_3, X_4\}$ . Pour un seuil de corrélation  $\theta=0.95$  par exemple, on obtient la matrice seuillée  $R_{seuil}$  suivante:

$$R_{seuil} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

En regroupant les composantes corrélées, on obtient les composantes connexes maximales suivantes:  $\{X_1\}$ ,  $\{X_2\}$  et  $\{X_3, X_4\}$ .

Supposons que la taille de subdivision est  $\delta=3$  et le nombre de classes  $C=3$ , alors le nombre de règles générées par les composantes connexes  $\{X_1\}$ ,  $\{X_2\}$  et  $\{X_3, X_4\}$  est égal à:

$$3 \cdot (3^1 + 3^1 + 3^2) = 45$$

Si une prémisses possible générée pour la composante  $\{X_1\}$  est telle que:  $X_1$  est dans  $rg_0$ , elle conduira alors aux 3 règles suivantes:

Règle 1: Si  $X_1$  est dans  $rg_0$  Alors  $y_1$  avec le degré  $\alpha_1$

Règle 2: Si  $X_1$  est dans  $rg_0$  Alors  $y_2$  avec le degré  $\alpha_2$

Règle 3: Si  $X_1$  est dans  $rg_0$  Alors  $y_3$  avec le degré  $\alpha_3$

où:

$$\alpha_1 = P(y_1 / X_1 \text{ est dans } rg_0)$$

$$\alpha_2 = P(y_2 / X_1 \text{ est dans } rg_0)$$

$$\alpha_3 = P(y_3 / X_1 \text{ est dans } rg_0)$$

Considérons une des règles de l'exemple précédent:

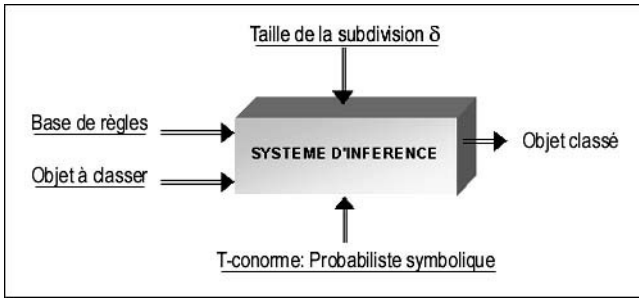
Règle 1: Si  $X_1$  est dans  $rg_0$  Alors  $y_1$  avec  $\alpha_1 \in ]0, 0.5[$

Après interfaçage homme-machine, la règle 1 prend la forme suivante:

Règle 1: Si  $X_1$  est dans  $rg\_0$  Alors  $y_1$  avec  $dg(\tau_1)=1$ .

**SYSTEME D'INFERENCE**

Le système d'inférence représenté par le schéma de la figure 2, consiste à classer de nouveaux objets n'appartenant pas à l'ensemble d'apprentissage. Le moteur d'inférence utilisé est du type classique. Le système d'inférence reçoit en entrée la base de règles fournie par le générateur, ainsi qu'un vecteur (de la base test non étiqueté par le nom de la classe à laquelle il appartient) décrivant l'objet à classer. Une fois que ce vecteur est discrétisé, le moteur d'inférence lui associe une classe d'appartenance avec un certain degré de croyance symbolique.



**Figure 2:** Schéma représentant le système d'inférence.

La méthode de raisonnement utilisée par le système d'inférence est exacte parce que les règles déclenchées sont celles dont les prémisses sont vérifiées exactement par le nouveau vecteur à classer. L'inférence exacte mise en œuvre par le moteur d'inférence est le Modus Ponens classique (l'inférence par déduction). Si l'on connaît la relation d'implication entre une prémisse et une conclusion ( $P \rightarrow C$ ) et si la prémisse  $P$  est vraie, alors le Modus Ponens permet de déduire la conclusion  $C$ .

**Traitement de l'incertitude**

Une fois que les règles sont déclenchées puis regroupées par classes, on associe à chaque classe un degré de croyance final qui est le résultat de la T-conorme probabiliste symbolique [14 -16]:

$$T(p, q) = \min(p+q, 6)$$

$p$  et  $q$  étant des degrés de probabilités symboliques.

Après avoir calculé tous les degrés de croyance finaux, la classe choisie qui sera associée au nouveau vecteur à classer est celle qui a un degré de probabilité symbolique maximal.

Dans le cas où le maximum des degrés de probabilités symboliques est atteint par plusieurs classes, la classe attribuée à l'objet à classer est celle qui a une fonction de pénalité maximale  $F_i$  définie par:

$$F_i = \max(\sum_{j=1}^p P_{i,j} - (M-1), 0)$$

$P_{i,j}$  désigne le degré de probabilité symbolique d'appartenance de l'attribut  $X_j$  à la classe  $y_i$ .

Considérons l'exemple suivant: soient le nombre d'attributs d'un ensemble d'apprentissage  $p = 4$ , le nombre de classes  $C = 3$ , le seuil de corrélation  $\theta = 0.95$  et les

domaines de variation  $R_1, R_2, R_3$  et  $R_4$  correspondant respectivement aux composantes  $X_1, X_2, X_3, X_4$ . Pour une taille de la subdivision  $\delta = 3$ , la discrétisation des attributs conduit à l'ensemble des sous-intervalles totalement ordonnés suivant:  $\{rg\_0, rg\_1, rg\_2\}$ .

Soit le vecteur à classer  $V = \{X_1, X_2, X_3, X_4\}$ . Sa discrétisation conduit au vecteur discrétisé  $V^{dis} = \{rg\_0, rg\_2, rg\_0, rg\_0\}$  et les règles déclenchées par le moteur d'inférence sont les suivantes:

Règle 2: Si  $X_1$  est dans  $rg\_0$  Alors  $y_1$  avec  $deg(\tau_4)=4$

Règle 1: Si  $X_1$  est dans  $rg\_0$  Alors  $y_2$  avec  $deg(\tau_2)=2$

Règle 17: Si  $X_2$  est dans  $rg\_2$  Alors  $y_1$  avec  $deg(\tau_1)=1$

Règle 15: Si  $X_2$  est dans  $rg\_2$  Alors  $y_3$  avec  $deg(\tau_5)=5$

Règle 20: Si  $X_3$  est dans  $rg\_0$  et Si  $X_4$  est dans  $rg\_0$ , Alors  $y_1$  avec  $deg(\tau_6) = 6$ .

Les règles déclenchées appartenant à la classe  $y_1, y_2, y_3$  conduisent aux ensembles suivants:

$A = \{\text{règle 2, règle 17, règle 20}\}$

$B = \{\text{règle 1}\}$

$C = \{\text{règle 15}\}$

En appliquant la conorme probabiliste symbolique  $T(p, q) = \min(p+q, 6)$  à l'ensemble  $A$ , on obtient:

$$T(4, 1) = 5$$

$$T(5, 6) = 6 \text{ avec } F_1 = 5$$

Les degrés de probabilités symboliques finaux ainsi que leurs fonctions de pénalité correspondantes sont les suivants:

$$T = 6 \text{ avec } F_1 = 5$$

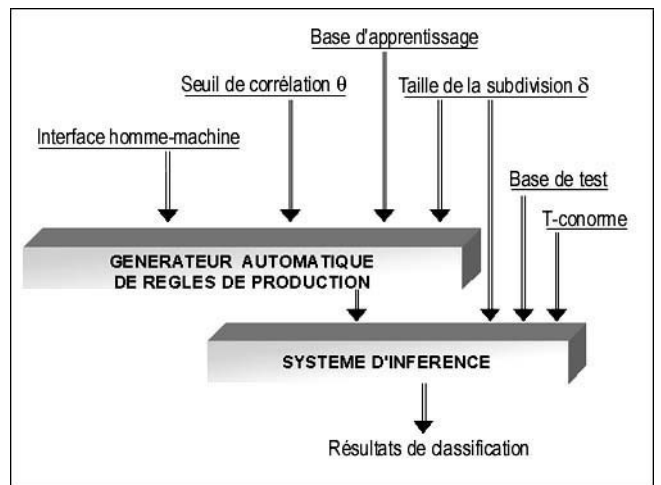
$$T = 2 \text{ avec } F_2 = 0$$

$$T = 5 \text{ avec } F_3 = 0$$

On peut donc conclure que le vecteur à classer appartient à la classe 1 avec un degré de probabilité symbolique égale à 6 et une fonction de pénalité égale à 5.

**IMPLEMENTATION ET RESULTATS**

La structure générale du système de classification symbolique est représentée par le schéma de la figure 3.



**Figure 3:** Schéma de l'architecture générale du système de classification symbolique.

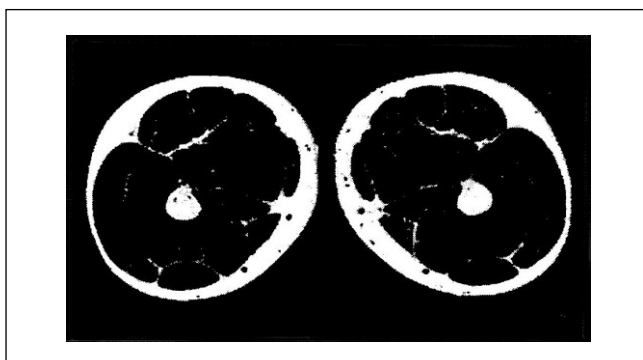
Pour valoriser notre approche symbolique, nous avons effectué des tests sur deux images: la cryosection de la

cuisse humaine et la texture. Nous rappelons à ce stade, que notre système de classification symbolique a été développé en langage C sous MS-DOS.

Pour valider la cohérence et la stabilité des résultats, nous utilisons la méthode de cross-validation stratifiée parce qu'elle permet de fournir une estimation non biaisée du taux de bonne classification, même dans le cas de petits échantillons d'apprentissage [20]. La méthode de cross-validation stratifiée consiste à diviser l'ensemble des exemples en  $n$  sous-ensembles égaux  $E_1, E_2, \dots, E_n$ , avec une répartition équitable des classes. A partir de  $E_1 \cup E_2 \cup \dots \cup E_{n-1}$ , le générateur de règles construit la base de règles qui est ensuite utilisée pour classer les éléments de  $E_n$ . L'application de ce même procédé par permutation circulaire, donne  $n$  pourcentages de classification. Le taux final de classification est la moyenne de ces  $n$  pourcentages.

### Image de cryosection humaine RVBXY

L'image que nous avons testé est la cryosection de la cuisse humaine provenant du site "National Library of Medicine". Les images numériques du corps humain disponibles sont celles d'un homme et d'une femme; elles sont acquises selon trois modes: l'imagerie tomographique, l'IRM (Imagerie par Résonance Magnétique) et la photographie couleur de cryosection. Le dernier mode consiste à immerger le corps dans un gel, le congeler puis le découper en couches de 1 mm pour l'homme et de 0.33 pour la femme qui sont arasées avant d'être photographiées. L'image observée sur la figure 4 a été acquise avec le dernier mode (photographie couleur de cryosection), puis filtrer et mise sous le format TIFF (Tag Image File Format). Cette image a une taille de 670\*415 pixels. Les attributs qui caractérisent un pixel sont les couleurs Rouge, Vert et Bleu qui constituent le modèle colorimétrique désigné par RVB. Les niveaux de couleurs varient entre 0 et 255.



**Figure 4:** Image d'une cryosection de cuisse humaine.

Une classification manuelle de l'image par un expert (médecin, anatomiste) a été faite. Quatre tissus ont été identifiés: graisse, os, moelle et muscle. Chacun de ses tissus correspond à une classe et chaque classe est représentée par un fichier de 300 pixels. L'échantillon obtenu est constitué de 1200 pixels, 300 pixels de chaque classe. Avec la méthode de cross-validation stratifiée d'ordre 4, quatre ensembles d'apprentissage sont obtenus dont chacun est constitué de 900 pixels; les ensembles de tests correspondant contiennent donc 300 pixels. Les tests

effectués ont montré que l'ajout des composantes X et Y, pour repérer la position géométrique d'un pixel et de tenir compte de son voisinage, améliore les résultats [13]. Les résultats obtenus avec l'image de la cryosection humaine RVBXY sont regroupés dans le tableau 1.

Seuil de corrélation	Cardinal de la subdivision	3	5	7
0.95	Nombre de règles générées	60	100	140
	Nombre de règles optimisé	33.75	44.75	53.75
	Temps du générateur (secondes)	1	2	3
	Taux de classification (%)	85.58	<b>99.08</b>	95.66
	Temps du classifieur (secondes)	0	0	0
0.90	Nombre de règles générées	87	255	567
	Nombre de règles optimisé	28.75	43.25	57.5
	Temps du générateur (secondes)	1.25	6.33	10.5
	Taux de classification (%)	85.74	<b>99.08</b>	90.99
	Temps du classifieur (secondes)	0	0	0
0.80	Nombre de règles générées	132	540	1428
	Nombre de règles optimisé	25.75	43	62
	Temps du générateur (secondes)	2	10	27
	Taux de classification (%)	85.99	<b>99.08</b>	92.74
	Temps du classifieur (secondes)	0	0	0
0.50	Nombre de règles générées	336	2520	9632
	Nombre de règles optimisé	21.75	40.5	58
	Temps du générateur (secondes)	6	48	183
	Taux de classification (%)	86.83	97.83	97.00
	Temps du classifieur (secondes)	0	0	0

**Tableau 1:** Résultats obtenus avec l'image de la cryosection humaine RVBXY.

En regardant attentivement les résultats du tableau 1, on remarque que les taux de bonne classification obtenus pour un cardinal de subdivision de 5 sont presque égaux à 100%. Le meilleur taux de classification est de 98.08%, tandis que le nombre de règles correspondant est de 43. Les temps de traitement (de génération et de classification) sont faibles, du moins lorsque le nombre de règles n'est pas trop élevé.

### Image des textures

La figure 5 montre une image constituée de deux microtextures différentes. Un prétraitement (le calcul des différentes corrélations locales) de l'image initiale a donné naissance à une série de 8 images; chacune de ces images est le résultat d'une détection d'un attribut particulier. Un pixel est alors décrit par un vecteur à 8 attributs [21].

Ainsi, l'image des deux textures de la figure 5 est représentée par deux classes dont chacune est décrite par 8 fichiers contenant chacun les valeurs des pixels pour l'une des 8 composantes. L'échantillon obtenu est constitué de 400 pixels de chaque classe. La méthode de cross-validation stratifiée d'ordre 4 conduit à 4 ensembles d'apprentissage contenant chacun 600 pixels et 4 ensembles de tests de 200 pixels chacun [1]. Les résultats obtenus avec l'image des textures sont regroupés dans le tableau 2.

Les résultats du tableau 2 révèlent que l'approche symbolique présente de très bons résultats, mais quelques fois avec une dépréciation qui varie de 0.25% à 11.8%. Le meilleur taux de bonne classification obtenu est de 97.25%; il correspond à un nombre de règles égale à 44.5.

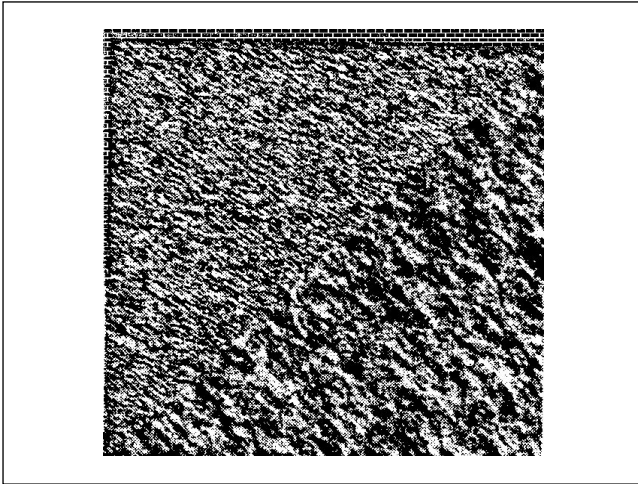


Figure 5: Image des textures.

Seuil de corrélation	Cardinal de la subdivision	3	5	7
0.95	Nombre de règles générées	48	80	112
	Nombre de règles optimisé	25	36.25	44.5
	Temps du générateur (secondes)	1	2	3
	Taux de classification (%)	87.37	94.62	97.25
	Temps du classifieur (secondes)	0	0	0
0.90	Nombre de règles générées	54	110	182
	Nombre de règles optimisé	23.75	35.5	44.75
	Temps du générateur (secondes)	1	2	4
	Taux de classification (%)	83.75	94.50	93.25
	Temps du classifieur (secondes)	0	0	0
0.80	Nombre de règles générées	91.5	427.5	1403.5
	Nombre de règles optimisé	22	34.5	47
	Temps du générateur (secondes)	1.75	8	24.75
	Taux de classification (%)	91.87	94.62	95.75
	Temps du classifieur (secondes)	0	0	0
0.50	Nombre de règles générées	426	5040	26474
	Nombre de règles optimisé	20.5	39	63
	Temps du générateur (secondes)	7.75	89.75	444.75
	Taux de classification (%)	91.37	92.37	92.37
	Temps du classifieur (secondes)	0	0	2

Tableau 2: Résultats obtenus avec l'image de textures.

### CONCLUSION

Les tests expérimentaux réalisés sur les deux images (cryosection de la cuisse humaine et textures), nous ont permis d'effectuer une première validation de notre approche symbolique. Les performances de classification obtenues par notre modèle symbolique sont très bonnes. On peut dire que notre modèle symbolique présente une approche qualitative de classification intéressante et très satisfaisante. Contrairement à l'approche numérique, il permet d'obtenir un résultat plus lisible et fournit des explications convaincantes. Notons, toutefois, que les résultats de classification dépendent du type d'image (de la base de données), le nombre de règles dépend du nombre d'attributs et le temps de traitement dépend non seulement du nombre de règles générées mais aussi de la vitesse du processeur.

### REFERENCES

- [1]- Borgi A. "Apprentissage supervisé par génération de règles: le système SUCRAGE" Thèse de l'Université de Paris 6, (1999).
- [2]- Ishibuchi H., Nozaki K., Tanaka H., "Distributed representation of fuzzy rules and its application to pattern classification", *Fuzzy Sets and Systems*, North-Holland, Vol. 52, (1992), pp. 21-32.
- [3]- Ishibuchi H., Nozaki K., Yamamoto N., Tanaka H., "Construction of fuzzy classification systems with rectangular fuzzy rules using genetic algorithms", *Fuzzy Sets and Systems*, North-Holland, Vol. 65, (1994), pp. 237-253.
- [4]- Ishibuchi H., Nozaki K., Yamamoto N., Tanaka H., "Selecting Fuzzy If-Then Rules for Classification Problems Using Genetic Algorithms", *IEEE Transactions on Fuzzy Systems*, Vol. 3, N°3, August (1995).
- [5]- Nozaki K., Ishibuchi H., Tanaka H., "Selecting Fuzzy Rules with Forgetting in Fuzzy Classification Systems", *Proceedings of the Third IEEE International Conference on Fuzzy Systems*, Orlando, Vol. 1, (1994), pp. 618-623.
- [6]- Ishibuchi H., Nozaki K., Tanaka H., "Efficient fuzzy partition of pattern space for classification problems", *Fuzzy Sets and Systems*, North-Holland, Vol. 59, (1993), pp. 295-304.
- [7]- Zighed D.A., Auray J.P., Dura G., "SIPINA: Méthode et logiciel", Editions Alexandre Lacassagne, Mathématiques appliquées, N°2, (1992).
- [8]- Quinlan J.R. "Learning efficient classification procedures and their application to chess and games", in: R.S. Michalski, J.G. Carbonell et T.M. Mitchell editors, *Machine Learning: An Artificial Intelligence Approach*. Morgan Kaufman Publishers, Vol.1, (1983), pp. 463-482.
- [9]- Quinlan J.R., "Improved Use of Continuous Attributes in C4.5", *Journal of Artificial Intelligence Research*, 4, (1996), pp. 77-90.
- [10]- Breiman L., Friedman J.H., Olshen R.A., Stone C.J., "Classification and regression trees", Chapman and Hall, (1984).
- [11]- Marsala C., "Apprentissage inductif en présence des données imprécises: construction et utilisation d'arbres de décision flous", Thèse de Doctorat de l'Université Paris 6, janvier (1998).
- [12]- Vernazza G., "Image Classification By Extended Certainty Factors", *Pattern Recognition*, Vol. 26, N° 11, Pergamon Press Ltd, (1993), pp.1683-1694.
- [13]- Borgi A., Bazin J. M., Akdag H., "Supervised Classification by Automatic Rules Generation", *The Fourth World Congress on Expert Systems, Application of Advanced Information Technologies*, Mexico City, (1998), pp. 561-568.
- [14]- Seridi H., Bannay-Dupin F., Akdag H., "Qualitative operators for dealing with uncertainty", *Proceeding of Fuzzy-Neuro Systems'98*, Computational Intelligence, Edited by C. Freksa, infix, (1998), pp. 202-209.
- [15]- Seridi H., Akdag H., "Le traitement de l'incertain par une approche qualitative", *Proceedings of IPMU'98*, Paris, (1998), Vol. 2, pp. 1684-1690.
- [16]- Seridi H., Akdag H., "A Qualitative Approach for Processing Uncertainty", in: *Uncertainty in Intelligent and Information Systems*, Edited by B. Bouchon-Meunier, R.R. Yager and L.A. Zadeh in: *Advances in Fuzzy Systems-Applications and Theory*, World Scientific, Vol. 20, (2000), pp. 46-57.
- [17]- Pearl J., "Numerical Uncertainty In Expert Systems", *Readings in Uncertainty Reasoning*, Edited by G. Shafer and J. Pearl, Morgan Kaufman publishers, Inc. California, (1990), pp. 255-258.

- [18]- Heckermans D., "Probabilistic Interpretations for MYCIN's Certainty Factors", *Uncertainty in Artificial Intelligence*, L.N. Kanal & J.F. Lemmer (Editors), North-Holland, (1986), pp. 167-196.
- [19]- Zadeh L., "Fuzzy Sets as a Basis for a Theory of Possibility", *Fuzzy Sets and Systems*, 1, North-Holland Publishing company, (1978), pp. 3-28.
- [20]- Kohavi R., "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection", *Proceedings of fourteenth International Joint Conference on Artificial Intelligence, Canada, Vol. 2*, (1995), pp. 1137-1145.
- [21]- Ould Ahmedou M.L., Bonnet N., "Etude comparative de différentes méthodes de classification non supervisée. Application à la segmentation d'images multivariées", *Proceedings of IPMU'98, Paris, Vol. 2*, (1998), pp. 1677-1683. □