

## **IDENTIFICATION DE PISTES RADAR PAR L'ANALYSE DES DONNEES SYMBOLIQUE-NUMERIQUE.**

**Jean-François GrandIn et Catherine Jacq  
THOMSON-CSF / Division RCM  
178, Boulevard Gabriel Péri  
92242 Malakoff Cedex**

Dans les systèmes de défense modernes, les avancées du domaine de l'électronique ont fait croître de plusieurs ordres de grandeur le volume et la complexité des informations manipulées. De plus, toutes les décisions doivent être prises très rapidement et leurs conséquences sont de plus en plus coûteuses, eu égard aux moyens humains, techniques et financiers mis en jeu.

L'accroissement considérable de la puissance des moyens informatiques dans tous les domaines (acquisition, communication, stockage, calcul) permet aujourd'hui d'utiliser en temps réel sur une machine de coût modeste des techniques d'interprétation et d'analyse autrefois dédiés aux super-calculateurs.

Ces techniques vont jouer un rôle fondamental dans tous les systèmes où l'analyse, la classification et l'identification d'objets divers sont des fonctions de première nécessité (C3I, NCTR, ELINT...).

Les plus prometteuses parmi ces techniques sont les méthodes d'Apprentissage Automatique Symbolique-Numérique. Elles sont applicables à une grande variété de problèmes : discrimination, classification, détection, optimisation, contrôle, fusion de données..., et sont particulièrement adaptées aux applications Défense, où les maîtres mots sont détection, identification, guidage, systèmes multi-capteurs.

## PARTIE I

### APPRENTISSAGE AUTOMATIQUE SYMBOLIQUE NUMÉRIQUE

#### Les techniques numériques

Les techniques d'analyse numérique de données sont utilisées pour analyser les données disponibles, puis réaliser des regroupements et produire des décisionnelles dans le cadre de l'identification.

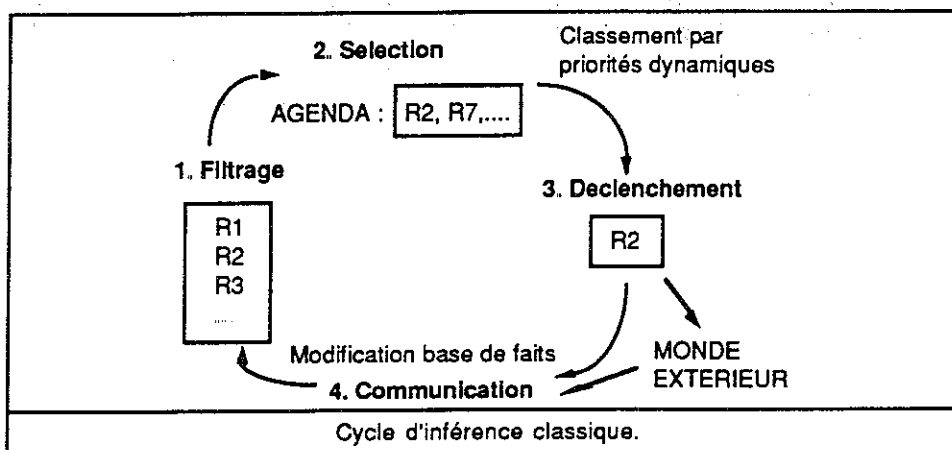
#### Les premiers systèmes experts

Les recherches en IA symbolique ont été focalisées dans un premier temps autour du raisonnement formel et plus précisément autour des systèmes experts et moteurs d'inférences. A travers les systèmes experts, l'Intelligence Artificielle cherche non pas à imiter l'intelligence humaine mais l'un de ses aspects : la démarche logique.

Le schéma suivant définit les mécanismes de base des premiers systèmes experts.

*P est une proposition logique si on sait lui associer une valeur VRAIE ou FAUSSE  
Disposant d'un ensemble de propositions des formules peuvent être employées pour créer de nouvelles propositions (non A, A ou B, A et B, etc. ...).  
Parmi ces formules deux d'entre elles sont des principes de résolution déduits de la définition de l'implication ( $A \Rightarrow B \Leftrightarrow (\text{non } A \text{ ou } B)$   
modus ponens - Déduction - (chaînage avant) ( $A \Rightarrow B$  et A) entraîne B  
modus tollens - Induction - (chaînage arrière) ( $A \Rightarrow B$  et non B) entraîne non A*

Ainsi disposant d'un ensemble de faits observés ou produits précédemment (A, B, non C ...) et de règles ( $A \Rightarrow B$  ...) fournies généralement par un Expert (d'où l'appellation système expert) d'autres faits sont produits lors du cycle d'inférence.



Dans ces systèmes on produit des faits, pas des règles. Les règles sont fournies lors de la conception du système expert par l'Expert du domaine. Très vite on s'est aperçu que le recueil des connaissances expertes était le véritable "goulot d'étranglement" de la démarche conceptuelle. Le propos

des premières méthodes d'apprentissage automatique a été de remédier à ce problème en produisant automatiquement non plus des faits mais de nouvelles règles.

## L'Apprentissage Automatique Symbolique

Les mécanismes logiques sont complétés par d'autres producteurs de règles et de faits plus complexes combinant à la fois des méthodes de généralisation, spécialisation avec les techniques de déduction et d'induction.

<u>Exemples de techniques de généralisation:</u>	
<ul style="list-style-type: none"> <li>• par suppression d'un terme :</li> </ul>	avant : $[fréquence = [1550,2250]] \wedge [Li = [0.8,1.5]]$ après : $[fréquence = [1550,2250]]$
<ul style="list-style-type: none"> <li>• par création d'une disjonction :</li> </ul>	avant : $[type = \{mono\}]$ après : $[type = \{mono\}] \vee [type = \{multi\}]$
<ul style="list-style-type: none"> <li>• par variabilisation d'une constante :</li> </ul>	avant : $[modulation = \{diversité\}]$ après : $[modulation = \{X\}]$
<ul style="list-style-type: none"> <li>• généralisation par connaissance experte :</li> </ul>	(ex : transitivité de Prioritaire) avant : $Prioritaire(mission1, mission5) \wedge Prioritaire(mission5, mission3) \wedge$ $Prioritaire(mission1, mission3)$ après : $Prioritaire(mission1, mission5) \wedge Prioritaire(mission5, mission3)$
<u>Exemples de techniques de spécialisation:</u>	
Toutes les règles de généralisation en sens inverse et les règles d'exceptions	
$\left. \begin{array}{l} A \Rightarrow B \\ A \wedge D \wedge \neg C \end{array} \right\} \dashrightarrow A \wedge \neg D \Rightarrow C$	
Avant :	$déte\text{c}ter(Radar\_X) \Rightarrow brouillage\_y$ $déte\text{c}ter(Radar\_X) \wedge Ami(Radar\_X) \wedge NON\ brouillage\_y$
Après :	$déte\text{c}ter(Radar\_X) \wedge NON\ Ami(Radar\_X) \Rightarrow brouillage\_y$
Techniques de généralisation et de spécialisation.	

Le schéma précédent explique comment on combine des règles pour produire d'autres règles. A partir de 1988, des travaux sont réalisés sur des systèmes regroupés sous le vocable "systèmes d'apprentissage automatique symbolique-numériques" intégrant intimement des connaissances symboliques et numériques et où le contrôle est effectué par des mécanismes logiques et statistiques.

En effet il est nécessaire d'adjoindre des critères statistiques qui vont permettre d'évaluer les règles produites par les mécanismes logiques pour sélectionner les plus pertinentes et les plus efficaces.

*Entropie (à minimiser)*

$$E = - \sum_i p_i \log(p_i)$$

*Signifiante:*

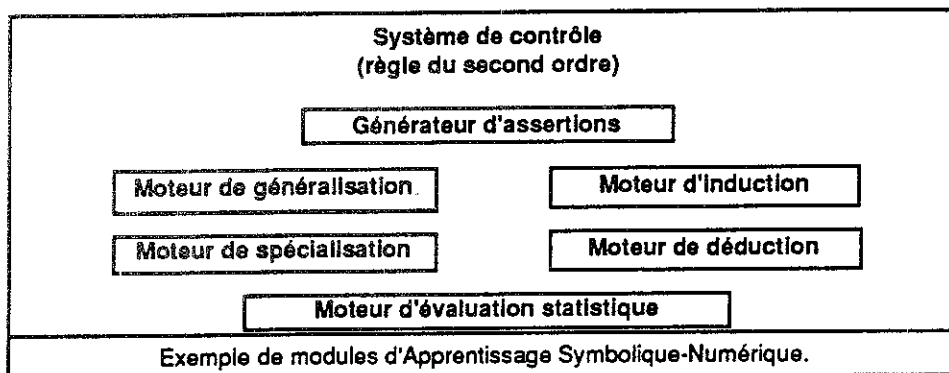
$$S = 2 \sum_{i=1}^n f_i \log(f_i/e_i) \rightarrow \chi^2(n-1)$$

*si "fi = ei"*

---

Exemples de critères statistiques.

La combinaison des différents moteurs d'inférence (moteur de généralisation, spécialisation, déduction, induction) conduisant à des moteurs plus complexes et l'apparition de critères d'évaluation statistiques des bases de faits et des bases de règles sont les éléments de base des techniques d'apprentissage Symbolique-Numérique.



### Intérêt de l'approche

L'Apprentissage Automatique Numérique-Symbolique intègre à la fois les techniques de prises en compte de connaissances expertes et les méthodes robustes de l'analyse numérique des données. Il mixe inférences statistiques et logiques et permet de manipuler des mesures, des règles expertes, des catalogues et des modèles.

La génération de connaissance intelligible réutilisable dans le procédé d'apprentissage permet d'envisager le développement dans les systèmes de Défense de composants intelligentes (capable de synthétiser et d'évaluer leur environnement puis de décider du bon mode de réaction).

Par exemple pour un avion pénétrant, il serait possible, à distance de sécurité, d'exciter judicieusement les radars adverses, d'apprendre leur modes de réponse, de comprendre leur enchaînement et d'utiliser au mieux cette information pour réaliser une séquence judicieuse de Contre-Mesures pendant la traversée de la zone adverse.

THOMSON-CSF a en cours plusieurs travaux pour développer et évaluer de tels systèmes. Les plus avancés sont :

- l'identification de signatures radar en contre-mesures
- le projet apprentissage automatique.

## PARTIE II

### APPLICATION À L'IDENTIFICATION DE RADARS EN CONTRE-MESURES

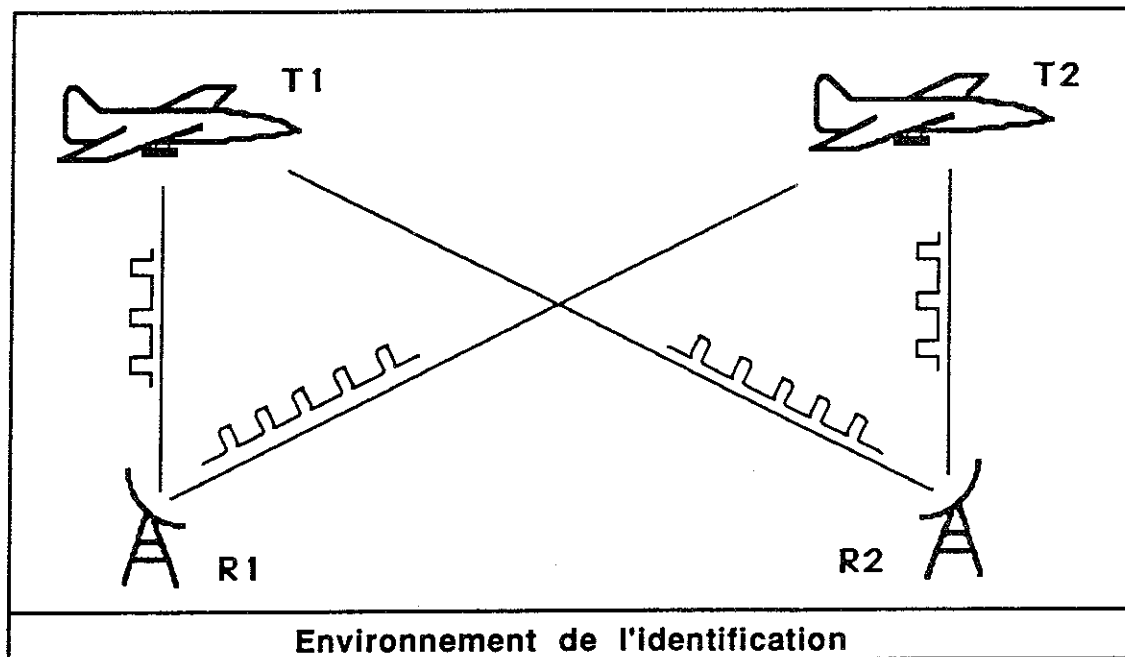
L'identification de radars s'appuie sur les formalismes établis par le professeur Edwin DIDAY. Les conseils et les propositions du professeur Edwin DIDAY nous ont permis dans le cadre de cette application, de prendre le recul nécessaire, de formaliser solidement l'application, de trouver de nombreuses solutions inovantes à nos problèmes et résultats appréciés dans le cadre industriel d'obtenir un algorithme numérique-symbolique efficace.

#### LE PROBLEME

Une action aérienne doit être impérativement soutenue par une manoeuvre électronique adaptée. Celle-ci repose sur une connaissance préalable précise et constamment maintenue à jour, de la situation électronique radar de l'adversaire. On parle de "Météo électronique"

Cette météo électronique est constituée par la représentation sur une carte des émetteurs localisés et identifiés. La localisation renseigne sur la position possible d'un émetteur radar, l'identification renseigne sur le type du radar et la fonction assurée par ce radar. Plusieurs centaines de radars peuvent être écoutés sur une mission aéroporté, ils devront être associés à l'une des plusieurs milliers de classes potentielles.

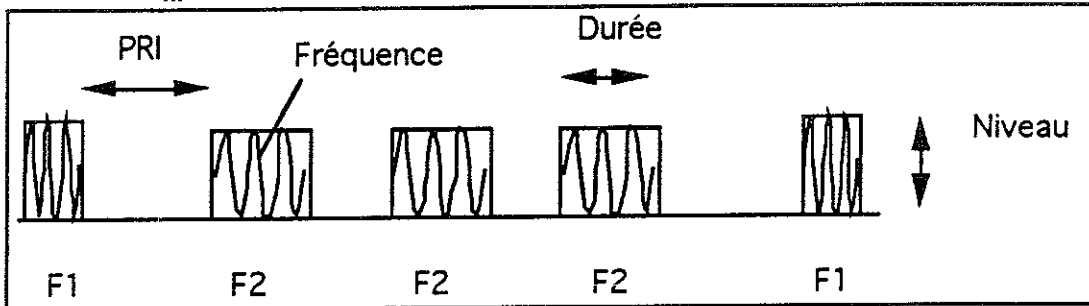
Dans le but de synthétiser la météo-électronique un capteur aéroporté enregistre les impulsions émises par les radars en émission à des instants (T1,T2) successifs d'acquisition. Les impulsions provenant de la même émission sont regroupées sous l'appellation de *piste*.



Une piste synthétise l'ensemble des impulsions recueillies pendant la mission, et caractérise un radar par sa localisation géographique et par un descriptif des motifs d'émissions.

Le motif d'émission est caractérisée par des ensembles de valeurs correspondant aux mesures de certains paramètres parmi lesquels:

- Fréquence(s) nominale(s) (ensemble de valeurs de fréquence)
- Période(s) de répétition des Impulsions (ensemble de valeurs)
- Largeur(s) d'impulsions (ensemble de valeurs)



A cette liste de paramètres mesurables, il faut ajouter un ensemble de paramètres qualitatifs caractérisant les synthèses d'informations.

Il est à noter que la principale difficulté réside dans le fait que les radars modernes utilisent des modes de fonctionnement complexes pouvant comporter une ou plusieurs fréquences, aux valeurs déterministes ou aléatoires et émises simultanément ou successivement. De même pour les Périodes de Répétition des Impulsions.

### LA FONCTION IDENTIFICATION CLASSIQUE

Les systèmes d'identification utilisent le plus souvent des algorithmes de comparaison entre les données mesurées et des fourchettes de paramètres stockées dans une bibliothèque. Ces systèmes sont sensibles aux fausses mesures et ne permettent pas l'identification des radars modernes à scénarios d'émission complexes, qui nécessitent des traitements spécifiques.

Ceci est du :

- au fait que certaines connaissances particulières des experts ne sont pas prises en compte comme, par exemple les incohérences générées par les fausses mesures.
- au fait que les bruits sont localisés dans une partie de l'espace de représentation (dépendent fortement de ce que l'on peut appeler le contexte)
- au fait que les métriques, les critères, et les paramètres à privilégier dépendent de la classe d'origine à identifier (localité des décisions et des métriques).

## SOLUTION SYSTÈME EXPERT

La solution opérationnelle aujourd'hui est le recours à un expert humain qui est le plus souvent capable d'identifier un radar à la seule vue d'un listing de paramètres. La connaissance des contraintes propres à la technologie ou à l'utilisation du radar lui permet souvent de lever les ambiguïtés. Malheureusement, de tels experts sont rares et souvent indisponibles pour diverses raisons (comment embarquer un expert dans un avion monoplace !!!).

L'observation et les interviews d'un expert ont permis de cerner les différentes étapes implicitement ou explicitement parcourues pour mener à bien l'identification d'un signal radar. De nombreuses connaissances sont manipulées, parfois de natures très différentes et nécessitant des représentations fondamentalement différentes (règles - algorithmes - base de données).

## UTILISATION DES STATISTIQUES

Les statistiques conventionnelles sont utilisées soit pour analyser localement les données mesurées et produire des décisionnelles locales, soit pour valider statistiquement les connaissances et les décisionnelles élaborées par l'apprentissage numérique-symbolique.

## L'APPRENTISSAGE AUTOMATIQUE NUMÉRIQUE-SYMBOLIQUE

L'introduction dans le système de l'apprentissage numérique-symbolique a pour objectifs :

- de prendre en compte à la fois la connaissance experte sur les pistes (sous forme de hiérarchies conceptuelles, modélisation du bruit...), la connaissance experte sur la discrimination de pistes (sous forme de règles d'identification pré définies, règles de discrimination entre certains radars ...), les bases de données empiriques (sous forme de bases de pistes identifiées) et des bases d'objets (les bases de signatures radar).
- de produire des décisionnelles locales à partir de l'examen des régularités statistiques présentées par les données, et des connaissances à disposition (signatures, localité des bruits, règles expertes.....)
- d'être capable de détecter l'apparition d'une nouvelle signature non-répertoriée dans le but de s'adapter à une évolution du terrain opérationnel.

## PARTIE III

### SISYPHE

Sisyphé est un algorithme d'apprentissage automatique numérique-symbolique. Il se situe donc au confluent de l'apprentissage automatique et de l'analyse des données.

L'apprentissage automatique est une technique issue de l'intelligence artificielle. Elle est apparue à la suite du succès obtenu par les systèmes experts d'identification.

La nécessité de fournir une base de règles expertes en entrée du système, la non disponibilité des experts dans certains domaines, le caractère évolutif de la connaissance furent à l'origine de la création d'algorithmes de génération automatique de base de règles.

Les pionniers en la matière sont AQ de Michalski et ID3 de Quinlan. Ces deux algorithmes représentaient dès lors les deux tendances fortes de l'apprentissage automatique: les algorithmes ascendants et les algorithmes descendants.

Un algorithme descendant est un algorithme qui part de tout l'espace des descripteurs et qui aboutit à l'identification de zones particulières de ce même espace.

A l'inverse un algorithme ascendant part de zones spécifiques de l'espace et agrandit ces zones le plus possible pour les étiqueter.

Sisyphé est un algorithme mixte, descendant puis ascendant. L'idée sous-jacente est la prise en compte du caractère progressif de la discrimination. Une discrimination tout d'abord descendante permet de scinder le problème de discrimination en sous problèmes plus spécifiques (démarche usuelle des experts), la discrimination ascendante est réservée à la résolution plus fine de ces problèmes locaux.

Cette approche se justifie également par la complexité souvent exponentielle des algorithmes ascendants. La restriction du nombre d'individus, du nombre de classes ainsi que de la taille de l'espace des généralisés permet d'éviter les explosions combinatoires.

L'analyse des données traite usuellement des tableaux individus-variables, une voie nouvelle est apparue avec l'analyse des données numérique-symbolique. Dans ce cadre l'analyse des données est appliquée à des objets structurés et probabilisés (ou possibilisés ou crédibilisés), de nouveaux algorithmes sont conçus pour pouvoir traiter ces objets et en extraire de l'information.

Sisyphé est l'un de ces algorithmes, il hérite de l'analyse des données la prise en compte de critères statistiques tels que le risque de Bayes, l'entropie, l'inertie (inter et intra), le  $\chi^2$ , le critère de Kolmogorov-Smirnov ....

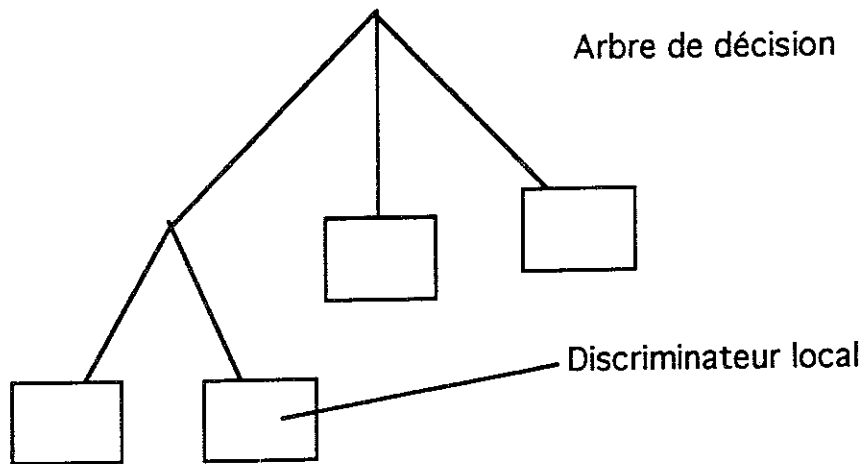
L'utilisation de ces indices a pour but de garantir une certaine robustesse à l'algorithme.

La prise en compte de probabilités rattachées aux objets, de coûts d'erreur d'affectation et d'estimations locales de densités est également un plus par rapport aux techniques de l'intelligence artificielle.



## Structure de Sysiphe

La structure du discriminateur généré est donc la suivante :



A cette structure centrale sont associées d'autres structures utilisées pour générer l'arbre et les discriminateurs locaux (génération et validation de règles locales),

en particulier à chaque nœud :

- des graphes de Conway traduisant des hiérarchies sur les classes et sur les variables, (présentés dans la suite)

- une classification pyramidale qui minimise la perte de l'information discriminante sur les assertions présentes au nœud, (présenté dans l'article Discrimination d'objets symboliques par segmentation et apprentissage local. Les troisièmes journées " Symboliques-Numériques" Paris 14-15 Mai 1992.

et à chaque nœud terminal :

- des moteurs d'apprentissage

- un moteur de validation statistique de faits.

### Avantages de Sysiphe :

On retiendra de cette courte introduction les avantages suivants pour SISYPHE :

- localité des décisionnelles (adaptation au contexte).
- maîtrise de l'explosion combinatoire.
- robustesse et qualification des décisions.
- capacité d'apprentissage.

## DESCRIPTIF

Sisyphé est un générateur automatique de discriminateurs d'individus capable de prendre en compte l'apparition de nouvelles classes, les données manquantes lors des phases d'apprentissage et de généralisation ainsi que la connaissance experte caractérisant le domaine et la discrimination des classes dans ce domaine.

Sisyphé prend en entrée :

- Une base d'individus identifiés (les pistes radars)
- Une base d'objets théoriques (les signatures radars)
- La connaissance du domaine exprimé sous la forme de contraintes structurés dans un graphe de Conway
- La connaissance experte exprimé sous formes de règles
- Des modélisations de bruit sur certains paramètres et des graphes de connaissance (hiérarchie conceptuelle).

Sisyphé fournit en sortie :

- Une base de règles discriminantes structurées sous la forme d'un arbre.
- De nouveaux objets théoriques lorsque cela s'avère nécessaire.

## LE FORMALISME

Nous présentons ci-dessous les trois principaux types d'objets utilisés par l'algorithme. Ce formalisme est basé sur celui du professeur Diday.

### **Un événement élémentaire (e) :**

- Son intention est définie sur une seule variable par un ensemble de modalités si la variable est qualitative ou un intervalle de valeurs si la variable est quantitative.

exemple :  $e = [\text{couleur} = \{\text{blanc}, \text{bleu}\}]$

- Son extension dans un cas simple est obtenue par l'application de la fonction indicatrice de son intention sur l'ensemble des individus.

exemple :

Soient :

$A = \{w, x, y, z\}$  un ensemble d'individus

et  $G = \{(w, \text{vert}), (x, \text{blanc}), (y, \text{jaune}), (z, \text{bleu})\}$

le graphe de la relation couleur sur A.

L'extension de e est :  $\{x, z\}$

Remarque : La fonction d'extension peut être différente de la fonction caractéristique de l'intention de l'événement élémentaire, ceci à fin de prendre en compte des notions de proximité entre les modalités d'une même variable. La fonction extension est alors la fonction composée d'une fonction de généralisation de l'intention et de la fonction caractéristique.

Par exemple :

Soit  $M(\text{couleur1}, \text{couleur2})$  la matrice de dissimilarité entre les modalités de la variable couleur, définie comme suit :

couleur1 couleur2	vert	blanc	jaune	bleu
vert	0	1	1	1
blanc		0	0,5	1
jaune			0	1
bleu				0

La fonction extension  $f_e$  peut alors être la fonction caractéristique de l'ensemble des modalités dont la dissimilarité avec l'intention de l'événement élémentaire est inférieure ou égale à une constante  $k$  donné.

$$a \longrightarrow \begin{cases} 1 & \text{si } \min\{M(\text{couleur}(a), c) / \forall c \in \{\text{blanc}, \text{bleu}\}\} \leq k \\ 0 & \text{sinon} \end{cases}$$

où  $a$  est un individu.

L'extension de  $e$  (pour  $k = 0,5$ ) est alors  $\{x, y, z\}$

Ce cas peut être vu comme la généralisation, dans l'espace des variables, de l'événement élémentaire  $e$  à  $[\text{couleur} = \{\text{blanc}, \text{bleu}, \text{jaune}\}]$ .

Le choix de la constante de généralisation peut alors être fait en fonction de la nature des variables discriminantes.

Dans le cas d'une variable quantitative dont le bruit peut être modélisé par une loi normale, la fonction extension devient alors la fonction caractéristique de l'intervalle agrandi à plus ou moins  $k$  écarts-types. Cette généralisation est indépendante de l'ensemble des individus.

En utilisant ces événements élémentaires on construit des objets symboliques plus compliqués, les assertions :

### Une assertion (a) :

- Son intention est définie par une conjonction d'événements élémentaires.

exemple :  $a = [\text{couleur} = \{\text{blanc}\}] \wedge [\text{poids} = [25, 50]]$

- Son extension est l'intersection de tous les événements élémentaires qui la composent, elle est donnée par la fonction produit des fonctions extensions.

Une fois définie la notion d'assertion l'on peut définir l'objet symbolique règle :

### Une règle (r) :

- Son intention est un couple ordonné d'assertions .

exemple :  $[\text{couleur} = \text{orange}] \text{ } 3 \text{ } [\text{forme} = \text{sphérique}] \Rightarrow [\text{nom} = \text{mandarine}]$

$r = ( \text{assertion 1} ; \text{assertion 2} )$

- Une règle admet deux fonctions d'extension.

La fonction d'extension positive qui est le minimum des deux fonctions d'extensions des assertions formant la règle (l'extension positive est l'ensemble des exemples de la règle).

La fonction d'extension négative est le minimum de la fonction extension de la première assertion et de la négation de la fonction extension de la deuxième assertion (l'extension négative est l'ensemble des contre-exemples de la règle).

exemple :

individus	couleur	forme	taille	nom
I 1	orange	sphérique	petit	mandarine
I 2	jaune	oblongue	grand	banane
I 3	jaune	sphérique	grand	pamplemousse
I 4	jaune	sphérique	petit	citron
I 5	orange	sphérique	grand	orange
I 6	orange	sphérique	petit	mandarine

L'extension positive de r est donnée par:

$$(\min([\text{couleur} = \text{orange}] \text{ } 3 \text{ } [\text{forme} = \text{sphérique}], [\text{nom} = \text{mandarine}]))^{-1}(1) = \{ I 1, I 6 \}$$

L'extension négative de r est donnée par:

$$(\min([\text{couleur} = \text{orange}] \text{ } 3 \text{ } [\text{forme} = \text{sphérique}], \text{non}([\text{nom} = \text{mandarine}])))^{-1}(1) = \{ I 5 \}$$

NB: on appelle extension d'une règle l'union de son extension positive et de son extension négative.

Remarque : Les assertions et les règles étant construites à partir des événements élémentaires intègrent de la même façon les dissimilarités entre modalités ou les fonctions modélisatrices du bruit.

### L'ALGORITHME :

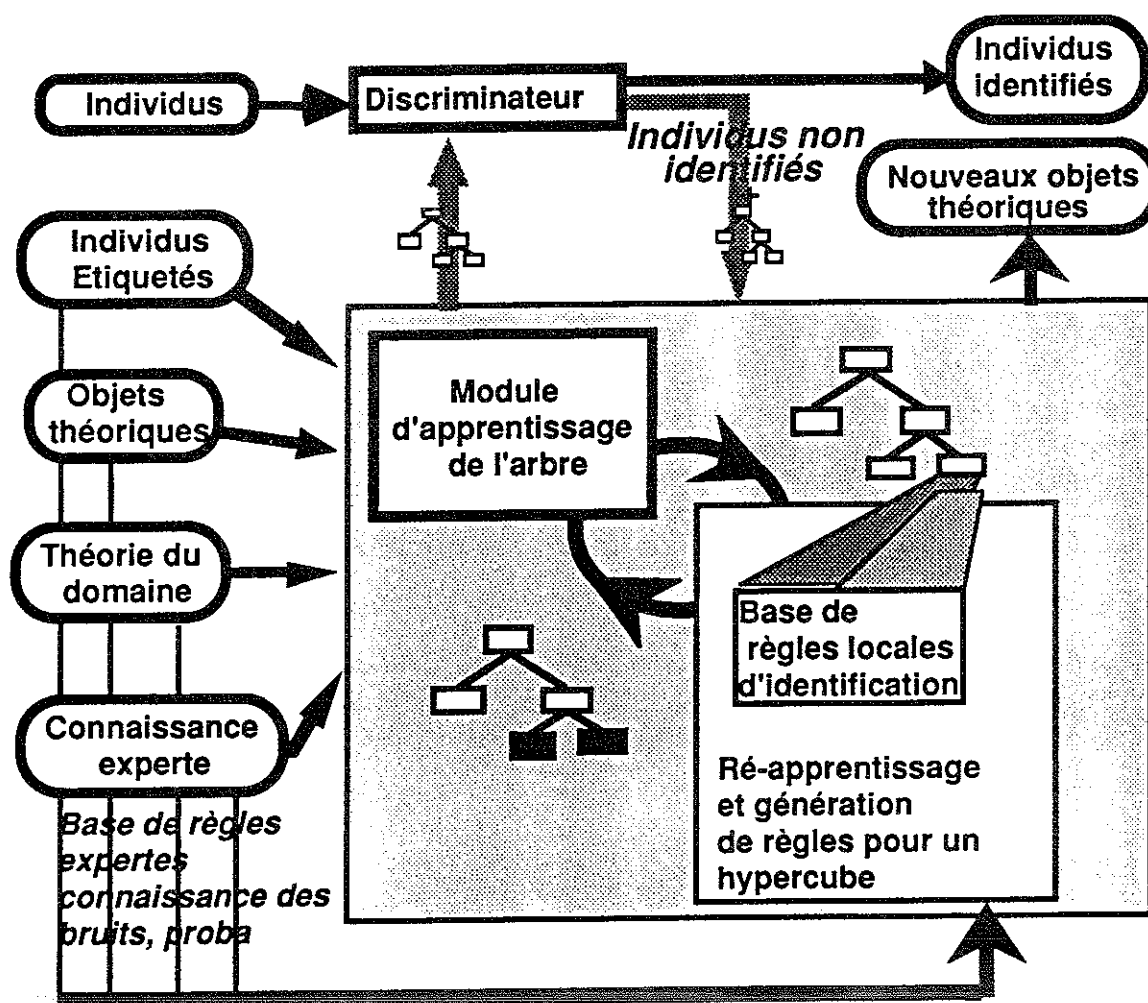
L'algorithme est à la fois descendant et ascendant.

Structure générale :

Sysiphe s'appuie sur deux idées principales :

- La segmentation qui produit une découpe récursive de l'espace des caractéristiques, ou autrement dit qui découpe le problème en sous-problèmes élémentaires.
- L'apprentissage local qui vise à raffiner la solution locale en prenant en compte un certain nombre de connaissances expertes qui vont être mises à profit localement.

Le schéma suivant montre les différentes phases de l'algorithme :



## LA SEGMENTATION :

On construit à partir de la base d'assertions et de la base d'apprentissage des individus, un arbre de segmentation d'assertions.

A chaque noeud de l'arbre on liste toutes les assertions qui intersectent avec le noeud.

Pour chacune des variables présentes dans l'une au moins des assertions attachées au noeud on génère une liste structurée d'événements élémentaires.

- Pour les variables quantitatives on génère une hiérarchie ou une pyramide d'intervalles (dans le cas où l'on autorise les recouvrements);
- Pour les variables qualitatives on prend en compte les hiérarchies conceptuelles s'il y en a.

Chacune de ces structures est indiquée .

Dans le cas quantitatif , l'indice de dissimilarité est une fonction de la différence entre le nombre d'assertions discriminées par l'union de deux intervalles et le nombre d'assertions discriminées par les deux intervalles si les deux intervalles ne sont pas fusionnés ainsi que le nombre total d'assertions présentes sur les deux intervalles .

Dans le cas qualitatif l'indice est celui de la hiérarchie conceptuelle des modalités de la variable.

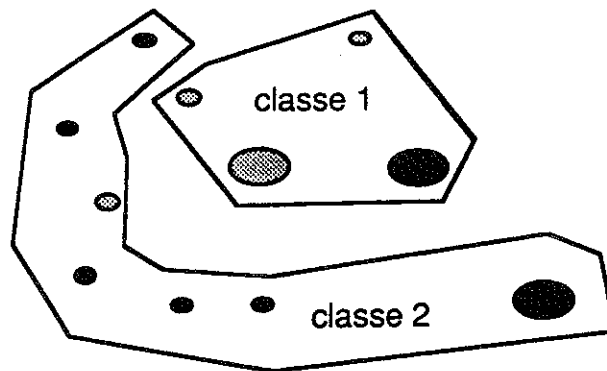
En cas d'égalité on prend en compte l'ordre de la droite réelle, ou un ordre sur les modalités si la variable est qualitative ordonnée.

Chaque valeur de l'indice induit une découpe de la variable (un partitionnement s'il s'agit d'une hiérarchie, un recouvrement s'il s'agit d'une pyramide).

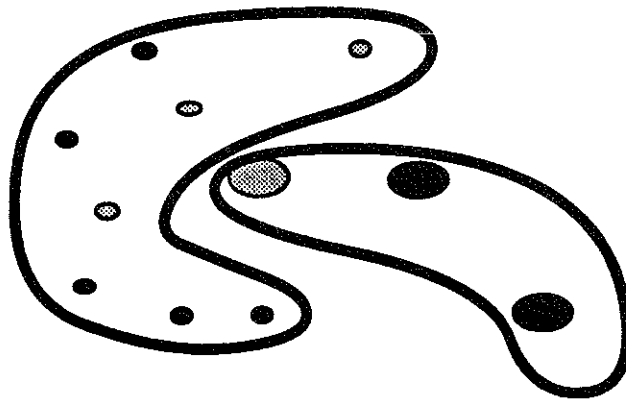
On optimise alors le critère dérivé du  $\chi^2$  (critère du  $\chi^2$  pour une hiérarchie) sur la base d'apprentissage des individus.

Le critère du  $\chi^2$  traduit la distance entre une partition cible d'un ensemble d'individus (induite par la variable à expliquer) et une partition induite par une variable explicative ,ici un niveau de la hiérarchie. Le but étant de minimiser cette distance.

exemple partition induite par la variable à expliquer :

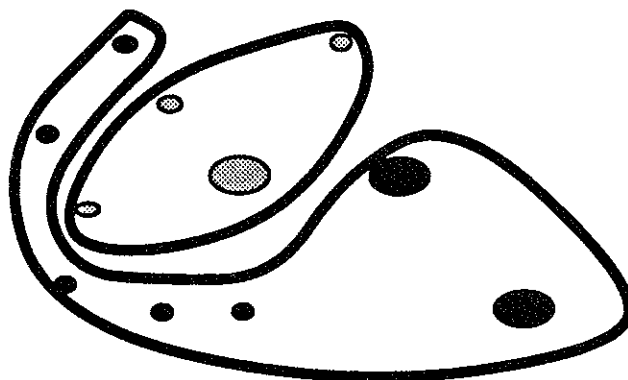


● : un individu



● : un individu

partition induite par la variable : taille

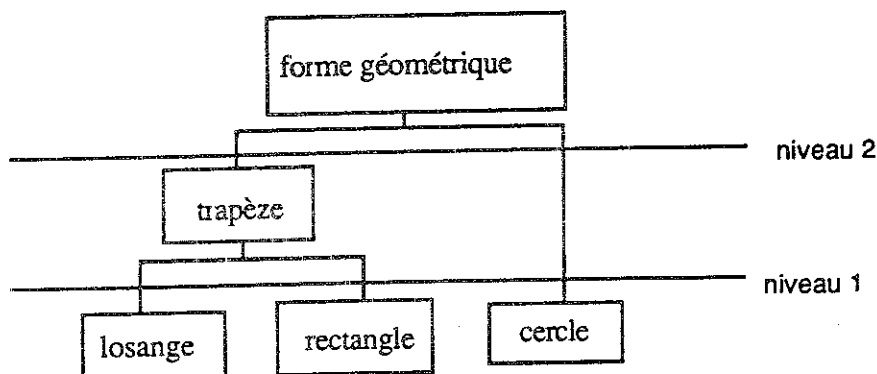


● : un individu

partition induite par la variable : couleur

Ici on voit que la partition induite par la variable couleur est plus proche de la partition à expliquer que la partition induite par la variable taille, c'est cette notion de proximité que caractérise le  $\chi^2$ , ie caractérisé la classe 2 comme étant les individus noir et la classe un comme étant les individus blancs est moins faut que de caractérisé la classe 1 comme étant les individus de grandes tailles et la classe deux comme étant les individus de petites tailles. (on a minimisé l'erreur).

Le calcul de ce critère est effectué pour chaque variable à expliquer sur les partitions induites soit pour un niveau d'une hiérarchie conceptuelle sur la variable soit par un niveau de la pyramide des intervalles.



ici par exemple la partition induite par le niveau 1 est (losange,rectangle,cercle), celle induite par le niveau 2 étant : (trapèze, cercle)

En pratique on n'évalue le critère que pour les indices correspondant à des paliers, pour lesquels le cardinal du codage induit est inférieur à un seuil et pour lesquels l'espérance du  $\chi^2$  sur les paliers inférieurs est plus petit ou égal au  $\chi^2$  du palier .

• Le calcul local du codage des paramètres quantitatifs est fait par une sélection de niveau dans classification pyramidale des assertions . L'indice de dissimilarité calculé traduit la perte d'information discriminante induit par le regroupement de deux assertions.

Un algorithme numérique classique prend en compte un pré-codage pour les variables quantitatives. SISYPHE travaillant à la fois à partir d'une bibliothèque de signatures et d'une base d'individus , le codage des paramètres quantitatifs est recalculé à chacun des noeuds de l'arbre en fonction des classes restant à discriminer à ce noeud. Cela permet l'utilisation dans l'algorithme de métriques locales calculées de façon à minimiser la perte d'information induite par un codage sur un sous-ensemble de classes restant à séparer.

L'élagage de l'arbre est effectué a posteriori.

Il a pour but de minimiser le risque de Bayes sur la base d'entraînement. Certaines assertions sont d'extension nulle sur cette base, le risque minimisé est donc :  $R_{\alpha,\beta}(A) = R_{\alpha}(A) + \beta (A^{\circ})$  où  $A^{\circ}$  est le nombre d'assertions dont l'extension est vide dans les noeuds terminaux de l'arbre [3].



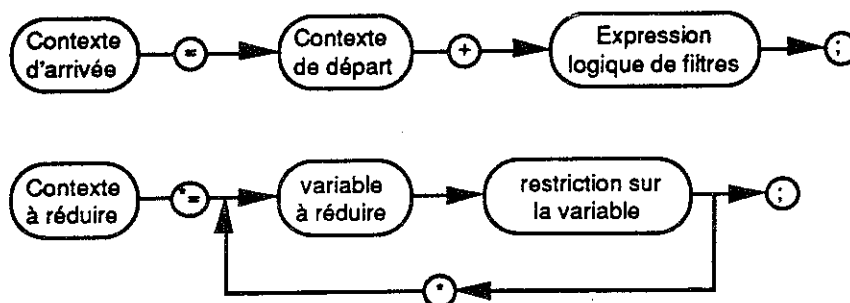


## Le filtre du domaine :

Le filtre du domaine est un graphe de Conway utilisé lors de la construction de l'arbre pour restreindre la liste des variables à considérer pour un noeud donné.

Il traduit en fait les relations entre les variables ou entre des valeurs particulières de ces variables.

Ces relations sont des relations globales d'impossibilité (variables mère-filles ou mères-filles lorsque les contextes sont obtenus par la conjonction de propriétés portant sur des variables différentes).



Le compilateur comprend deux types d'instructions :

- Description d'un lien entre deux contextes :  $C1 = C0 + \text{expression}$ ;  
 $C0$  = contexte de départ,  
 $C1$  = contexte d'arrivée,  
expression = expression logique de conditions à vérifier pour passer du contexte  $C0$  au contexte  $C1$ . L'expression peut contenir les opérateurs logiques **and** et **or** ainsi que des regroupements par parenthèses.
- Réduction de l'espace local d'un lien :  $C * = \text{restriction} * \dots$ ;  
 $C$  = contexte dont il faut réduire l'espace local,  
restriction = restriction d'une variable à une partie de son domaine de définition.

## Exemples de lien mère-fille dans le domaine RADAR

La prise en compte d'informations sur l'espace de paramètres permet par exemple d'utiliser, dans certaines zones de l'espace des paramètres, une variable non licite en dehors de cette zone.

La Période de répétition des impulsions n'a pas de sens lorsque le type d'émission est CW.

## Le filtre des classes :

Le filtre des classes est également un graphe de Conway. Chaque contexte correspond à une règle experte de discrimination entre des classes particulières.

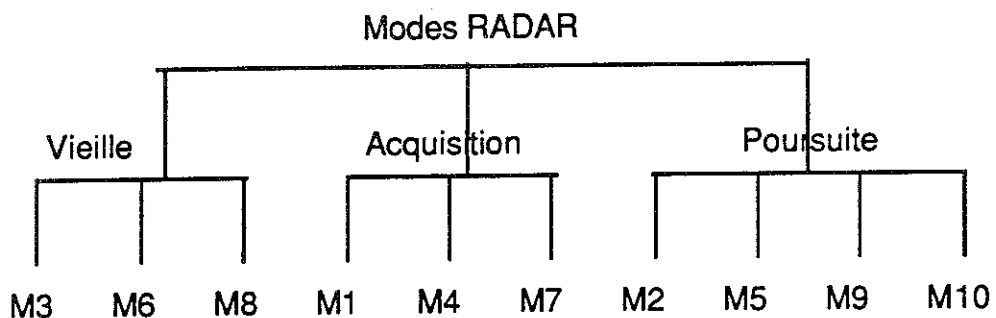
Ce filtre est utilisé pour enrichir, par la connaissance de l'expert, les bases de règles attachées aux noeuds terminaux de l'arbre. On peut aussi traduire des hiérarchies de classes, modélisant des priorités dans l'ordre de discrimination des classes.

Exemple de prise en compte de hiérarchies sur les classes :

La prise en compte d'une hiérarchie sur les classes permet d'utiliser une classification des modes radars en fonction des types de contre-mesures associés. Par exemple on peut regrouper les modes radars de veilles, les modes radars d'acquisition et ceux de poursuites. L'identification des modes radars se fait alors en respectant la structure hiérarchique des types de modes.

exemple :

Soit la classification simplifiée des dix modes RADAR suivants :



L'identification du mode radar M4 se fera par une pré identification d'un mode RADAR d'acquisition puis une spécialisation de cette identification en l'identification du mode RADAR M4.

Avantage de la méthode :

La prise en compte des coûts d'erreurs d'affectation au niveau de la fonction radar identifiée et non du mode. Ceci permet une minimisation du risque d'erreur calculé en fonction du temps octroyé pour fournir la décisionnelle. Autrement dit si l'on ne dispose pas d'assez de temps pour effectuer l'identification exacte du mode RADAR, on dispose déjà de l'information concernant sa fonction et on peut donc adapter les contre-mesures en conséquence.

Exemples de filtre :

$C2^* = q0 \text{ in } [K1, K2];$

$C1 = C0 + f1;$

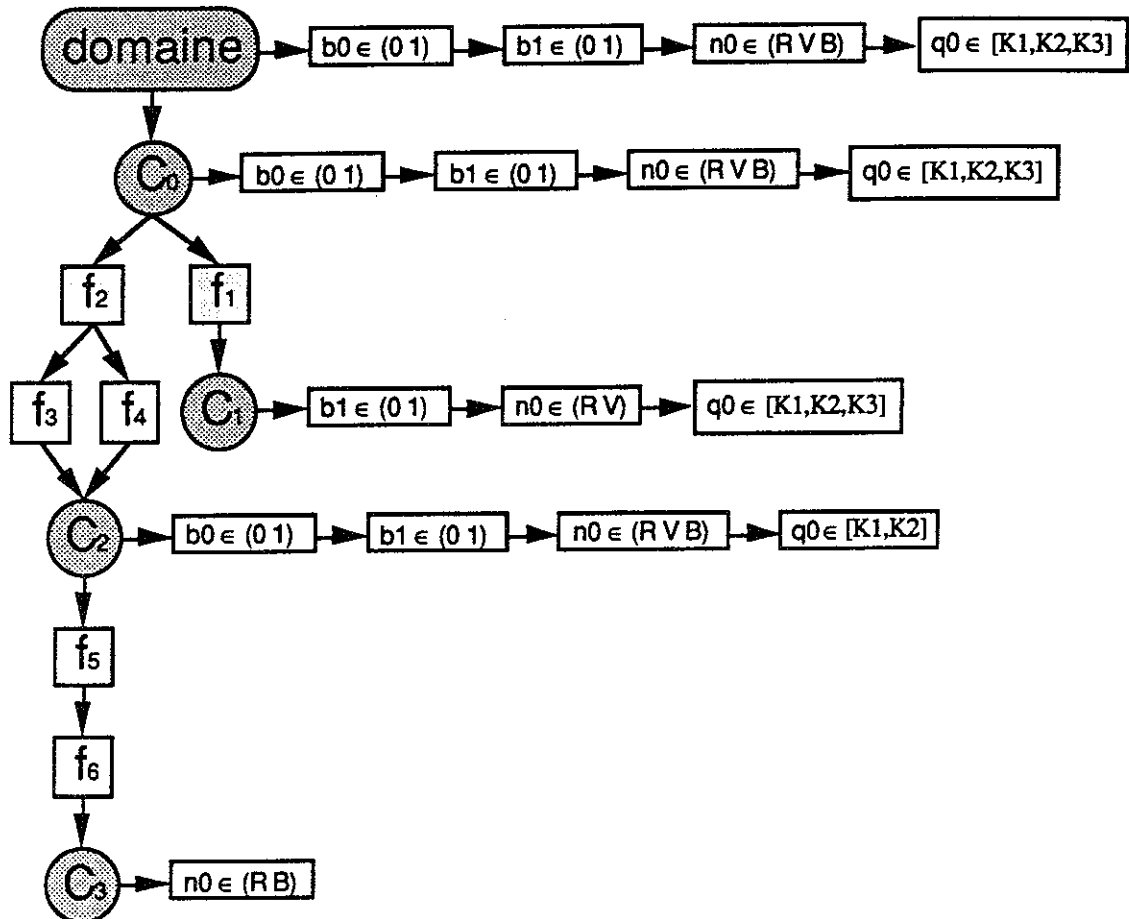
$C2 = C0 + f2 \text{ and } (f3 \text{ or } f4);$

$C3 = C2 + f5 \text{ and } f6;$

$C1^* = b0 \text{ in } () * n0 \text{ in } (R V);$

$C2^* = q0 \text{ in } [K1, K2];$

$C3^* = b0 \text{ in } () * b1 \text{ in } () * n0 \text{ in } (R B) * q0 \text{ in } [];$



A la sortie du module d'apprentissage de l'arbre on considère que l'information uniquement contenue dans les individus de la base d'apprentissage a été prise en compte.

Afin d'intégrer également toute celle comprise dans les intentions des assertions, on utilise des techniques d'apprentissage sur chacun des noeuds terminaux de l'arbre.

## L'APPRENTISSAGE LOCAL :

L'apprentissage local s'effectue séparément sur chaque noeud terminal de l'arbre fourni par le module d'apprentissage de l'arbre.

Un noeud terminal de l'arbre est caractérisé par :

- Une conjonction d'événements élémentaires appelée étiquette du noeud, elle correspond à l'assertion caractérisant la branche de l'arbre aboutissant au noeud.
- Un ensemble d'assertions obtenues par restriction des assertions intersectant avec le noeud (intersection entre ces assertions et l'étiquette du noeud).
- Un ensemble de règles expertes obtenues en utilisant le filtre des classes.
- Un ensemble d'individus appartenant à la base d'entraînement (Ils ont été affectés à ce noeud de l'arbre car ils appartiennent à l'extension de son étiquette).

Le module d'apprentissage est constitué d'un système de contrôle qui commande six opérateurs, chacun d'entre eux étant affecté à une tâche particulière.

Une base de règles d'ordre deux est utilisée pour évaluer la situation du noeud, l'ordre d'utilisation des opérateurs d'apprentissage est donnée en fonction de chacune des situations possibles.

Les taux de recouvrement des exemples et des contre-exemples des règles sont évalués sur la base d'individus et les techniques d'apprentissage sont utilisées par binôme de règles.

On laisse inchangées les règles dont le cardinal de l'extension sur la base des individus est inférieure au seuil de véracité. On considère alors tous les binômes de règles de conclusions identiques que l'on tente de généraliser. L'ordre utilisé est celui induit par le critère de véracité.

On passe ensuite au cas des règles de conclusions différentes. On commence par la règle dont le seuil de véracité est le plus faible, on considère les règles dont la conclusion correspond à la classe majoritaire sur l'extension négative de la règle.

On applique alors le moteur de spécialisation. Le critère de véracité ne peut donc qu'augmenter. Le critère de signifiante lui décroît. Si le critère de signifiante des règles générées est inférieur au seuil la spécialisation est rejetée, sinon la spécialisation est acceptée.

L'opérateur de généralisation d'assertions permet d'obtenir le généralisé de deux assertions par union des intentions.

La remontée dans une hiérarchie conceptuelle ou la suppression d'un événement élémentaire sont effectuées en commençant par la généralisation de l'événement élémentaire qui entraîne un accroissement maximum du cardinal de l'extension de l'assertion, puis si elle est acceptée par l'opérateur statistique on relance l'opérateur de généralisation sur la nouvelle assertion.

L'opérateur de spécialisation d'assertions génère à partir de deux assertions leur intersection et leurs complémentaires.

Les deux opérateurs d'induction et de déduction sont utilisés pour générer de nouvelles règles à partir des règles expertes.

L'opérateur d'évaluation statistique permet d'évaluer une liste de critères statistique :

- la véracité : fonction du cardinal de l'extension d'une assertion par rapport au cardinal du noeud.
- la complétude [1] (l'opérateur de complétude étant la fermeture du graphe des assertions)
- la simplicité [1] (la simplicité caractérise comme son nom l'indique la simplicité d'une description par rapport aux individus qu'elle recouvre)
- l'entropie [4] (l'entropie mesure la dispersion des classes)
  - la signifiante [4] la signifiante mesure l'écart entre une distribution aléatoire des classes sur un hypercube et la distribution réelle ( la loi uniforme sur chaque paramètre étant prise comme base).
- un critère de taille maximum d'une assertion (basé sur la loi minimum de x lois binomiales pour calculer l'espérance du cardinal de l'assertion générée en fonction du cardinal du noeud) .

Ce dernier critère est utilisé pour réduire la liste des assertions générées par l'opérateur de génération de nouvelles assertions.

Tous ces critères sont évalués sur les données de la base d'entraînement rattachées à l'étiquette du noeud.

Le générateur d'assertion est basé sur l'algorithme de Clark et Niblett [4] avec l'hypercube comme domaine restreint, il est utilisé lors de l'apparition d'une ou de plusieurs nouvelles classes en phase de discrimination afin d'apprendre localement les assertions caractérisant ces classes.

Lorsque l'opérateur de génération de nouvelles assertions a été utilisé dans deux noeuds frères de l'arbre la branche de l'arbre qui part du noeud père est réapprise par le module d'apprentissage de l'arbre.

### Avantages pratiques :

- La prise en compte de la connaissance supplémentaire fournie par l'expert.

exemple : apprentissage locale , l'expert fournissant une règle particulière de discrimination entre deux modes radars

- La prise en compte dans l'identification d'une base de signatures des classes à discriminer.

Ceci est particulièrement adapté pour cette application puisque l'on dispose d'une bibliothèque technique d'identification, dans laquelle chaque mode radar est décrit comme une union d'hypercubes de l'espace des paramètres.

La bibliothèque technique est en effet la seule façon de prendre en compte la description de modes radars pour lesquels on ne dispose pas de pistes identifiées.

Inversement les vols sur lesquels on est amené à travailler comprennent des pistes radar qui proviennent de radars ne figurant pas encore dans la bibliothèque technique, il faut donc que le générateur de discriminateur puisse apprendre également à partir d'individus et non seulement à partir de signatures.

- L'apprentissage d'une ou de plusieurs signatures de nouvelles classes à partir des données identifiées comme n'appartenant pas à une classe connue.
- Le réapprentissage du discriminateur afin de prendre en compte cette nouvelle signature RADAR
- Cette capacité permettant la maintenance du discriminateur et de la bibliothèque technique RADAR

## **PARTIE IV**

### **CONCLUSION**

Sysiphe est un système d'identification et d'apprentissage utilisant à la fois des techniques symboliques et des techniques numériques. Ce système tient compte de façon intégrée et cohérente des éléments d'information suivants:

- Une base de données d'individus (des interceptions réalisées),
- Une base de données de signatures (des modèles prototypiques de radar)
- Une base des connaissances expertes (des règles de discrimination, de la connaissance sur les variables, des descriptions des classes).

Le système résultant a la capacité :

- De reconnaître des individus (impulsions radars) comme une occurrence d'une signature prototypique (signature radar de la bibliothèque) ou une occurrence proche d'individus préalablement identifiés (des interceptions RADAR antérieures),
- De détecter l'apparition d'un nouvel individu (impulsions provenant d'un nouveau mode radar) et de le distinguer d'une anomalie ponctuelle de mesure,
- D'apprendre une nouvelle signature prototypique (signature d'un nouveau radar dans la bibliothèque technique)
- De mettre à jour le procédé d'identification (nouvelle fonction d'identification de radars)..

Les principaux avantages sont :

- La capacité à traiter localement et donc de façon adaptée les différentes classes cibles.
- La capacité à tenir compte simultanément des informations décrites sous-forme symbolique ou numérique.

NB : Le résultat de l'algorithme est lié à la qualité de la base d'assertions fournie en entrée.

- L'aptitude à traiter les liens entre les variables à travers le filtre du domaine ainsi qu'à intégrer les règles expertes au niveau local.

L'aptitude à prendre en compte les hiérarchies sur les classes.

- La faculté à s'adapter au bruit. Ce dernier étant pris en compte au niveau de l'extension des assertions et également lors de la création ou lors du réapprentissage de l'arbre.
- La capacité à apprendre de nouvelles classes et à les prendre en compte dans le discriminateur.
- La capacité à apprendre de nouvelles assertions et à les prendre en compte dans le discriminateur.



Un prolongement, à l'étude de l'algorithme, est la prise en compte d'assertions modales de l'intérieur (cf E. Diday)

Cet algorithme a été conçu afin de prendre en compte les spécificités de l'application traitement de données E.L.I.N.T de RCM.

Il peut cependant s'appliquer de la même façon à tout problème de génération automatique d'un discriminateur, lorsque l'on dispose d'une base d'objets pré-définis, d'un ensemble d'individus d'apprentissage et que l'on désire maintenir à jour à la fois le discriminateur et la base d'objets.

Pour plus de détails sur l'algorithme, nous renvoyons le lecteur à l'article:

Discrimination d'objets symboliques par segmentation et apprentissage local. Les troisièmes journées " Symboliques-Numériques" Paris 14-15 Mai 1992.

### **Bibliographie**

- Diday E. (1991) Towards a statistic of Intentions for knowledge Analysis.  
WOCFAI ' 91 PARIS
- Bertrand P, Diday E, (1989) Une généralisation des arbres hiérarchiques :  
Les représentations pyramidales  
Revue Statistique Appliquée 1990XXXVIII p53-78
- Breiman L, Friedman J.H, Ohlsen R.A, Stone C.J (1984)  
Classification and Regression Trees Wadsworth & Brooks
- Clark. P, Boswell. (1991) R Rule Induction with CN2 :  
Some recent Improvements.  
Proceedings of the European Working Session on  
learning (Portugal)
- Ganascia J. G. (1991) Charade Apprentissage de bases de connaissances.  
Inférence Inductive de Données Symboliques-Numériques  
eds : Y. Kodratoff & Diday
- Gascuel O (1989) Inductive learning, numerical criteria and  
combinatorial optimization, some results.  
Data Analysis, Learning Symbolic and numeric knowledge  
(Antibes)

