

Réseaux de Neurones de Classement d'Objets dans un Modèle de Représentation de Connaissances

Cyr Gabin Bassolet, Ana Simonet et Michel Simonet
Laboratoire TIMC-IMAG, Faculté de Médecine de Grenoble
38706 LA TRONCHE CEDEX - FRANCE
e-mail : *Cyr-Gabin.Bassolet@imag.fr*,
(Ana,Michel).Simonet@imag.fr Fax +33 76-51-86-67

Mots clés Incertitude, base de connaissances, bases de données, orienté-objet, réseau de neurones.

Résumé Le modèle des p-types, implanté par un système appelé Osiris, est un modèle de représentation des connaissances centré-objet. Dans ce modèle, un utilisateur décrit une base de connaissances à travers un mécanisme de vues décrites par des propriétés logiques. La détermination pour un objet donné des vues pour lesquelles l'ensemble de ces propriétés logiques sont satisfaites est la problématique même du classement. Une analyse de la description des vues permet la construction de réseaux de classement d'objets dont l'architecture est totalement déterminée dans cette phase d'analyse.

Keywords Uncertainty, knowledge bases, databases, object-oriented, neural networks.

Abstract The Osiris system implements an objet data model wich was designed to enable users to share knowledge through a view mechanism. Views, wich play the role os subclasses in an object perspective, are defined by their logical properties. Determining the views whose properties are satisfied by a given instance is the very mechanism of classification in knowledge bases. A in-depth static analysis of the views and their logical properties allows the construction at compile time of a network to classify the objects of the base.

1. Introduction

Le modèle des p-types, modèle de représentation de connaissances centré-objet implanté par le système Osiris, a été défini dans la période de plein développement des bases de données relationnelles [SIM84]. Il a été défini dans le cadre de la théorie des types abstraits algébriques et diffère des bases de données objet standard dont la proposition de standard élaborée par le groupe ODMG est un bon représentant [CAT 94]. Une différence importante est la définition des classes par leurs propriétés logiques, qui sont exprimées dans le système Osiris par des assertions dont la forme générale correspond à des clauses de Horn dont les littéraux sont des prédicats sur les domaines des attributs. Cette approche s'apparente à celle des logiques terminologiques [GRE91], avec toute fois des différences notables. En particulier, l'implication n'est pas prise en compte pour la définition des concepts dans les langages usuels (Classic, Back, Loom) implantant une logique terminologique.

Etant donnée une base de connaissances décrite à travers des vues ayant chacune des propriétés logiques les définissant, l'utilisation de la base consiste à situer un objet donné dans cette description. Les propriétés logiques décrivant les vues sont alors assimilées à des contraintes que doivent satisfaire les valeurs des attributs d'un objet pour qu'il appartienne à la vue. Dans ce problème de classement, deux cas sont à distinguer : le cas où l'utilisateur décrivant l'objet connaît toutes les valeurs d'attributs de l'objet (objet totalement instancié), et le cas où l'utilisateur ne connaît pas toutes les valeurs d'attribut de l'objet (objet partiellement instancié). L'analyse des propriétés logiques décrivant les vues nous a conduit à résoudre ce problème de classement d'objets par des réseaux de neurones construits et exacts. En effet, l'architecture de

ces réseaux tels qu'ils seront présentés, est entièrement définie à la compilation de la base. Contrairement aux réseaux de neurones utilisés après apprentissage, ces réseaux ne font pas d'erreur de classement.

2. Les p-types

Les p-types permettent la modélisation d'un univers de concepts hétérogènes. Ces concepts seront organisés à travers des hiérarchies séparées de classes, telles que chaque hiérarchie de classe décrit un même concept.

Le modèle des p-types est fondé sur la théorie algébrique des types et sur la genericité [SIM84]. Très succinctement, la théorie algébrique se définit de la façon suivante :

Spécifier une algèbre hétérogène A équivaut à donner un triplet $\langle S, F, E \rangle$ tel que :

- S est une collection d'ensembles : support de A; l'ensemble dont l'algèbre sert à définir les éléments est nommé l'ensemble d'intérêt;
- F est une collection de fonctions indexées par S^*XS ;
- E est un ensemble d'axiomes que tous les éléments de l'algèbre doivent satisfaire.

notations $\langle S, F \rangle$: signature Σ de l'algèbre ; $\langle S, F, E \rangle$: Σ -E-algèbre.

La notion d'enrichissement est obtenu par un morphisme d'inclusion.

$$s : \Sigma \rightarrow \Sigma'$$

$$A \rightarrow A'$$

avec $A = \Sigma$ -E-algèbre, $A' = \Sigma'$ -E'-algèbre tel que $E \subset E'$.

Un p-type est un type abstrait qui définit un ensemble d'objets (P désignant le mot "Partagé") susceptibles d'être perçus selon plusieurs points de vue par diverses catégories d'utilisateurs. Sa définition part de la description la plus générale des éléments de cet ensemble (ensemble d'intérêt général appelé vue minimale), puis à l'aide de l'enrichissement (spécialisation) par introduction d'attributs et/ou d'assertions, autorise la description de sous-ensembles particuliers. Un élément de cet ensemble général de départ a donc potentiellement tous les attributs utilisés pour la description du p-type.

Dans ce modèle, chaque famille [MIN75][KAU87](ici une hiérarchie de vues) est donc organisée de façon hiérarchique en partant d'une vue minimale (le chef de la famille) puis de vues spécialisées à partir des vues déjà définies en suivant cette même méthode. Les familles sont organisées de sorte qu'une vue n'hérite que de vues de la même famille, assurant ainsi sur la base un regroupement de ses éléments en sous-ensembles homogènes [MAS89][BRA85](correspondant chacun à une même sémantique).



Fig. 1 Représentation hiérarchique d'un p-type (type partagé)

Pour qu'un objet appartienne à un p-type, il doit satisfaire la condition d'appartenance à la vue minimale A. La construction de la hiérarchie provient de la description que le concepteur a donnée de la base de connaissance et peut être complétée de façon explicite par un algorithme de subsomption de concepts comme dans les logiques terminologiques [GRE91].

2.1. Exemple

L'exemple est écrit dans le langage Osiris qui implante le modèle des p-types [SIM94].

p-type PERSONNE

attributes

Nss, Nom, S-Militaire : STRING;
Adresse : ADRESSE;
Enfants : set of PERSONNE;
Sexe : CHAR;
Age : INT;

assertions

a1 $0 \leq \text{Age} \leq 120$
a2 $\text{Age} < 17 \Rightarrow S_Militaire = 'non'$
a3 $\text{Age} \geq 18 \text{ et } \text{Sexe} = 'M' \Rightarrow S_Militaire \in \{ \text{sursitaire, oui, exempté} \}$
a4 $\text{Sexe} \in \{ 'M', 'F' \}$

end;

/* Spécification de quatre autres vues de *personne* : */

view ÉLÈVE : PERSONNE /* élève hérite de la définition de personne */ ... end;

view PROFESSEUR : PERSONNEend;

view PROFESSEUR-AGRÉGÉ : PROFESSEURend;

view ENSEIGNANT-ÉLÈVE : PROFESSEUR, ÉLÈVE /* héritage multiple */end;

Noter que la vue Enseignant-élève n'est pas nécessaire pour exprimer qu'un individu peut être à la fois élève et professeur. Le mécanisme des vues suffit car une instance peut être modèle de plusieurs vues. La vue Enseignant-élève est utile pour exprimer des propriétés supplémentaires et donc pour désigner un sous-ensemble de l'intersection ÉLÈVE et PROFESSEUR.

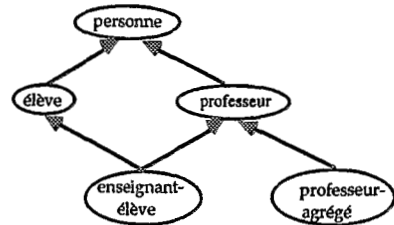


Fig. 2 Graphe d'appartenance de la famille d'objets *personne*

3. Le problème du classement d'objets

L'objectif principal est de pouvoir identifier un objet, c'est-à-dire déterminer la ou les vues auxquelles il appartient. Cette problématique est précisément celle du classement dans les bases de connaissances. Deux cas sont à distinguer :

- 1) L'objet que l'on veut classer a tous ses attributs instanciés.
- 2) On ignore certaines valeurs d'attribut lors de la description de l'objet.

Ces deux cas aboutissent à des objets partiellement instanciés. Il n'est plus possible de donner a priori toutes les vues auxquelles l'objet *appartient*, mais il sera intéressant d'énumérer les vues auxquelles *il pourrait appartenir* (vues potentielles), en plus de celles qui sont certaines. Il serait également intéressant de pouvoir dire pour deux vues potentielles laquelle est la plus

probable. La prise en compte de l'incertitude est un objectif important dans la conception des réseaux de neurones de classement que nous présenterons.

4. Réseaux de neurones de classement dans les p-types

Ce qui nous intéresse dans l'utilisation d'un réseau de neurones de classement d'objets dans le modèle des p-types, est de pouvoir en tirer une fonction de classement qui nous dispensera de parcourir à chaque fois la description des vues pour en vérifier les contraintes logiques qui les décrivent, pour les valeurs des attributs de l'objet à classer. Cette fonction est réalisée par un réseau de neurones. L'alternative du choix entre l'utilisation d'un réseau de classement après apprentissage et celle d'un réseau construit devant tenir compte de l'importance du fait d'effectuer des classements exacts d'objets. Or un réseau d'apprentissage (réseau multicouches, par exemple entraîné grâce à l'algorithme de rétropropagation du gradient) lors de sa phase de généralisation effectuée souvent même dans le meilleur des cas des erreurs qui sont de l'ordre de 10% des associations effectuées. La construction d'un réseau exact passe nécessairement par l'étude de l'espace des objets du point de vue des propriétés logiques décrivant les vues. Cette étude conduit à une partition du domaine de chaque attribut en sous-domaines stables

4.1. Définition 1: Sous-domaines stables des attributs

On peut construire pour chaque attribut d'une famille donnée une partition de son domaine de définition par tous les prédicats sur son domaine (qui n'ont que cet attribut dans leur description). Ces prédicats se trouvent définis lors de la description des vues du p-type. Suite à cette partition du domaine des attributs, lorsque l'on considère un attribut particulier, deux objets qui ont des valeurs de cet attribut appartenant à un même élément de la partition auront la même évaluation (même valeur de vérité) pour les prédicats sur cet attribut. Ces éléments de partition du domaine sont donc stables pour les prédicats sur le domaine de l'attribut ; c'est pourquoi ils portent le nom de sous-domaines stables. Par conséquent tous les objets se projetant sur une même combinaison de sous-domaines stables ont un même comportement pour une vue, i.e. ils ont la même valeur de vérité par rapport à l'ensemble d'assertions définissant la vue puisque ce sont elles qui ont conduit à cette partition.

Exemple :

Considérons l'exemple de la description de la vue *Personne* où nous ne nous intéressons qu'aux trois attributs *Sexe*, *S-Militaire* et *Âge* et aux assertions sur ceux-ci :

Partition du domaine des attributs

Le domaine de définition de l'attribut *Age* est partitionné en trois sous-domaines par les assertions *a1*, *a2* et *a3* où interviennent trois prédicats sur l'attribut *Age* :

$$\Delta(\text{AGE}) = [\text{INF} ; 17[\cup [17 ; 18[\cup [18 ; 120] = d11 \cup d12 \cup d13$$

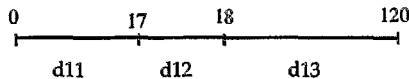


Fig.3 Partition du domaine de définition de l'attribut AGE

4.1. Définition 2 : les Eq-classes

Le prolongement à l'espace des objets de la partition du domaine de chaque attribut en sous-domaines stables définit une partition dont les éléments sont des n-uplets de sous-domaines stables qui sont appelés Eq-classes. L'ensemble des eq-classes, espace produit cartésien des sous-domaines stables, est appelé l'Espace de Classement du p-type.

Exemple de représentation :

Lorsque l'on travaille avec une famille d'objets à trois attributs, l'espace des objets est l'espace tridimensionnel. En reprenant l'exemple précédent avec les trois attributs AGE, S-MILITAIRE et SEXE avec les partition en trois, deux et deux sous-domaines stables respectivement, l'espace des objets est partitionné en 12 Eq-classes. C'est ce qu'illustre la figure ci-contre (Fig; 4);

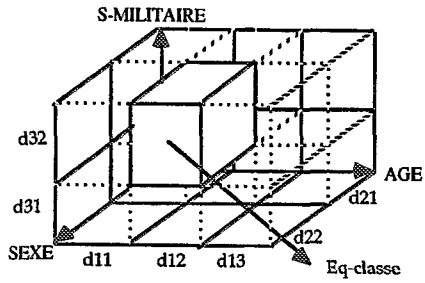


Fig. 4 Partition de l'espace des objets en Eq-classes

4.2. Réseau SDS-Eq-classes-Vues

Nous nous plaçons dans une situation où l'on dispose d'une distribution de probabilité de chaque attribut sur son domaine, exprimée sous la forme des probabilités p_{ij} (la probabilité que la valeur $attr_i$ de l'attribut $ATTR_i$ apparaisse sur le sous-domaine stable d_{ij}). On fait par ailleurs l'hypothèse (forte) que les valeurs des attributs sont indépendantes entre elles. La prise en compte des dépendances entre les attributs fait l'objet d'une étude en cours. Soit une famille d'objets à N attributs. Appelons n_j le nombre de sous-domaines stables de l'attribut $ATTR_j$ et p_{ij} la probabilité que la valeur $attr_i$ de l'attribut $ATTR_j$ apparaisse sur le sous-domaine stable d_{ij} .

Pour un i fixé compris entre 1 et N on a $\sum_{j=1}^{n_j} p_{ij} = 1$. Ces faits constituent notre hypothèse de départ. Soit une Eq-classe $F = (d1k_1, d2k_2, \dots, dnk_n)$. La probabilité qu'un objet se projette sur cette Eq-classe est $P(o \in F) = p_{1k_1} p_{2k_2} \dots p_{nk_n} = \prod_{i=1}^N p_{ik_i}$, où p_{ij} est la probabilité associée au sous-domaine stable d_{ij} (Fig. 5).

En appelant F_V l'ensemble des Eq-classes valides pour la vue V , on peut affirmer que la probabilité pour qu'un objet appartienne à cette vue est : $P(o \in V) = \sum P(o \in F)$ avec F parcourant F_V (Fig. 6).

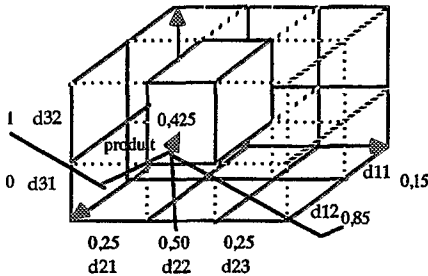


Fig. 5 Produit des probabilités associées aux sous-domaines définissant l'Eq-classe.

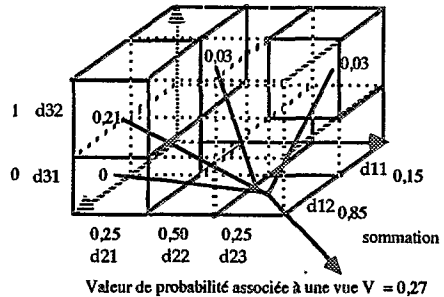


Fig. 6 Exemple de vue validée par quatre Eq-classes (sur 12)

Un réseau à trois couches permet d'intégrer les coefficients de probabilité de la façon suivante :

- Une **première couche en entrée** est constituée d'autant de cellules que de sous-domaines stables. On a alors deux situations possibles :

- Soit l'attribut ATTRi est connu, et il se projettera sur un seul sous-domaine stable qui aura la valeur 1 et les autres la valeur 0.

- Soit l'attribut ATTRi est inconnu, et toutes les cellules correspondant chacune à un sous-domaine stable de ATTRi recevront la valeur de probabilité associée au sous-domaine stable qu'il représente.

La deuxième couche est composée de N_2 neurones correspondant chacun à une Eq-classe déterminée : chaque Eq-classe est représentée par au plus un neurone (une Eq-classe qui n'est jamais valide n'est pas représentée). Chacun des neurones de cette deuxième couche est relié à un nombre de cellules de la première couche qui constituent ses entrées. Ces neurones prendront la valeur 1 si et seulement si les attributs de l'objet permettent de dire qu'il est sur la Eq-classe représentée par ce neurone de la deuxième couche.

Leur sortie calculée est le **produit de leurs entrées**.

La troisième couche est composée d'autant de neurones qu'il y a de vues. Ces neurones seront reliés seulement aux cellules de la deuxième couche qui peuvent lui faire prendre la valeur 1.

La sortie de ces neurones est la **somme des valeurs** qu'ils reçoivent en entrée.

$$\text{sortie} = \sum_i e_i$$

Ne sont représentées dans la deuxième couche que les Eq-classes validant au moins une vue.

Ce réseau à une complexité en taille égale au nombre d'Eq-classes obtenues après la partition de l'espace des objets. Sa complexité

est donc en $O\left(\prod_{i=1}^N n_i\right) = O(\bar{n}_i^N)$ où \bar{n}_i est le

nombre moyen de sous-domaines stables par attribut.

L'architecture du réseau est représentée sur la figure 7 ci-contre.

Du fait de sa taille exponentielle ce réseau n'est pas utilisé dans le système Osiris.

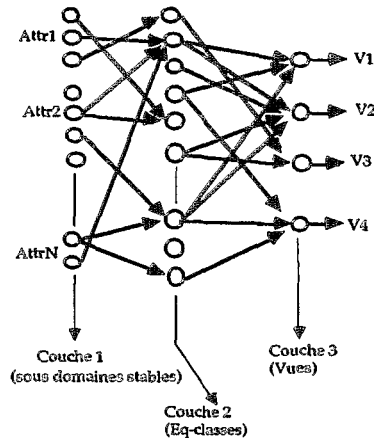


Fig. 7 Exemple de réseau pour un p-type de quatre vues.

4.3. Approche par les assertions : Réseau pseudo-exponentiel

Nous considérons maintenant une autre approche, fondée sur les assertions, qui permet d'aboutir à un réseau dont la complexité en taille varie du polynomial à l'exponentiel selon les cas. Soit a_j une assertion écrite dans une vue V . La forme générale de a_j est *Gauche* \Rightarrow *Droite* où *Gauche* et *Droite* sont respectivement des conjonctions d'assertions élémentaires pour ce qui concerne la partie *Gauche*, et des disjonctions d'assertions élémentaires sur les attributs pour la partie *Droite*.

Les assertions sont donc de la forme :

$a_j : P_1(\text{ATTR}_1) \text{ et } P_2(\text{ATTR}_2) \text{ et... et } P_n(\text{ATTR}_n) \Rightarrow P_{n+1}(\text{ATTR}_{n+1}) \text{ ou... ou } P_{n+k}(\text{ATTR}_{n+k})$
 où $P_i(\text{ATTR}_i)$ est une disjonction de prédicats portant sur un même attribut ATTR_i , dont la forme générale est : $(\text{ATTR}_i \in \text{Domaine}_1 \text{ ou } \text{ATTR}_i \in \text{Domaine}_2 \text{ ou ... } \text{ATTR}_i \in \text{Domaine}_k)$.

Après la détermination des sous-domaines stables de chaque attribut, chaque prédicat $P_i(\text{ATTR}_i)$ est transformé en l'assertion disjonctive : $(\delta_{ia'_1} \text{ ou } \delta_{ia'_2} \text{ ou... ou } \delta_{ia'_p})$

où $\alpha'_k \in \{1, 2, \dots, n_j\}$ (n_j : nombre total de sous-domaines stables de ATTR_j).

$a_j = P_1'(d_1) \text{ et } P_2'(d_2) \text{ et... et } P_n'(d_n) \Rightarrow P_{n+1}'(d_{n+1}) \text{ ou... ou } P_{n+k}'(d_{n+k})$

avec $P'(d_j) = (\delta_{ja'_1} \text{ ou } \delta_{ja'_2} \text{ ou... ou } \delta_{ja'_p})$

En projetant chaque valeur d'attribut attr_j sur son sous-domaine stable, on obtient un vecteur de dimension N : $(\delta_{1\alpha_1}, \delta_{2\alpha_2}, \dots, \delta_{N\alpha_N})$ où α_i est le numéro de sous-domaine stable de attr_i . Ce vecteur, obtenu par projection des valeurs d'attribut sur leur sous-domaine stable, représente l'Eq-classe de l'objet. Si l'on appelle ξ_i l'ensemble des sous-domaines stables de ATTR_i , le produit cartésien $\xi_1 \times \xi_2 \times \dots \times \xi_N$ désigne l'ensemble de toutes les Eq-classes, qui constitue l'espace de classement. En considérant la table de vérité de l'implication, $(p \Rightarrow q)$ est équivalent à $(\neg p \vee q)$ on peut caractériser les Eq-classes non valides (\bar{a}_i) (que nous noterons

aussi $E(\bar{a}_i)$) pour a_i . Cela revient à décrire le produit cartésien qui est :

$(\bar{a}_i) = E_1 \times E_2 \times \dots \times E_N$ où $E_i \subset \xi_i$ est déterminé comme suit :

$E_i = \xi_i$ si ATTR_i n'intervient pas dans l'assertion a_j ;

$E_i = \{d_{ij} \in P'(d_i) \text{ si } \text{ATTR}_i \text{ intervient en partie gauche de } a_j\}$;

$E_i = \{d_{ij} \notin P'(d_i) \text{ si } \text{ATTR}_i \text{ intervient en partie droite de } a_j\}$.

Cet ensemble (\bar{a}_i) des Eq-classes non valides pour l'assertion a_j a été déterminé grâce à la table de vérité de " $p \Rightarrow q$ " (*Gauche* vraie et *Droite* fausse). C'est le seul cas où l'assertion est fausse. L'ensemble $E^{(a_i)}$ des Eq-classes qui sont valides pour l'assertion a_j ne s'énumère pas

sous forme d'un seul produit cartésien. Les ensemble $E^{(a_i)}$ et $E(\bar{a}_i)$ forment une partition de

l'espace E des objets : $E = E^{(a_i)} \cup E(\bar{a}_i)$. Une approche probabiliste permet d'avoir une idée du degré de satisfaction de l'assertion pour la description partielle de l'objet.

4.4. Probabilité de satisfaction d'une assertion

Nous appelons $P_O(a_j)$ la probabilité que l'objet O satisfasse l'assertion a_j . De même on appellera $P_O(\bar{a}_i)$ la probabilité que l'objet o ne satisfasse pas l'assertion a_j .

On a : $P_O(a_j) = P(o \in E^{(a_i)})$, $P_O(\bar{a}_i) = P(o \in E(\bar{a}_i))$ avec la relation $P_O(a_j) = 1 - P_O(\bar{a}_i)$.

$E(\bar{a}_i) = E_1 \times E_2 \times \dots \times E_N$ où E_i est un ensemble de sous-domaines stables de ATTR_i .

La probabilité que l'objet o appartienne à $E(\bar{a}_i)$ sous les hypothèses posées au départ est le produit des probabilités que chaque attribut Attr_i de l'objet se projette sur un sous-domaine d_i appartenant à E_i . En appelant S_i la somme des probabilités p_{ij} élémentaires associées aux sous-

domaines stables d_{ij} appartenant à E_i , on a : $P_0(a_j) = 1 - S_1 * S_2 * S_3 * ... * S_N$, probabilité que l'assertion a_j soit vraie pour les valeurs d'attribut de l'objet o .

Remarque : $S_i = 1$ lorsque E_i est égal à ξ_i , ensemble de tous les sous-domaines stables de $Attr_j$. Rappelons que ce cas se produit entre autres lorsque $ATTR_j$ n'intervient pas dans l'assertion.

Pour une vue V avec n assertions a_j et sachant les probabilités $P_0(a_j)$, il faut déterminer la probabilité que l'objet respecte toutes les assertions de la vue.

4.5. Probabilité d'appartenance d'un objet à une VUE

Pour être un élément de la vue V , un objet o doit avoir des valeurs d'attribut satisfaisant chacune des assertions a_j décrivant la vue V . La condition nécessaire et suffisante pour qu'un objet appartienne à une vue est que cet objet se situe dans l'intersection des ensembles d'Eq-classes validant chaque assertion a_j de la vue : $P(o \in V) = P(o \in (E^{(a_1)} \cap E^{(a_2)} \cap ... \cap E^{(a_n)}))$ qu'on note $P_0(a_1 \cap a_2 \cap ... \cap a_n)$, où $E^{(a_i)}$ est l'ensemble des Eq-classes qui valident a_j .

La formule que nous utilisons met à profit la complémentarité :

$$P(o \in V) = 1 - P(\overline{a_1} \cup \overline{a_2} \cup ... \cup \overline{a_n})$$

Ceci est plus facile à calculer du fait que les ensembles $(\overline{a_i})$ sont énoncés sous forme d'un seul produit cartésien. Nous proposons une représentation possible de l'ensemble $(\overline{a_1} \cup \overline{a_2} \cup ... \cup \overline{a_n})$; Il s'agit de chercher un ensemble équivalent de produits cartésiens (b_i) tels que $\bigcup_{i=1}^n \overline{a_i} = \bigcup_{i=1}^K b_i$ avec $b_i \cap b_j = \emptyset$. On a alors $P(o \in V) = 1 - P(\overline{a_1} \cup \overline{a_2} \cup ... \cup \overline{a_n}) = 1 - \sum_{i=1}^K P(b_i)$.

L'intersection de produits cartésiens se fait aisément et donne un seul produit cartésien, mais ceci n'est pas le cas pour l'union.

Intersection de deux produits cartésiens :

$$E^{(\overline{a_i})} = E_1 \times E_2 \times ... \times E_N \quad E^{(\overline{a_j})} = E_1' \times E_2' \times ... \times E_N'$$

On a $(\overline{a_i}) \cap (\overline{a_j}) = E^{(\overline{a_i})} \cap E^{(\overline{a_j})} = E_1'' \times E_2'' \times ... \times E_N''$ avec $E_i'' = E_i \cap E_i'$.

Union de produits cartésiens :

Soient n produits cartésiens $E^{(\overline{a_i})}$ $i=1..n$ notés $E^{(\overline{a_i})} = E_1^i \times E_2^i \times ... \times E_N^i$. E_j^i étant un ensemble de sous-domaines stables de l'attribut $ATTR_j$ et N le nombre d'attributs du p-type.

1. Utiliser les ensembles E_j^i pour effectuer une partition de ξ_j (ensemble des sous-domaines stables de $ATTR_j$). Deux sous-domaines stables d_{jp} et d_{jq} appartiennent à un élément de cette partition si et seulement si ils apparaissent toujours simultanément dans les ensembles E_j^i pour $i=1..n$.

2. Effectuer une énumération explicite en produits cartésiens des ensembles $E^{(\overline{a_i})}$ selon cette nouvelle partition.

3. Faire l'union en créant une liste des $E^{(\overline{a_i})}$ ainsi éclatés, en éliminant les éléments répétés (intersection des différents produits cartésiens).

Résultat : On aboutit à une liste de produits cartésiens qui sont deux à deux disjoints entre eux et qui représente les ensembles de départ. On verra plus loin l'évaluation de la complexité en taille de cette liste. L'explication plus détaillée de ce procédé est donnée dans [BAS96].

Pour calculer $P_0(a_i)$, l'ensemble des Eq-classes non valides pour (a_i) étant décrit sous la forme d'un produit cartésien

$E^{(\bar{a}_i)} = E_1 \times E_2 \times \dots \times E_N$, on associe à chaque E_j un neurone pour le calcul de S_j , somme des probabilités associées aux sous-domaines de E_j quand $S_j \neq 1$, puis un neurone qui calcule $P_0(a_i)$ qui est $1 - S_j$.

Le sous-réseau de calcul des $P_0(a_i)$, est le suivant :

Les poids w_{ij} associés à la connexion du neurone d'entrée correspondant à d_{ij} vers le neurone calculant S_j sont définis comme suit : $w_{ij} = 1$ si $d_{ij} \in E_j$ et $w_{ij} = 0$ sinon.

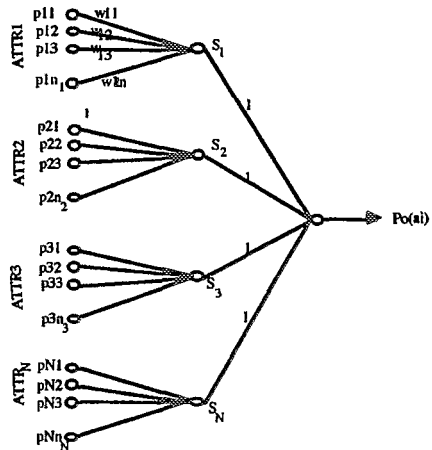


Fig.8 Sous-réseau de calcul de $P_0(a_i)$

Puisque $P(o \in V) = 1 - P(\bar{a}_1 \cup \bar{a}_2 \cup \dots \cup \bar{a}_n)$, la probabilité d'appartenance d'un objet à une vue sera la somme des probabilités d'appartenance aux ensembles disjoints obtenus pour la représentation de $\bar{a}_1 \cup \bar{a}_2 \cup \dots \cup \bar{a}_n$. Pour la construction du réseau, on associe à chaque produit cartésien $E_1 \times E_2 \times \dots \times E_N$, $(N+1)$ neurones pour le calcul de la probabilité que l'objet décrit en entrée appartienne à cet ensemble.

Un neurone calcule pour E_j la somme $S_j = \sum p_{ij}$ avec p_{ij} associé à d_{ij} appartenant à E_j . Un dernier neurone (associé au produit cartésien proprement dit) calculera le produit des valeurs de sortie S_j . La somme des sorties respectives des neurones associés aux produits cartésiens représentant $\bar{a}_1 \cup \bar{a}_2 \cup \dots \cup \bar{a}_n$ sera effectuée par un neurone, qui est celui associé à la vue où étaient décrites les assertions considérées. Sa valeur de sortie est donc la probabilité que l'objet décrit en entrée du réseau appartienne à la vue correspondante. Prendre en compte l'héritage revient simplement à considérer qu'une vue dérivée comprend toutes les assertions de ses vues ancêtres. En taille, ce réseau peut être exponentiel du fait que la recherche des produits cartésiens (b_i) peut aboutir au nombre d'Eq-classes et dans ce cas il rejoint le premier réseau présenté. Par contre ce dernier réseau est basé sur la ressemblance qu'il peut y avoir dans la description des différentes assertions des vues. En tirant profit de ce fait, il peut se réduire en une taille polynomiale[BAS96]. Le meilleur cas (théorique) étant celui où toutes les assertions sont identiques.

4.6. Un réseau approché de classement d'objets

Pour mettre à profit la possibilité de déterminer la probabilité de satisfaction d'une assertion par un objet donné, on pourrait utiliser les sorties des sous-réseaux associés aux calculs de ces valeurs de probabilité respectives pour effectuer par un neurone associé à la vue où sont décrites ces assertions (assertions propres plus celles héritées) une fonction de transition donnant la valeur 0 lorsqu'une de ses entrées est nulle, i.e., une assertion est fautive : vue invalide; - 1 quand toutes les entrées sont égales à 1 (toutes les assertions satisfaites) : vue valide;

- une valeur quelconque comprise entre 0 et 1 dans les autres cas c'est-à-dire lorsque toutes les entrées sont différentes de 0 et ne sont pas toutes égales à 1 : vue potentielle.

Une fonction de transition simple consiste à prendre pour sortie de ce neurone le minimum de ses entrées. On obtient un réseau qui classe bien les objets sur des vues mais qui donne une valeur comprise entre 0 et 1 pour les vues potentielles qui ne correspondent plus à une valeur de probabilité exacte. L'un des intérêts de ce réseau est sa complexité en taille, qui est faible :

$$\sum_{i \in \{1, \dots, N\}} n_i + (N + 1) * n_a + n_v$$
 neurones ; avec n_i : nombre de sous-domaines stables de l'attribut ATTR_i ; N : nombre d'attributs ; n_a : nombre d'assertions du p-type et n_v : nombre de vues.

5. Conclusion

Le modèle des p-types est un modèle de représentation de connaissances centré-objet. Il permet à travers un mécanisme de vues de décrire un concept et d'affiner à travers l'héritage, la description du concept général. L'analyse des propriétés logiques décrivant les vues permet la construction de réseaux de neurones de classement d'objet. Ceci a l'intérêt de pouvoir éviter des vérifications dynamiques pour la détermination des vues d'un objet. On aboutit à une fonction de classement distribuée à travers des processus élémentaires effectués par les différentes cellules du réseau. Cette vérification statique de la description de la base permet d'envisager l'étude de la cohérence de la base, ce qui est souvent un outil intéressant pour le concepteur d'une base de connaissances.

La fonction d'un système expert revient à effectuer du classement de cas qui sont totalement ou partiellement décrits. En comparaison, le modèle des p-types est un modèle où la notion de chaînage avant et arrière disparaît. L'identification des classes d'un objet ne s'effectue pas par remise en cause de la validité de vues : les vues qui étaient valides ou non valides le resteront, tandis que celles qui étaient potentielles deviendront soit valides ou invalides. C'est en cela qu'il diffère également de la méthode de maintien de la cohérence dans les TMS et ATMS (Truth Maintenance Systems et Assumption Truth Maintenance Systems)[DOY78][MAR90].

6. Bibliographie

- [BAS96] C.-G. Bassolet, Réseaux de Neurones de Classement dans le modèle des P-types, Rapport de Recherches IMAG, Grenoble, France, Dec. 1995.
- [BIR73] BIRTWISTLE G., DALL O., MYHRHAUG B., and NYGAARD K. Simulation begin. Petrocelli Charter, New York, 1973.
- [BOU91] BOURRET P., REGGIA J., SAMUELIDES M. Réseaux neuronaux. Ed. Tecknea, nov 1991.
- [BRA77] BRACHMAN R.J. What's in a concept : Structural Foundations for Semantic Networks. International Journal of Man-Machine Studies, 1977.
- [BRA85] BRACHMAN R.J., LEVESQUE H.J. Readings in knowledge representation. Morgan Kaufmann Publishers, Inc. 1985.
- [CAT94] R. G. G. Cattell, T. Atwood, J. Duhl, G. Ferran, M. Loomis, D. Wade, Object Database Standard : ODMG-93, Morgan Kaufmann Publishers, 1994.
- [DOY78] DOYLE J., Truth maintenance systems for problem solving, Massachusetts Institute of Technology, Cambridge, 1978.
- [GOL83] GOLBERG A., ROBSON D. Smalltalk-80, The langage and its Implementation. Addison-Wesley, Reading, Massachusetts, 1983.
- [GRE91] R. M. MacGregor, The evolving Technology of Classification -based Systems, Principles of Semantic Networks : Explorations in Representation of Knowledge, pp 385-400, J. F. Sowa ed., Morgan Kaufmann Publishers, 1991.

- [HEW73] HEWITT C.E., BISHOP P., STREIGER. A Universal Modular ACTOR Formalism for Artificial Intelligence. In Proceedings of the 3rd IJCAI, Stanford, California, 1973.
- [KAU87] KAUFMANN A. Nouvelles logiques pour l'intelligence artificielle. Ed Hermès, Paris, 1987.
- [MAS89] MASINI G., NAPOLI A., COLNET D., LEONARD D., TOMBRE K. Les langages à objets. Ed InterEditions, Paris, 1989.
- [MAR90] MARTINS J.P. Truth maintenance systems for problem solving : proceedings, Ed. Reinfrack, 1990.
- [MIN75] MINSKY M.A. A framework for representing knowledge. in WINSTON P. (ed.), The psychology of computer vision, Mc Graw-Hill, New york, pp 211-277, 1975.
- [PAU92] PAUGAM-MOISY H., Optimisation des réseaux de neurones artificiels : analyse et mise en oeuvre sur ordinateurs massivement parallèles. Thèse ENS Lyon, Janvier 1992.
- [SIM84] SIMONET A. Types abstraits et bases de données. Formalisation de la notion de partage et analyse statique des contraintes d'intégrité. Thèse, Université Grenoble1, 1984.
- [SIM89] SIMONET A., SIMONET M. Les Eq-classes : classes d'équivalence induites par les dépendances inter-attributs dans les familles d'objets. Rapport de recherche, IMAG, ARTEMIS, 1989.
- [SIM94] SIMONET A., SIMONET M. *Objects with Views and Constraints : from Databases to Knowledge bases*, OOIS'94 (Object-Oriented Information Systems), London Dec. 1994, Springer Verlag, pp 182-197.