

Examen d'Intelligence Artificielle

19 juin 2008

Durée 2h00 - Tous documents autorisés

Nom :

Notations /22

Exercice 1.1 : 1Pt

Exercice 1.2 : 1Pt

Exercice 1.3 : 2Pt

Exercice 1.4 : 1Pt

Exercice 1.5 : 2Pt

Exercice 1.6 : 1Pt

Exercice 2.1 : 1Pt

Exercice 2.2 : 2Pt

Exercice 3 : 4Pt

Exercice 4.1 : 1Pt

Exercice 4.2 : 2Pt

Exercice 5.1 : 1Pt

Exercice 5.2 : 1Pt

Exercice 5.3 : 2Pt

Notes globales :

1 Apprentissage supervisé et non supervisé

On cherche à apprendre le *concept "Faire un Barbecue"* à partir de l'observation des conditions météorologique . la table 1 nous montre un ensemble d'apprentissage. Une instance est décrite par les attributs : $\{CIEL, TEMP, HUMI, VENT\}$ définis respectivement sur les domaines suivants : $D(CIEL) = \{ensoleille, couvert, pluvieux\}$, $D(TEMP) = \{eleve, moyenne, basse\}$, $D(HUMI) = \{forte, normale, moyenne\}$, $D(VENT) = \{oui, non\}$.

NUM	CIEL	TEMP.	HUMI.	VENT	Concept
1	ensoleillé	élevé	forte	non	1
2	ensoleillé	élevé	forte	oui	1
3	couvert	élevé	normale	non	0
4	pluvieux	moyenne	normale	non	0

TAB. 1 – Description par extension du concept "Faire un Barbecue"

Exercice 1.1 Donner le résultat de l'application de FIND-S.

Corrigé 1.1 $\langle \text{ensoleillé, élevé, forte, ?} \rangle$

Exercice 1.2 Un nouvel exemple positif, donc pour lequel $c(x)$ vaut 1, se présente : $\langle \text{ensoleillé, élevé, normale, non} \rangle$. L'hypothèse résultat est-elle consistante avec cet exemple ? Que peut-on en déduire ?

Corrigé 1.2 Non l'hypothèse n'est pas consistante avec cet exemple. Donc cette hypothèse n'est pas assez générale car elle est induite seulement à partir des exemples positifs.

Exercice 1.3 Donner le résultat de l'application de candidate-elimination sur les quatre exemples d'apprentissage. En déduire l'espace de versions.

Corrigé 1.3

$$S = \langle \text{ensoleille, leve, forte, ?} \rangle$$

$$G = \{ \text{ensoleille, ?, ?, ?} \rangle, \langle \text{?, ?, forte, ?} \rangle \}$$

Exercice 1.4 Coder la base d'exemples E sous forme de base de transactions, pour cela nous proposons l'ensemble d'items suivant :

$$E = D(CIEL) \cup D(TEMP) \cup D(HUMI) \cup D(VENT) \cup \{1, 0\}$$

Par exemple : l'exemple numéro 1 serait codé par la transaction : $\langle \text{ensoleillé élevé forte non 1} \rangle$ de longueur 5.

Corrige 1.4 $\langle \text{ensoleillé élevé forte non 1} \rangle$

$\langle \text{ensoleillé élevé forte non 1} \rangle$

$\langle \text{ensoleillé élevé forte oui 1} \rangle$

$\langle \text{couvert élevé normale non 0} \rangle$

$\langle \text{pluvieux moyenne normale non 0} \rangle$

Exercice 1.5 Extraire les sous-ensembles d'items fréquents en appliquant l'algorithme *apriori*, avec $minSupp = 2$.

Corrigé 1.5 $L1 = \{\text{enseleillé, élevé, forte, normale, non}, 1, 0\}$

$L2 = \{\{\text{enseleillé élevé}\}, \{\text{enseleillé forte}\}, \{\text{élevé forte}\}, \{\text{enseleillé } 1\}, \{\text{élevé } 1\}, \{\text{forte } 1\}, \{\text{normale } 0\}, \{\text{non } 0\}\}$

$L3 = \{\{\text{enseleillé élevé forte}\}, \{\text{enseleillé élevé } 1\}, \{\text{élevé forte } 1\}, \{\text{normale non } 0\}\}$
 $L4 = \{\{\text{enseleillé élevé forte } 1\}\}$

Exercice 1.6 Trouver les règles d'association ayant comme résultat (ou conclusion) *l'item 1* avec $minConf = 1$; comparer avec le résultat de Candidate-elimination.

Corrigé 1.6 enseleillé élevé forte $\rightarrow 1$
enseleillé élevé $\rightarrow 1$
enseleillé forte $\rightarrow 1$
élevé forte $\rightarrow 1$
enseleillé $\rightarrow 1$
forte $\rightarrow 1$

On retrouve l'espace de version !!

2 Application des algorithmes arc-consistances : problème simple

Soit le problème suivant : $X = \{X_1, X_2, X_3\}$, $D(X_1) = D(X_3) = \{0, 1\}$, $D(X_2) = \{0, 1, 2\}$ avec les contraintes : $c_1 : X_1 < X_3$ et $c_2 : X_3 < X_2$

Exercice 2.1 Est ce que le système est arc-consistant ? justifier.

Corrigé 2.1 Non le système n'est pas arc-consistant, par exemple pour la valeur $X_1 = 1$ il n'y a pas une valeur de X_3 qui ne viole pas la contrainte c_1 .

Exercice 2.2 Appliquer l'algorithme AC3 pour rendre le système arc-consistant. Y-a t-il une solution au système ?

Corrigé 2.2

$$reviser(C_1, X_1) \rightarrow D(X_1) = \{0\}$$

$$reviser(C_2, X_3) \rightarrow D(X_3) = \{0, 1\}$$

$$reviser(C_1, X_3) \rightarrow D(X_3) = \{1\}$$

$$reviser(C_2, X_2) \rightarrow D(X_2) = \{2\}$$

Le système est consistant et la solution est $X_1 = 0$, $X_2 = 2$, $X_3 = 1$.

3 La résolution inverse

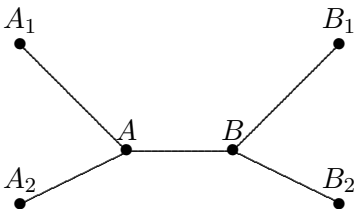
Nils J. Nilsson, un personnage important de l'IA, dans son livre "Introduction to Machine Learning" essaie d'expliquer à Toto les techniques d'utilisation des bus pour se déplacer d'un endroit à un autre.

Il y a d'abord des gares centrales des bus et on peut aller directement d'une gare centrale à n'importe quelle autre.

À chaque gare centrale sont rattachées des gares satellites et on peut aussi aller directement d'une gare centrale à une de ses gares satellites rattachées.

Par contre, on ne peut pas aller directement ni d'une gare satellite à une autre gare satellite, ni d'une gare satellite à une gare centrale autre que celle à laquelle est rattachée.

Par exemple sur la figure suivante, A et B sont des gares centrales, A1 et A2 sont des gares satellites de A et B1, B2 sont des gares satellites de B.



On peut aller directement de A à B, de A1 à A, de B2 à B, mais on ne peut pas aller directement de A1 ni à B, ni à B1.

Toto doit comprendre le concept "aller directement d'une gare à une autre". Pour l'aider on vous donne la base des connaissances suivante

$$X = \{\text{direct}(a,a1), \text{direct}(a,b), \text{direct}(b,b1)\}$$

et la base des connaissances supplémentaire

$$B = \{\text{centrale}(a), \text{centrale}(b), \text{satellite}(a,a1), \text{satellite}(b,b1)\}$$

En appliquant la résolution inverse, établir les deux clauses du concept $\text{direct}(X, Y)$.

Solution de l'exercice 3

4 Arbre de décision

Soit un échantillon de 200 patients se répartissant en 2 classes : M pour malade et B pour bonne santé. Deux attributs, gorge irritée et température, permettent de répartir les patients dans chacune des classes suivant le tableau suivant :

	Gorge irritée	Gorge non irritée
Température $< 37,5$	$6B,37M$	$91B,1M$
Température $\geq 37,5$	$2B,21M$	$1B,41M$

Exercice 4.1 Quel(s) arbre(s) de décision peut-on construire à partir de ces données en utilisant le critère du gain d'information (justifier le choix des attributs sans effectuer les calculs).

Exercice 4.2 On ajoute le critère d'arrêt suivant à l'algorithme de construction de l'arbre : si 90% des exemples d'un noeud sont d'une même classe, alors ce noeud devient une feuille de cette classe. Quel arbre obtient-on ? Quel est son taux d'erreur ?

5 Inférence grammaticale

Soit $R = \{b, aa, aba\}$ un ensemble de mots sur l'alphabet $\Sigma = \{a, b\}$. On veut construire une grammaire régulière qui reconnaît l'ensemble des mots de R . Nous allons donc inférer un automate d'états finis (équivalent aux grammaires régulières) selon la technique de l'agrégation des suffixes :

Exercice 5.1 Construire l'automate canonique minimal à partir de R .

Exercice 5.2 Construire le tableau des distances entre ensembles de suffixes.

Exercice 5.3 Appliquer l'algorithme de réduction du tableau par fusion d'états.