

Slide 1

**GloBo**

Fabien Torre ([fabien@lri.fr](mailto:fabien@lri.fr))  
Équipe Inférence et Apprentissage  
Laboratoire de Recherche en Informatique  
Université Paris-Sud, Orsay

CAP'99 - 17 juin 1999

Slide 2

**Plan de l'exposé**

- Apprentissage supervisé disjonctif et apprentissage non supervisé
- Algorithme stochastique de GloBo
- Critère de réussite
- Expérimentations

Slide 3

**Apprentissage supervisé conjonctif**

$E^+$		$E^-$
x,x,x,x,o,o,x,o,o		x,b,x,o,o,o,b,b,x
x,x,x,x,o,o,o,x,o		x,b,o,x,b,o,b,x,o
x,x,x,x,o,b,o,o,b		o,x,x,x,o,x,o,b,o
x,x,x,x,o,b,b,o,o		o,x,x,o,o,b,x,x,o
x,x,x,x,b,o,o,o,b		o,x,b,x,o,x,x,o,o
x,x,x,x,b,o,o,b,o		o,o,x,o,x,b,o,x,x
x,x,x,o,x,o,o,o,x		o,o,o,x,b,b,b,x,x
x,x,x,o,x,o,b,o,b		o,b,x,x,o,o,x,x,o
x,x,x,o,x,o,b,b,o		o,b,b,o,x,x,o,b,x
x,x,x,o,x,b,o,o,b		b,x,o,x,o,b,o,x,b
x,x,x,o,x,b,b,o,o		b,o,o,x,o,x,x,o,x

$mg(E^+) = x,x,x,?,?,?,?,?$

Slide 4

**Apprentissage supervisé disjonctif**

$E^+$		$E^-$
x,x,x,o,x,o,o,x,o		x,b,x,o,o,o,b,b,x
x,x,o,x,o,b,x,b,o		x,b,o,x,b,o,b,x,o
x,o,o,x,x,o,x,b,b		o,x,x,x,o,x,o,b,o
x,o,b,b,x,o,o,x,x		o,x,x,o,o,b,x,x,o
x,b,b,x,o,b,x,b,o		o,x,b,x,o,x,x,o,o
o,x,b,x,x,o,o,x,b		o,o,x,o,x,b,o,x,x
o,o,b,x,x,x,b,b,b		o,o,o,x,b,b,b,x,x
o,b,b,o,o,x,x,x,x		o,b,x,x,o,o,x,x,o
b,o,x,o,b,x,o,x,x		o,b,b,o,x,x,o,b,x
b,b,x,b,o,x,o,b,x		b,x,o,x,o,b,o,x,b
b,b,b,o,o,b,x,x,x		b,o,o,x,o,x,x,o,x

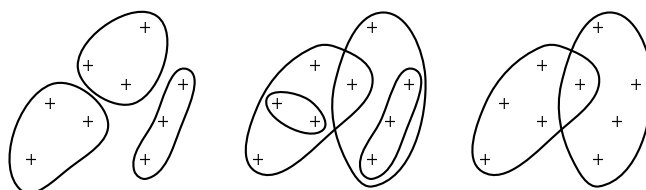
$mg(E^+) = ?,?,?,?,?,?,?,?$

Slide 5

### Difficultés de l'apprentissage supervisé disjonctif

Conjonctif : toutes les solutions sont équivalentes

Disjonctif : différentes couvertures des exemples positifs



Difficile pour les approches par couverture et pour les méthodes de type diviser pour régner [Boström, 1995].

Slide 6

### Problème du Morpion

Caractériser les fins de jeu au morpion qui sont gagnantes pour les croix : la solution est la disjonction des huit manières de faire des lignes de croix. On dispose de 958 exemples.

Expérimentations : on utilise 70 % des exemples pour apprendre, les 30 % restants pour le test [Aha, 1991].

default	65.3 %	IB1	98.1 %
NewID	84.0 %	IB3	82.0 %
CN2	98.1 %	IB3-CI	99.1 %
MBRtalk	88.4 %		

Slide 7

### Problème général

Étant donné un ensemble  $E$  et une propriété  $P$  sur les sous-ensembles de  $E$ , trouver une couverture minimale de  $E$  utilisant des sous-ensembles de  $E$  qui soient maximale-ment admissibles par rapport à  $P$ .

supervisé :  $P(S) \Leftrightarrow \forall n \in E^-, \text{mg}(S) \not\geq n$

non supervisé :  $P(S) \Leftrightarrow \forall (a, b) \in S, d(a, b) \leq d_{max}$

Slide 8

### Algorithme de GloBo : sélection stochastique

graine	autres positifs	→	paquet maximale-ment correct
$p_1$	$p_5 p_8 \cancel{p_7} p_{14} \dots$	→	$\{p_1, p_5, p_8, p_{14}, \dots\}$
$p_2$	<del><math>p_1</math></del> $p_3 \cancel{p_7} p_{12} \dots$	→	$\{p_2, p_3, p_{12}, \dots\}$
$p_3$	$p_7 \cancel{p_1} \cancel{p_5} \cancel{p_8} \dots$	→	$\{p_3, p_7, \dots\}$
$p_4$	$p_{13} p_3 \cancel{p_7} \cancel{p_8} \dots$	→	$\{p_4, p_{13}, p_3, \dots\}$
$p_5$	$p_8 \cancel{p_7} p_1 p_{14} \dots$	→	$\{p_5, p_8, p_1, p_{14}, \dots\}$
$\vdots$	$\dots$	→	$\dots$

Slide 9

**Paquets verrouillés et paquets condamnés**

<table border="1" style="width: 100%; height: 100%; border-collapse: collapse;"> <tr><td>×</td><td></td><td></td></tr> <tr><td>×</td><td>○</td><td></td></tr> <tr><td>×</td><td></td><td>○</td></tr> </table>	×			×	○		×		○	<table border="1" style="width: 100%; height: 100%; border-collapse: collapse;"> <tr><td>×</td><td>○</td><td>○</td></tr> <tr><td>×</td><td>×</td><td>○</td></tr> <tr><td>×</td><td></td><td></td></tr> </table>	×	○	○	×	×	○	×			<table border="1" style="width: 100%; height: 100%; border-collapse: collapse;"> <tr><td>×</td><td></td><td></td></tr> <tr><td>×</td><td>○</td><td></td></tr> <tr><td>×</td><td>○</td><td></td></tr> </table>	×			×	○		×	○		<table border="1" style="width: 100%; height: 100%; border-collapse: collapse;"> <tr><td>×</td><td>?</td><td>?</td></tr> <tr><td>×</td><td>?</td><td>?</td></tr> <tr><td>×</td><td>?</td><td>?</td></tr> </table>	×	?	?	×	?	?	×	?	?
×																																							
×	○																																						
×		○																																					
×	○	○																																					
×	×	○																																					
×																																							
×																																							
×	○																																						
×	○																																						
×	?	?																																					
×	?	?																																					
×	?	?																																					
<table border="1" style="width: 100%; height: 100%; border-collapse: collapse;"> <tr><td>×</td><td></td><td></td></tr> <tr><td>×</td><td>○</td><td>○</td></tr> <tr><td>×</td><td></td><td></td></tr> </table>	×			×	○	○	×			<table border="1" style="width: 100%; height: 100%; border-collapse: collapse;"> <tr><td>×</td><td>×</td><td>×</td></tr> <tr><td></td><td>○</td><td></td></tr> <tr><td></td><td></td><td>○</td></tr> </table>	×	×	×		○				○	<table border="1" style="width: 100%; height: 100%; border-collapse: collapse;"> <tr><td>×</td><td>?</td><td>?</td></tr> <tr><td>?</td><td>○</td><td>?</td></tr> <tr><td>?</td><td></td><td>?</td></tr> </table>		×	?	?	?	○	?	?		?									
×																																							
×	○	○																																					
×																																							
×	×	×																																					
	○																																						
		○																																					
×	?	?																																					
?	○	?																																					
?		?																																					

Slide 10

**Algorithme de GloBo : couverture minimale**

- Problème NP-difficile ;
- Heuristique : choisir systématiquement le paquet qui couvre le plus d'éléments non encore couverts [Paschos, 1997].

Slide 11

### Complexité de GloBo

- Construction d'un paquet maximalelement consistant :  $|E| - 1$  calculs d'agrégation et on fait un test de la propriété  $P$  pour chacune de ces agrégations.
- Couverture minimale :  $|E|^2$  [Paschos, 1997].
- GloBo supervisé :  $\Theta((|E^+| + |E^-|)^2)$  en calculs de moindres généralisés, et  $\Theta((|E^+| + |E^-|)^3)$  en tests de subsomption.
- Conditions d'application :  $P(A \cup B) \Rightarrow P(A) \wedge P(B)$  et unicité du moindres généralisé.

Slide 12

### Probabilité de réussite du clustering (1)

- $n$  nombre de sous-concepts à apprendre,
- $s$  nombre moyen d'éléments de  $E$  dans un sous-concept,
- $\alpha$  la probabilité de compatibilité avec la graine,
- $b$  la taille minimale d'un cluster verrouillé (hors graine).

événement	probabilité
verrouiller le cluster	$\alpha^b$
échouer pour $s$ exemples	$(1 - \alpha^b)^s$
réussir pour $n$ sous-concepts	$[1 - (1 - \alpha^b)^s]^n$

Slide 13

### Probabilité de réussite du clustering (2)

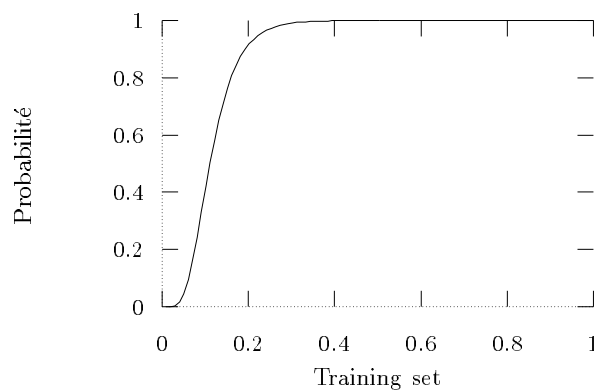
$$\left[1 - (1 - \alpha^b)^s\right]^n$$

- A posteriori, confiance dans la solution découverte.
- Notion de représentativité des exemples disponibles par rapport au concept à découvrir.
- No Free Lunch Theorem [Wolpert and Macready, 1995, Schaffer, 1994]: caractérisation des cas d'échec.

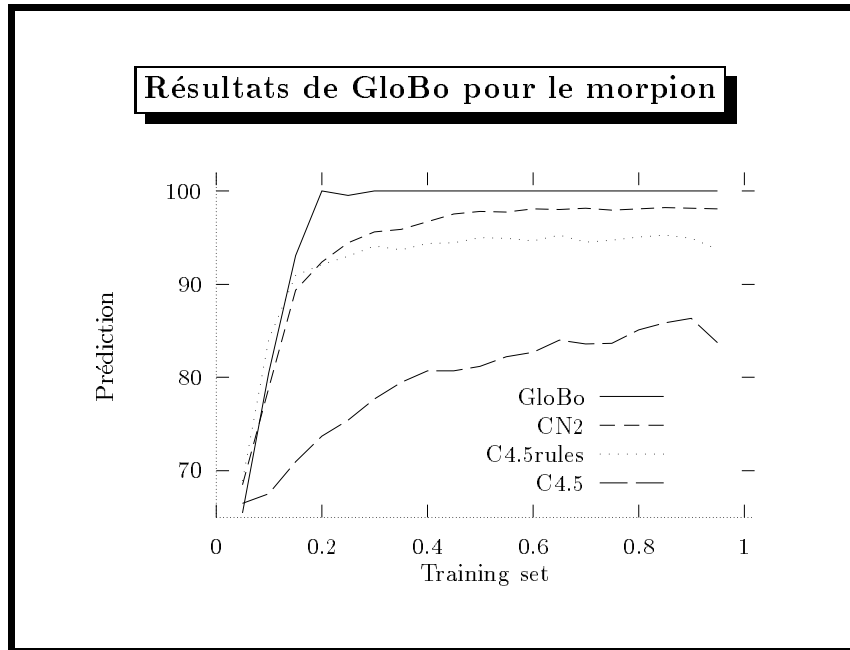
Slide 14

### Probabilité de succès de GloBo pour le morpion

$n = 8$ ,  $\alpha = \frac{1}{2}$ ,  $b = 2$  et  $s$  vaut 8% de la taille du dataset.



Slide 15



Slide 16

**Autres expérimentations**

Problème	Taille	Prédiction
Mushroom (avec négation)	3	100.0 %
Breast Cancer	4	93.0 %
Pima Indians Diabetes	24	68.8 %
Echocardiogram	3	88.2 %

PTE challenge [Srinivasan et al., 1997]: GloBo est premier ou deuxième, selon le critère considéré [Srinivasan et al., 1999a, Srinivasan et al., 1999b].

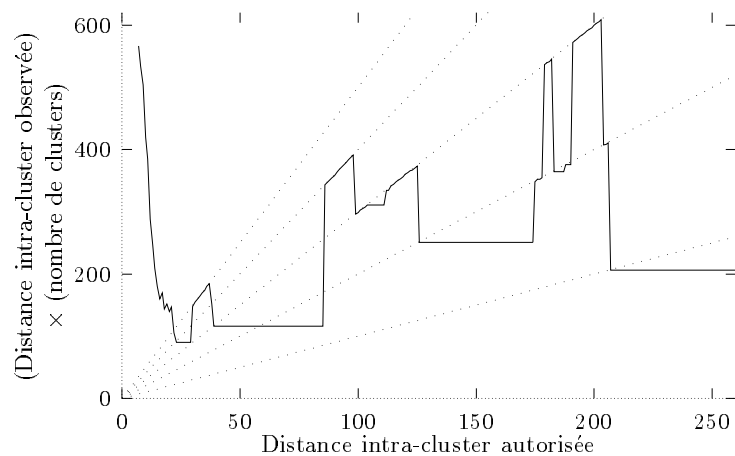


Slide 17

**Quadruped animals [Gennari et al., 1989]**

- Chaque instance appartient à l'une des quatre classes suivantes : chiens, chats, chevaux ou girafes.
- Chaque animal est décrit par 8 composantes : le cou, quatre jambes, le tronc, la tête, et la queue. Chacun de ces éléments est représenté par un cylindre, lui-même défini par 9 attributs.
- Utilisation de 100 exemples générés aléatoirement.
- Distance euclidienne.
- Minimiser à la fois le nombre de paquets et la distance intra-cluster : on observe le produit.

Slide 18

**Réglage empirique de la distance limite**

Slide 19

## Conclusion

Bilan :

- Algorithme pour le supervisé ou non ;
- Stochastique pour obtenir une complexité raisonnable ;
- Estimation du risque d'erreur de la procédure stochastique ;
- Moindre généralisé comme apprentissage conjonctif.

Perspectives :

- Données bruitées ;
- Autres langages.

Slide 20

## Références

- [Aha, 1991] Aha, D. (1991). Incremental constructive induction: An instance-based approach. In Proceedings of the Eighth International Workshop on Machine Learning, pages 117–121. Morgan Kaufmann.
- [Boström, 1995] Boström, H. (1995). Covering vs. divide-and-conquer for top-down induction of logic programs. In Proceedings of the 14th International Joint Conference on Artificial Intelligence, pages 1194–1200.
- [Gennari et al., 1989] Gennari, J. H., Langley, P., and Fisher, D. (1989). Models of incremental concept formation. Artificial Intelligence, 40:11–61.
- [Paschos, 1997] Paschos, V. T. (1997). A survey of approximately optimal solutions to some covering and packing problems. ACM Computing Surveys, 29(2):171–209.

## Slide 21

- [Schaffer, 1994] Schaffer, C. (1994). A conservation law for generalization performance. In Cohen, W. W. and Hirsh, H., editors, Proceedings 11th International Conference on Machine Learning, pages 259–265. Morgan Kaufmann.
- [Srinivasan et al., 1999a] Srinivasan, A., King, R., and Bristol, D. (1999a). An assessment of ILP-assisted models for toxicology and the PTE-3 experiment. In Džeroski, S. and Flach, P., editors, Proceedings of the 9th International Workshop on Inductive Logic Programming, volume 1634 of Lecture Notes in Artificial Intelligence. Springer-Verlag. À paraître.
- [Srinivasan et al., 1999b] Srinivasan, A., King, R., and Bristol, D. (1999b). An assessment of submissions made to the predictive toxicology evaluation challenge. In Proceedings of the 16th International Joint Conference on Artificial Intelligence. Morgan Kaufmann. À paraître.
- [Srinivasan et al., 1997] Srinivasan, A., King, R.,

## Slide 22

- Muggleton, S., and Sternberg, M. (1997). The predictive toxicology evaluation challenge. In Proceedings of the 15th International Joint Conference on Artificial Intelligence, pages 4–9. Morgan Kaufmann.
- [Wolpert and Macready, 1995] Wolpert, D. H. and Macready, W. G. (1995). No free lunch theorems for search. Technical Report SFI-TR-95-02-010, The Santa Fe Institute.