

# Apprentissage supervisé et Espace des Versions

Antoine Cornuéjols

CNAM-IEE  
et L.R.I., Université de Paris-Sud, Orsay (France)  
antoine@lri.fr  
http://www.iese.cnam.fr/~cornuejols/

## 0- Rappels

- Plusieurs types d'apprentissage
  - Supervisé
  - Non supervisé
  - Par renforcement
- Des choix
  - Objectif
  - Protocole
  - Critère de performance

### Aujourd'hui :

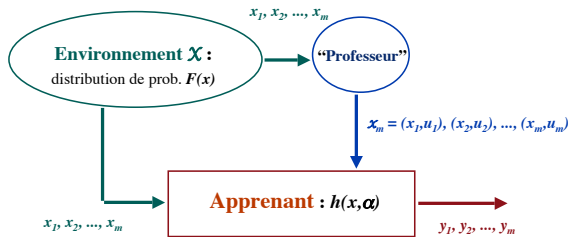
- Formalisation (un peu)
- Les grands principes et notions : **fondamental !!!**

## Cours 2 : L'apprentissage inductif supervisé

- I- Apprentissage supervisé de concept
- II- L'apprentissage comme jeu entre espaces
- III- Notions de base
- IV- Apprentissage et structuration de  $\mathcal{H}$
- V- Apprentissage et espace des versions

Voir les chapitres 1 & 4 de [Cornuéjols & Miclet « Apprentissage artificiel. Concepts et algorithmes ». Eyrolles, 2002]

## 1- Formulation du problème de l'induction



## 1- L'induction supervisée

- Si  $f$  est une fonction continue
  - Régression
  - Estimation de densité
- Si  $f$  est une fonction discrète
  - Classification
- Si  $f$  est une fonction binaire (booléenne)
  - Apprentissage de concept

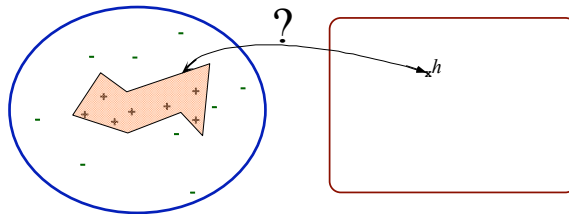
## 1- Induction supervisée de concept : exemple

	Aplati	Taille	Envergure	Couleur	Classe
Ex_1	VRAI	30	49	Roux	"+ (canard)"
Ex_2	FAUX	70	32	Gris	"- (manchot)"
Ex_3	VRAI	40	46	Orange	"+ (canard)"
Ex_4	FAUX	60	33	Orange	"- (manchot)"

- Induction d'une fonction indicatrice (à valeur dans  $\{-1,+1\}$ )

## 2- Apprendre = explorer un espace d'hypothèses

- Cas particulier de l'apprentissage de concepts



Espace des exemples :  $\mathcal{X}$

Espace des hypothèses :  $\mathcal{H}$

## 2- Trois questions fondamentales

### 1. Quel critère inductif ?

Quelle hypothèse devrait-on choisir étant donné l'échantillon d'apprentissage ?

### 2. Quel espace d'hypothèses ?

Quel espace d'hypothèses est approprié :

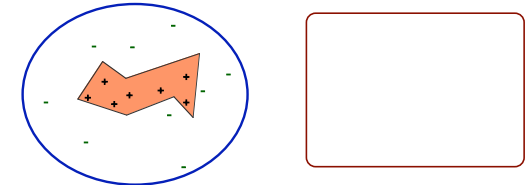
- Pour la tâche
- Pour que l'induction soit possible

### 3. Comment explorer l'espace des hypothèses ?

Résolution d'un problème d'optimisation :

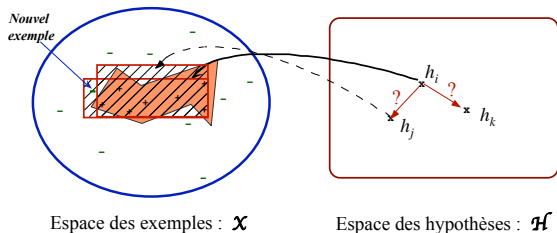
- recherche d'une bonne hypothèse dans un espace de possibilités afin de satisfaire le critère inductif

## 2- Apprendre = explorer un espace d'hypothèses



- Comment choisir une (des) hypothèse(s) ?
- Notion de biais de représentation

## 2- Apprendre = explorer un espace d'hypothèses



- Exploration de l'espace d'hypothèses
- Mesure de l'adéquation de l'hypothèse (critère de succès)

15/10/03

## 3- Réponses ... qualitatives

### 1. De quelle information doit-on disposer ?

- ➔ Compromis entre taille de l'échantillon d'apprentissage nécessaire ET "richesse" de l'espace d'hypothèses

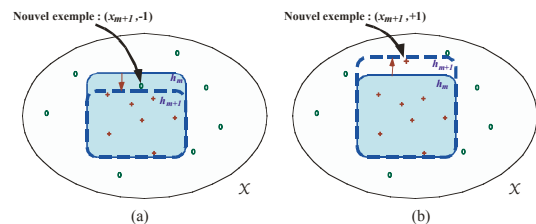
### 1'. Quel principe inductif ?

- ➔ Pas le principe "naïf" : minimiser l'erreur sur l'échantillon d'apprentissage en pariant que l'erreur sera également minimisée sur les exemples non vus
- ➔ Un **nouveau principe** : minimiser à la fois
  - ❑ l'erreur sur l'échantillon d'apprentissage
  - ❑ ET une mesure de la richesse de  $\mathcal{H}$

15/10/03

## 4- Induction et généralisation

### Comment corriger une hypothèse déficiente



15/10/03

## 2- Les interrogations fondamentales

### 1. Théorique : Sous quelles conditions est-il possible de résoudre le problème de l'induction ?

- ❑ De quelle information doit-on disposer ?
  - Dans les entrées (les exemples)
  - Dans l'espace d'hypothèse
- ❑ Quel principe inductif doit-on utiliser ?

### 2. Pratique : Comment explorer effectivement l'espace d'hypothèses ?

15/10/03

## 3- Réponses ... qualitatives

### 2. Comment explorer effectivement l'espace d'hypothèses $\mathcal{H}$ ?

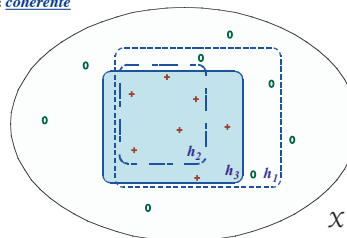
- ➔ Dépend de la structure dont on dispose sur  $\mathcal{H}$

- ❑ **Minimale** : mesure de performance et un voisinage
  - ➔ Gradient ± élaboré : Réseaux de neurones, recuit simulé, algorithmes génétiques
  - ➔ Méthodes de plus proches voisins ("Méthodes numériques")
- ❑ **Plus forte** : autres relations entre hypothèses (dans un langage)
  - Relations de généralité
  - Relations d'équivalence / à certains opérateurs
  - etc.
  - ➔ Méthodes d'exploration mieux dirigées ("Méthodes symboliques")

15/10/03

## 4- Couverture des exemples par une hypothèse

- $h_1$  : complète mais incorrecte
- $h_2$  : correcte mais incomplète
- $h_3$  : complète et correcte : cohérente



15/10/03

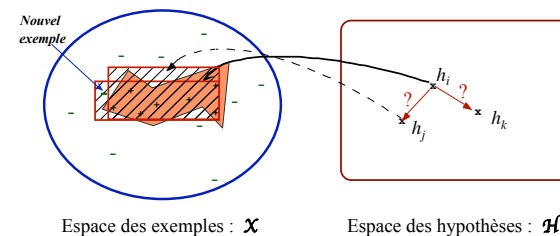
## 3- Notion de biais

Toute connaissance qui **restreint le champ des hypothèses** que l'apprenant doit considérer à un instant donné.

- **Biais de représentation**
  - ❑ On ne peut pas apprendre sans biais
  - ❑ Plus le biais est fort, plus l'apprentissage est facile
  - ➔ Bien choisir le biais
- **Biais de préférence**
  - ❑ Dû au contrôle de la recherche
  - ❑ Critère de choix entre hypothèse
    - Simplicité, complétude, intelligibilité, facilité d'évaluation, ...
  - ❑ Dû au protocole
    - Stratégie éducative (si apprentissage incrémental)

15/10/03

## 4- Apprendre = explorer un espace d'hypothèses

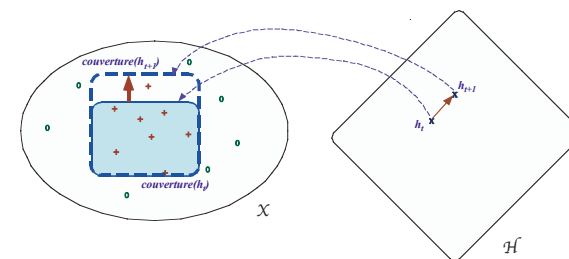


- Exploration de l'espace d'hypothèses
- ➔ guidée par les relations de généralités dans  $\mathcal{H}$

15/10/03

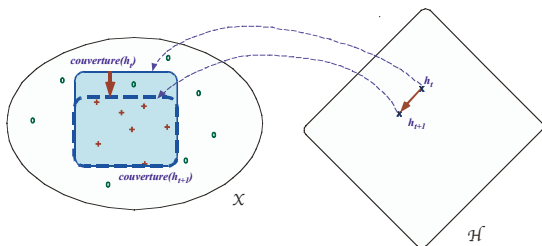
## 4- Relation d'inclusion et relation de généralité

### Vers la généralisation

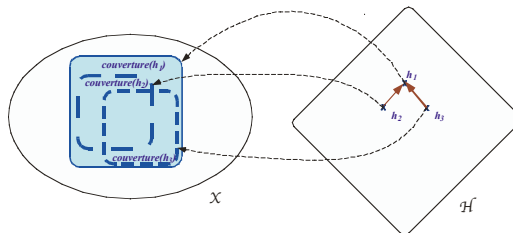


15/10/03

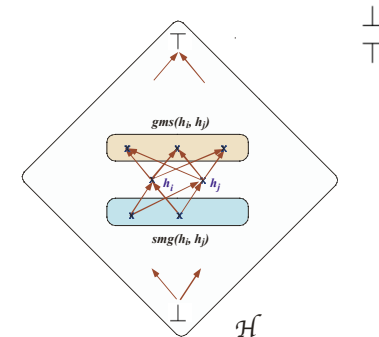
Vers la spécialisation



Relation de généralité dans H induite par la relation d'inclusion dans X



Ordre partiel dans H



4- Les opérateurs

Généralisation

- Transforme une description en une description plus générale

Spécialisation

- Duale de la généralisation
- (En général : produit une description qui est une conséquence logique de la description initiale)

Reformulation

- Transforme une description en une description logiquement équivalente

4- Opérateurs de généralisation

Règle d'abandon de conjonction

- $A \& B \rightarrow C \Rightarrow A \rightarrow C$   
ferrari & rouge  $\rightarrow$  coûteux  $\Rightarrow$  ferrari  $\rightarrow$  coûteux

Règle d'ajout d'alternative

- $A \rightarrow C \Rightarrow A \vee B \rightarrow C$   
ferrari  $\rightarrow$  coûteux  $\Rightarrow$  ferrari  $\vee$  rouge  $\rightarrow$  coûteux

Règle d'extension du domaine de référence

- $A \& [B=R] \rightarrow C \Rightarrow A \& [B=R'] \rightarrow C$   
grand & [couleur = rouge]  $\rightarrow$  coûteux  
 $\Rightarrow$  grand & [couleur rouge  $\vee$  bleu]  $\rightarrow$  coûteux

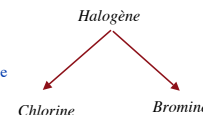
4- Opérateurs de généralisation

Règle de clôture d'intervalle

- $A \& [B=v_1] \rightarrow C \& A \& [B=v_2] \rightarrow C \Rightarrow A \& [B=v_1 \dots v_2] \rightarrow C$   
grand & [coût = 100]  $\rightarrow$  à acheter && grand & [coût = 150]  $\rightarrow$  à acheter  
 $\Rightarrow$  grand & [coût = 100 .. 150]  $\rightarrow$  à acheter

Règle de l'ascension dans l'arbre de hiérarchie

- $A \& [B=n_1] \rightarrow C \&& A \& [B=n_2] \rightarrow C \Rightarrow A \& [B=N] \rightarrow C$   
corrosif & [élément = chlorure]  $\rightarrow$  toxique  
corrosif & [élément = bromure]  $\rightarrow$  toxique  
 $\Rightarrow$  corrosif & [élément = halogène]  $\rightarrow$  toxique



4- Opérateurs de généralisation

Règle de variabilisation

- $F(a) \& F(b) \& \dots \rightarrow C \Rightarrow \exists v, F(v) \rightarrow C$   
grand(sommet\_objet) & grand(fond\_objet) & ...  $\rightarrow C$   
 $\Rightarrow \exists$  partie, grand(partie)  $\rightarrow C$

Règle de changement de conjonction en disjonction

- $A \& B \rightarrow C \Rightarrow A \vee B \rightarrow C$   
grand & rouge  $\rightarrow$  coûteux  $\Rightarrow$  grand  $\vee$  rouge  $\rightarrow$  coûteux

Règle d'extension du domaine de quantification

- $\forall v, F(v) \rightarrow C \Rightarrow \exists v, F(v) \rightarrow C$   
 $\forall$  partie, grand(partie)  $\rightarrow C \Rightarrow \exists$  partie, grand(partie)  $\rightarrow C$

4- Opérateurs de généralisation

Inversion de la résolution

- $A \& B \rightarrow C \&& \neg A \& D \rightarrow C \Rightarrow B \vee D \rightarrow C$   
vieux & grand  $\rightarrow C \&& \neg$  vieux & rouge  $\rightarrow C$   
 $\Rightarrow$  grand  $\vee$  rouge  $\rightarrow C$

Règle anti-extension

- $A \& [B=v_1] \rightarrow C \&& D \& [B=v_2] \rightarrow \neg C \Rightarrow [B \neq v_2] \rightarrow C$

Règle constructive de généralisation (modifiant les descripteurs)

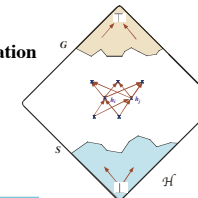
- $A \& B \rightarrow C \&& D \rightarrow C \Rightarrow A \& D \rightarrow C$

5- Représentation de l'espace des versions

Observation fondamentale :

L'espace des versions structuré par une relation d'ordre partiel peut être représenté par :

- sa borne supérieure : le **G-set**
- sa borne inférieure : le **S-set**



- G-set** = Ensemble de toutes les hypothèses les plus **générales** cohérentes avec les exemples connus
- S-set** = Ensemble de toutes les hypothèses les plus **spécifiques** cohérentes avec les exemples connus

... par mise à jour de l'espace des versions

**Idée :**

- maintenir le **S-set**
- et le **G-set**
- après chaque nouvel exemple

→ **Algorithme d'élimination des candidats**

**Initialiser S et G par (resp.) :**

- l'ensemble des hypothèses les plus *spécifiques* (les plus *générales*) cohérentes avec le 1er exemple positif connu.

**Pour chaque nouvel exemple (positif ou négatif)**

- mettre à jour S
- mettre à jour G

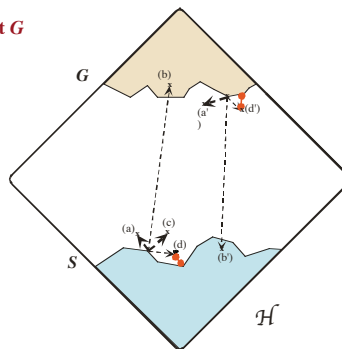
**Jusqu'à convergence**

ou jusqu'à ce que  $S = G = \emptyset$

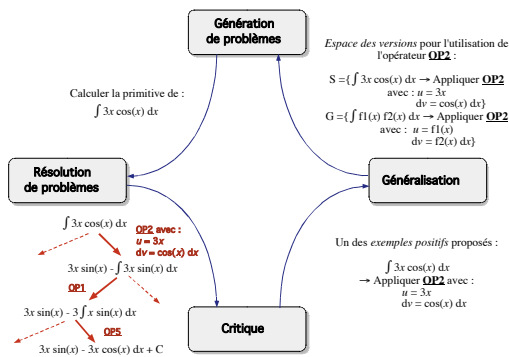
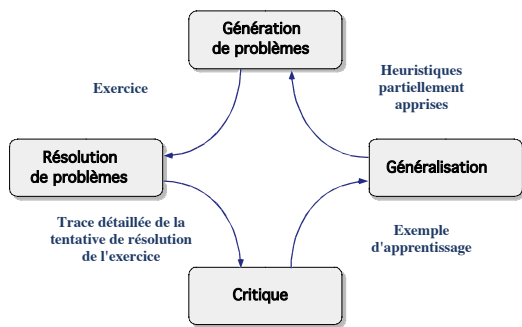
- $x_i$  est **néгатif**
  - **Éliminer** les hypothèses de S couvrant (indûment)  $x_i$
- $x_i$  est **positif**
  - **Généraliser** les hypothèses de S ne couvrant pas  $x_i$ , juste assez pour qu'elles le couvrent
  - Puis **éliminer** les hypothèses de S
    - couvrant un ou plusieurs exemples négatifs
    - plus générales que des hypothèses de S

- $x_i$  est **positif**
  - **Éliminer** les hypothèses de G ne couvrant pas  $x_i$
- $x_i$  est **néгатif**
  - **Spécialiser** les hypothèses de G couvrant  $x_i$ , juste assez pour qu'elles ne le couvrent plus
  - Puis **éliminer** les hypothèses de G
    - n'étant pas plus générales qu'au moins un élément de S
    - plus spécifiques qu'au moins une autre hypothèse de G

**Mise à jour des bornes S et G**



- **Incrementalité**
- **Complexité ?**
- **Utilisation si non convergence ?**
- **Que signifie  $S = G = \emptyset$  ?**
- **Possibilité d'"apprentissage actif" ?**
- **Que faire si les données sont bruitées ?**



- C'est surtout l'**induction supervisée** qui est étudiée
- Jeu entre *espace des exemples* et *espace des hypothèses*
- **On ne peut apprendre sans biais**
- **La réalisation de l'apprentissage dépend de la structuration de l'espace des hypothèses**
  - sans structure : méthodes par *interpolation*
  - notion de distance : méthodes par *gradient (approximation)*
  - relation d'ordre partiel : *exploration guidée (exploration)*