

COURS SUR

L'APPRENTISSAGE ARTIFICIEL

COURS MASTER IFI 2010/2011



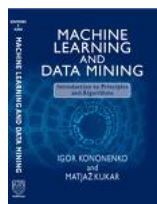
JEAN-DANIEL ZUCKER

DR À L'IRD UR GEODES
(MODÉLISATION MATHÉMATIQUES ET INFORMATIQUES DES SYSTÈMES COMPLEXES)
UMMISCO UMI 209



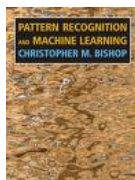
Administratif: 1/2 Module Apprentissage (18ECTS)²

- **Séance 1: Jeudi 25 Novembre** – INTRO GÉNÉRALE
 - Introduction, principe inductif, historique, formulation
 - Quelques mots sur l'apprentissage statistique
 - Espace des versions et algorithme
- **Séance 2: Lundi 6 Décembre** – APPRENTISSAGE SUPERVISÉ
- **Séance 3: Mardi 7 Décembre** – APPRENTISSAGE NON-SUPERVISÉ
- **Séance 4: Mardi 11 Janvier 2011** – ALGORITHMES ÉVOLUTIONNAIRES
- **Séance 5: Jeudi 14 Janvier 2011** – ALGORITHMES PAR RENFORCEMENT
- **Séance 6: Lundi 17 Janvier 2011** – MINI-PROJET



COURS N° 1

Principes généraux de l'INDUCTION



Cours 1: plan

1- Introduction à l'apprentissage et à l'induction

1.1- Visages de l'apprentissage

1.2- Des algorithmes

1.3- Apprentissage par exploration

2. Algorithme d'élimination des candidats



L'apprentissage pour une machine/un programme

5

Votre définition:

Définition

6

• Qu'est-ce que l'apprentissage pour une machine?

- “*Changement(s) dans un système lui permettant d'améliorer ses performances sur une tâche déjà vue ou similaire*” (Simon,83)

• Qu'est-ce qu'une tâche ? Quelles applications ?

- *Classification, Organisation de connaissances*
- *Résolution de problème, planification et action*

• Quels changements ?

- *Acquérir* de nouvelles connaissances
- *Réviser* ses connaissances (son comportement)
- *Optimiser* une fonction coût

• Comment mesurer les performances ?

- *Taux* de classification (tests), *cohérence*, *qualité* des solutions (concises, interprétables, ...), ...

1.1 Applications : Grand DARPA challenge (2005)

7



1.1 Applications : Grand DARPA challenge (2005)

8



1.1 Applications : Grand DARPA challenge (2005)

9

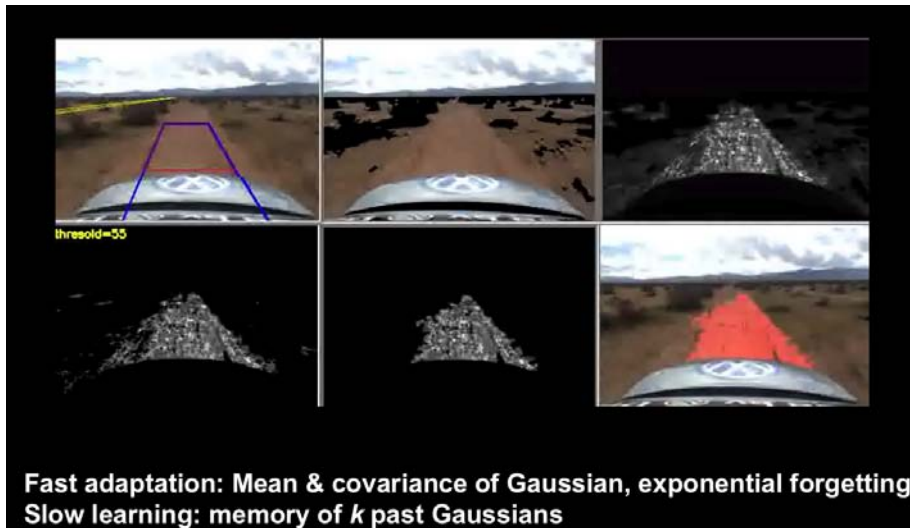


COURS APPRENTISSAGE N°1 Jean-Daniel ZUCKER

IFI 2011

1.1 Applications : Grand DARPA challenge (2005)

10

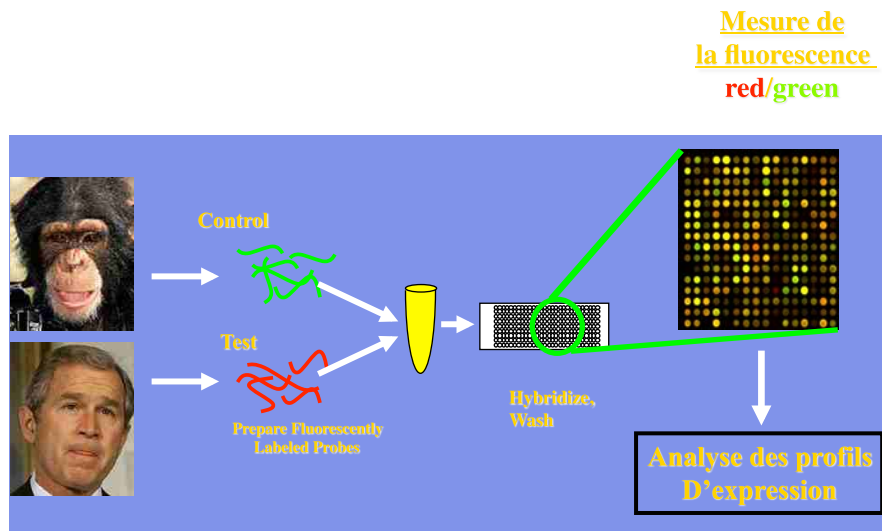


Fast adaptation: Mean & covariance of Gaussian, exponential forgetting
Slow learning: memory of k past Gaussians

COURS APPRENTISSAGE N°1 Jean-Daniel ZUCKER

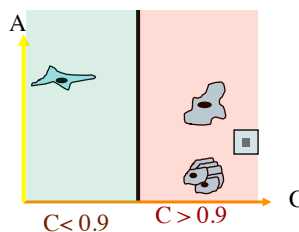
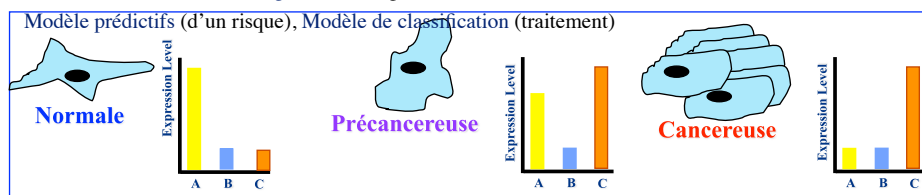
IFI 2011

1.1 Applications : Analyse de puces à ADN



1.1 Applications : Analyse de puces à ADN

- **Problème:** Construire des modèles pour la prédiction et classification à partir de données d'expression.
- **Utilisation :** Identification de gènes cibles (prédicteurs),



1.1 Applications : Apprendre à étiqueter des images

- Reconnaissance de visages

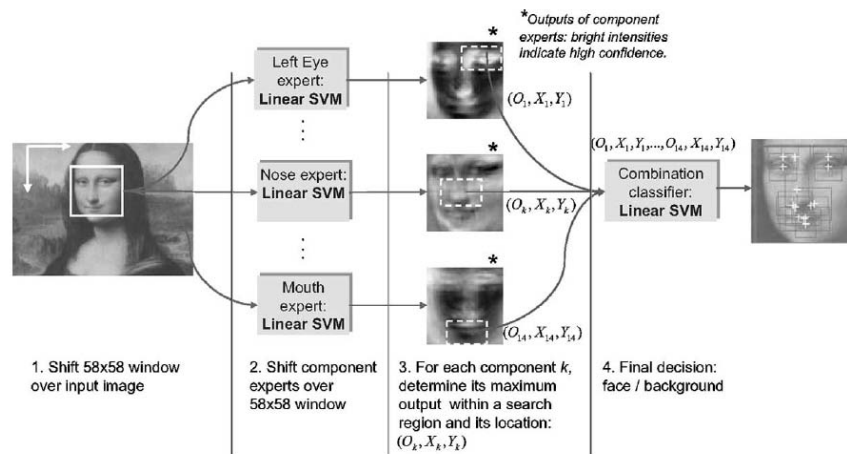
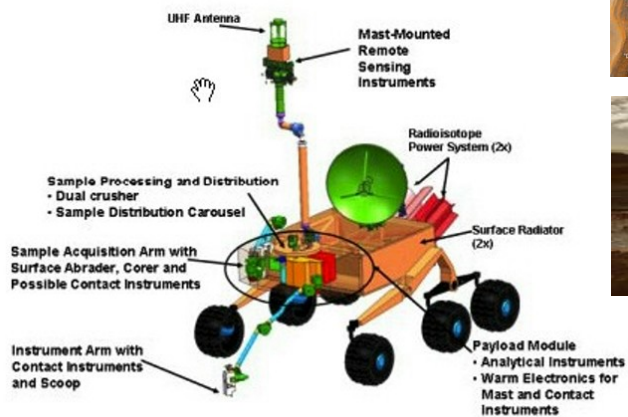


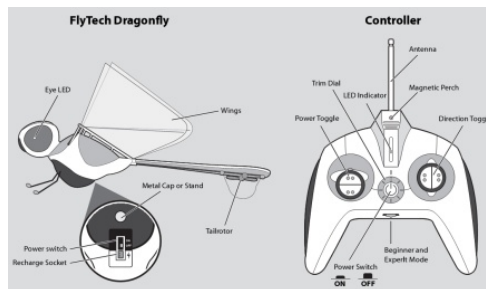
Fig. 4. System overview of the component-based face detector using four components.

1.1 Applications : Robot sur Mars



1.1 Applications : Robots

- Robot wowwee



1.1- Autres apprentissages

- Apprentissage pour la navigation

- Apprentissage de trajets (fourmis, abeilles)
 - Robots



- Discrimination

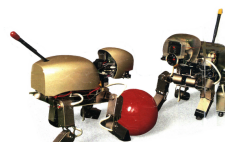
- Identification de sous-marins vs. bruits naturels
 - Identification de locuteur / de signature
 - Reconnaissance de l'écriture manuscrite, de la parole
 - Code postal



- Catégorisation SkyCat

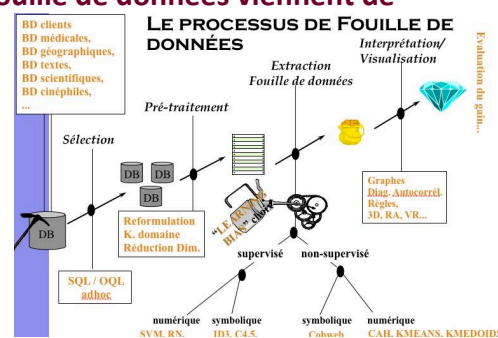
1.1- Autres apprentissages

- Révision de théorie
- Découverte scientifique
 - Découverte de régularités (en biochimie, ...)
- Apprendre à filtrer l'information
- Apprendre les préférences d'un utilisateur
- Apprendre à faire des résumés
- Apprendre à communiquer (e.g. Steels, ...)
- ...



1.1- Fouille de données & Analyse de sensibilité

- La plupart des algorithmes de Fouille de données viennent de l'apprentissage artificielle ...



- Pour analyser les données issues de simulateurs (comme GAMA !)

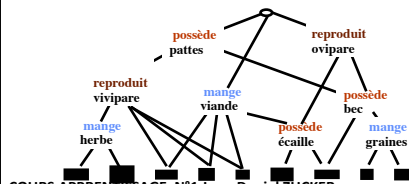
1.1. Pour résumer

Classification, Organisation

Diagnostiquer présence du SARS
 Accorder un prêt
 Rejeter des pièces sur une chaîne de prod.
 Quelle simulation minimise le nombre d'infectées



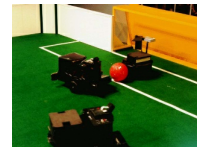
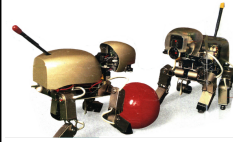
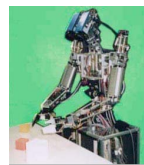
Poulet grippé Poulet sain



COURS APPRENTISSAGE N°1 Jean-Daniel ZUCKER

Résolution de problèmes, planification et action

Améliorer une heuristique
 Conduire une voiture
 Jouer au football



IFI 2011

Définition

- Qu'est-ce que l'apprentissage pour une machine?
 - "Changement(s) dans un système lui permettant d'améliorer ses performances sur une tâche déjà vue ou similaire" (Simon,83)
- Qu'est-ce qu'une tâche ?
 - Classification, Organisation de connaissances
 - Résolution de problème, planification et action
- Quels changements ?
 - Acquérir de nouvelles connaissances
 - Réviser ses connaissances (son comportement)
- Comment mesurer les performances ?
 - Taux de classification (tests), cohérence, qualité des solutions (concises, interprétables, ...), ...

COURS APPRENTISSAGE N°1 Jean-Daniel ZUCKER

IFI 2011

I.1. Quels “changements” pour apprendre?

Il y a deux manières principales pour améliorer les performances d'un système:

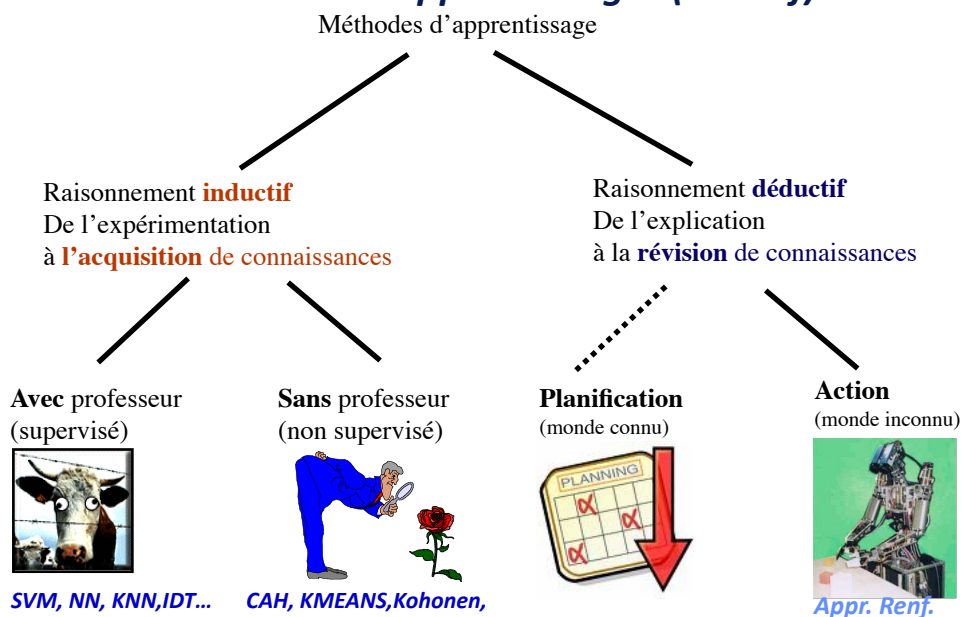
①- Acquérir de nouvelles connaissances

- ❑ *Acquérir de nouveaux faits (connaissances déclaratives)*
- ❑ *Acquérir de nouvelles capacités (connaissances procédurales)*

②- Adapter son comportement

- ❑ *résoudre les problèmes plus correctement,*
- ❑ *résoudre les problèmes plus efficacement, ...*

I.1. Une taxinomie des apprentissages (intuitif)



1.6. Différents types d'apprentissage

- Apprentissage par **renforcement** (RL)
- Apprentissage **symbolique** (Concept Learning)
- Apprentissage à partir de **cas** (CBR) ; K-PPV
- Algorithmes **génétiques** (Genetic Algorithm)
- Construction de **classifications** (Clustering)
- Apprentissage de programmes (prolog (ILP); PG)
- Apprentissage **neuronal** (NN)
- Apprentissage **probabiliste**, par cœur, par analogie
- Apprentissage par **explications** (EBL)
- Approches hybrides,

1. 1- Autres objectifs pour l'apprentissage artificiel ? (5)

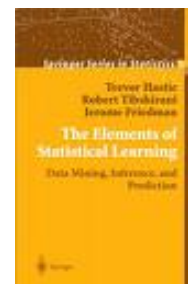
Mieux comprendre l'apprentissage humain :

- **Pour ne pas avoir à programmer**
 - Programmation par la démonstration
 - Programmation par l'exemple (e.g. l'EBL)
 - Programmation par échantillon d'apprentissage : induction
- **Pour mieux enseigner**
- **Pour savoir ce que d'autres intelligences pourraient apprendre :**
théorie générale de l'apprentissage

1.1- Qu'est-ce que la science de l'apprentissage artificiel

On étudie les apprentissages :

- naturels
- artificiels



→ des **théories** → **Apprentissage Statistique**

(s'appliquant à tout système apprenant)

→ des **méthodes** et des **algorithmes d'apprentissage**

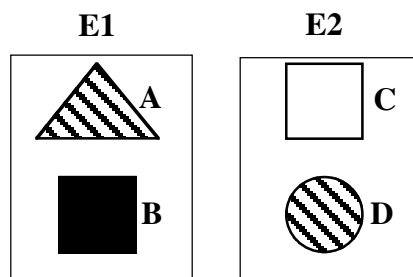
Implémentables sur machines

1. 1- Des scénarios

• **Induction**

- 1 2 3 5 ...
- 1 1 1 2 1 1 2 1 1 1 1 1 2 2 1 3 1 2 2 1 1 ...
- Comment ?
- Pourquoi serait-il possible de faire de l'induction ?
- Est-ce qu'un exemple supplémentaire doit augmenter la confiance dans la règle induite ?
- Combien faut-il d'exemples ?
- Quels scénarios ?

1. 1- Des scénarios



- Soient **deux exemples** dont les **descriptions** pourraient être :
 - *E1 : Un triangle rayé au-dessus d'un carré uni noir*
 - *E2 : Un carré uni blanc au-dessus d'un cercle rayé*
- ⇒ **Formuler une description générale de ces deux exemples**

1. 1- Des scénarios

Problème

Quel est le nombre a qui prolonge la séquence :

1 2 3 5 ... a ?

1. 1- Des scénarios

- **Solution(s).** Quelques réponses valides :

- $a = 6$. Argument : c'est la suite des entiers sauf 4.
- $a = 7$. Argument : c'est la suite des nombres premiers.
- $a = 8$. Argument : c'est la suite de Fibonacci
- $a = 2\pi$. (a peut être n'importe quel nombre réel supérieur ou égal à 5)
Argument : la séquence présentée est la liste ordonnée des racines du polynôme :

$$P = x^5 - (11 + a)x^4 + (41 + 11a)x^3 - (61 - 41a)x^2 + (30 + 61a)x - 30a$$

qui est le développement de : $(x - 1) \cdot (x - 2) \cdot (x - 3) \cdot (x - 5) \cdot (x - a)$

- **Généralisation**

Il est facile de démontrer ainsi que n'importe quel nombre est une prolongation correcte de n'importe quelle suite de nombre

*Mais alors ... comment faire de l'induction ?
et que peut-être une science de l'induction ?*

1. 1- Des scénarios

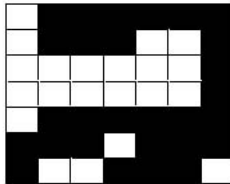
- **Exemples décrits par :**

- **nombre** (1 ou 2); **taille** (petit ou grand); **forme** (cercle ou carré); **couleur** (rouge ou vert)

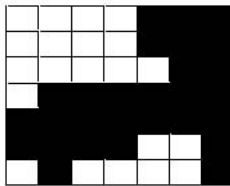
- **Les objets appartiennent soit à la classe + soit à la classe -**

| Description | Votre réponse | Vraie réponse |
|-------------------------|---------------|---------------|
| 1 grand carré rouge | | - |
| 1 grand carré vert | | + |
| 2 petits carrés rouges | | + |
| 2 grands cercles rouges | | - |
| 1 grand cercle vert | | + |
| 1 petit cercle rouge | | + |
| 1 petit carré vert | | - |
| 1 petit carré rouge | | + |
| 2 grands carrés verts | | + |

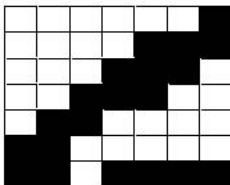
1. 1- Des scénarios



- Oui



- Oui



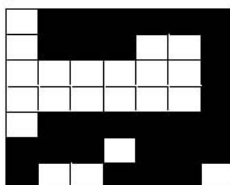
- Non

(de quel concept au fait ?)

IFI 2011

COURS APPRENTISSAGE N°1 Jean-Daniel ZUCKER

1. 1- Des scénarios



- Est-ce une tâche de reconnaissance de forme ? de caractères ?
- Comment coder les exemples ?

0111111011100100000010000001011111111101111001110

- Le choix de la représentation peut rendre l'apprentissage trivial

⇒ Mais comment faire ce choix ?

COURS APPRENTISSAGE N°1 Jean-Daniel ZUCKER

IFI 2011

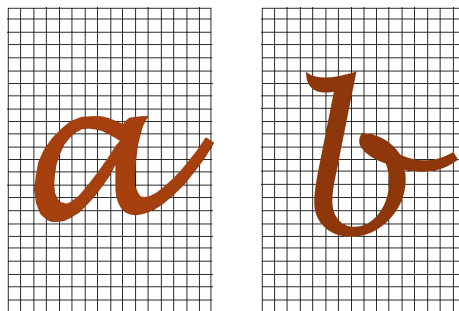
1. 1- Des scénarios

Apprendre par cœur ? IMPOSSIBLE !

➔ **Généraliser**

Apprentissage supervisé

Comment coder les formes ?



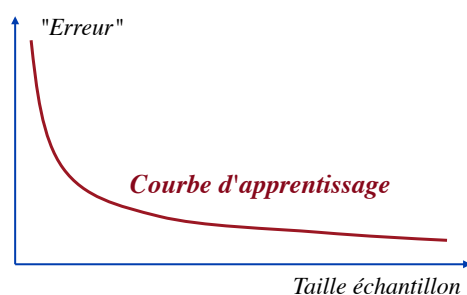
1. 1- Des scénarios

- **Extraction de caractéristiques (descripteurs, attributs)**

- ❑ **Éliminer** les descripteurs non pertinents
- ❑ **Introduction** de nouveaux descripteurs
 - Utilisation de connaissances a priori
 - Invariance par translation
 - Invariance par changement d'échelle
 - Histogrammes
 - Combinaisons de descripteurs
- ❑ **Ajouter** des descripteurs (beaucoup) !!

1. 1- Des scénarios

- **Quel critère de performance (de succès) ?**
 - Probabilité de misclassification
 - Risque
 - Nombre d'erreurs
- **Apprentissage sur un *échantillon d'apprentissage***
- **Test sur une *base de test***



Introduction à l'induction

- **Induction :**

Proposer des lois générales à partir de l'observation de cas particuliers

1.1- Types d'apprentissages

1. Apprentissage *supervisé*

À partir de l'*échantillon d'apprentissage* $S = \{(x_i, u_i)\}_{1,m}$

on cherche une loi de dépendance sous-jacente

- Par exemple une fonction h aussi proche possible de f (fonction cible)
tq : $u_i = f(x_i)$
- Ou bien une distribution de probabilités $P(x_i, u_i)$

afin de prédire l'avenir

1.1 - L'induction supervisée

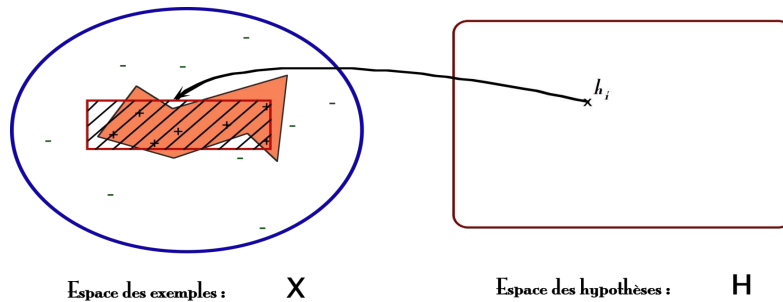
- Si f est une *fonction continue*
 - Régression
 - Estimation de densité
- Si f est une *fonction discrète*
 - Classification
- Si f est une *fonction binaire* (booléenne)
 - Apprentissage de concept

2. ESPACE DES VERSIONS

Rappel : apprentissage de concept

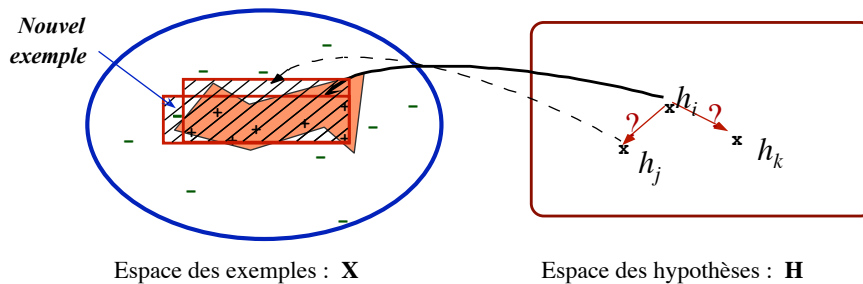
- **Un échantillon d'apprentissage** $S = \{(x_i, y_i)\}_{1..m}$
 - y_i : {appartient au concept, n'appartient pas au concept}
- **Objectif** : définir la fonction h (hypothèse) telle que $h(x_i)$ soit le plus proche possible de y_i

Apprentissage par exploration de l'espace des hypothèses



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

Apprentissage par exploration de l'espace des hypothèses

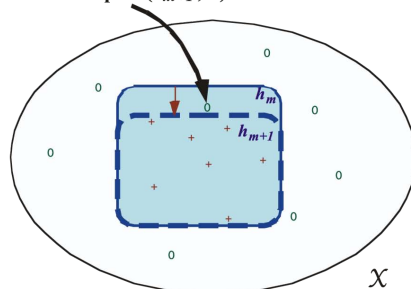


- ❖ Exploration de l'espace d'hypothèses
 - ◆ guidée par les relations de généralités dans H
- ❖ Mesure de l'adéquation de l'hypothèse

Apprentissage par exploration de l'espace des hypothèses

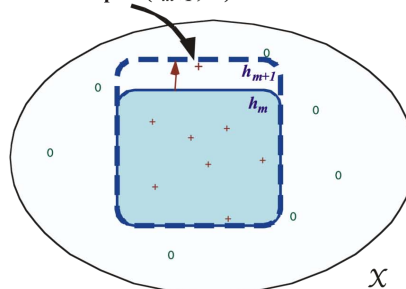
Comment corriger une hypothèse défectueuse ?

Nouvel exemple : $(x_{m+1}, -1)$



(a)

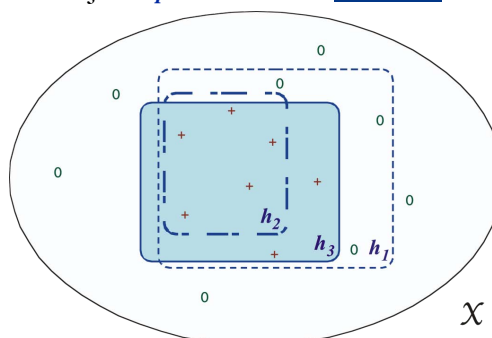
Nouvel exemple : $(x_{m+1}, +1)$



(b)

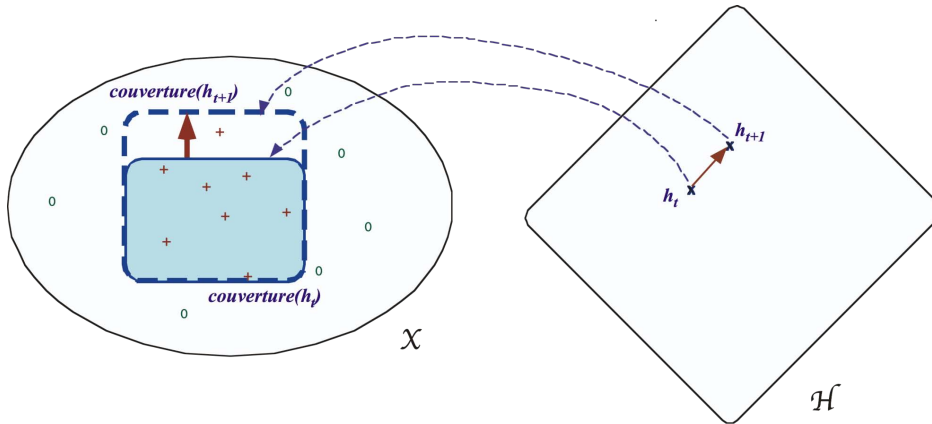
Couverture des exemples par une hypothèse

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



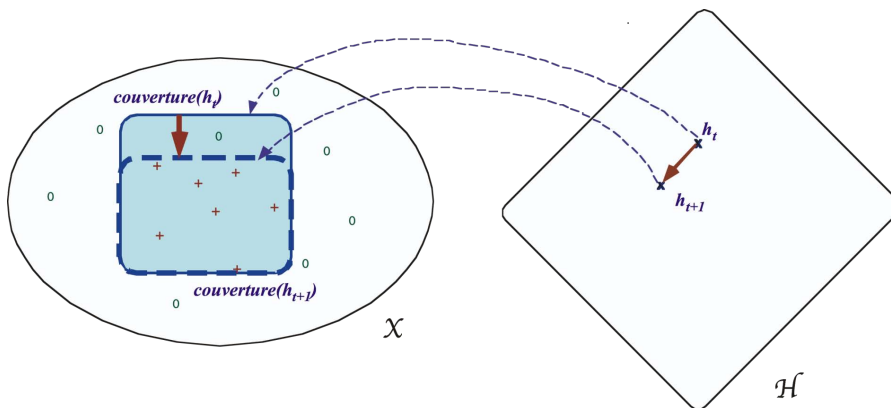
Relation d'inclusion et relation de généralité

Vers la généralisation

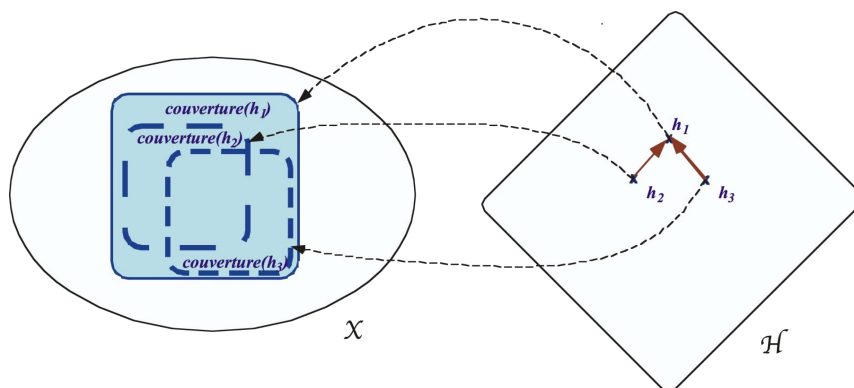


Relation d'inclusion et relation de généralité

Vers la spécialisation

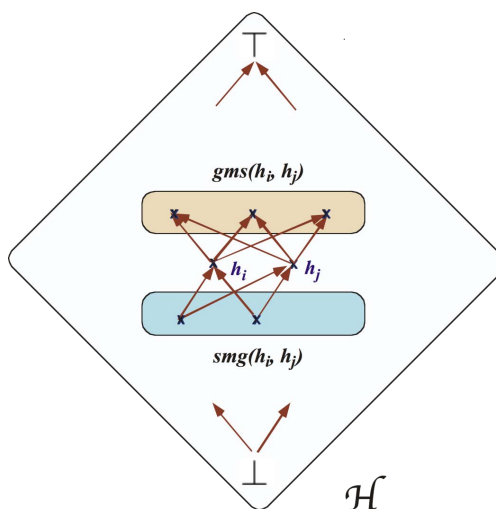


Relation d'inclusion et relation de généralité



Treillis de généralisation dans H

Ordre partiel dans H



Opérateurs

- **Généralisation**

- Transforme une description en une description plus générale

- **Spécialisation**

- Duale de la généralisation
- Transforme une description en une description moins générale

- **Reformulation**

- Transforme une description en une description **logiquement équivalente**

167 *Opérateurs de généralisation*

- **Règle d'abandon de conjonction**

- $A \wedge B \rightarrow C \Rightarrow A \rightarrow C$

ex : Ferrari \wedge rouge \rightarrow coûteux \Rightarrow Ferrari \rightarrow coûteux

- **Règle d'ajout d'alternative**

- $A \rightarrow C \Rightarrow A \vee B \rightarrow C$

ex : Ferrari \rightarrow coûteux \Rightarrow Ferrari \vee rouge \rightarrow coûteux

- **Règle d'extension du domaine de référence**

- $A \wedge [B = R] \rightarrow C \Rightarrow A \wedge [B = R'] \rightarrow C$

**ex : grand \wedge [couleur = rouge] \rightarrow coûteux
 \Rightarrow grand \wedge [couleur = rouge \vee bleu] \rightarrow coûteux**

Opérateurs de généralisation

- Règle de clôture d'intervalle

$$\square (A \wedge [B=v_1] \rightarrow C) \& (A \wedge [B=v_2] \rightarrow C) \Rightarrow A \wedge [B=v_1 \dots v_2] \rightarrow C$$

ex : (grand \wedge [coût = 100] \rightarrow à acheter)

& (grand \wedge [coût = 150] \rightarrow à acheter)

\Rightarrow grand \wedge [coût = 100 .. 150] \rightarrow à acheter

- Règle de l'ascension dans l'arbre de hiérarchie

$$\square (A \wedge [B=n_1] \rightarrow C) \& (A \wedge [B=n_2] \rightarrow C) \Rightarrow A \wedge [B=N] \rightarrow C$$

ex : (corrosif \wedge [élément = ^{Halogène}chlorine] \rightarrow toxique)

& (corrosif \wedge [élément = ^{Chlorine}bromine] \rightarrow toxique)

\Rightarrow corrosif \wedge [élément = ^{Chlorine}halogène] \rightarrow toxique

Opérateurs de généralisation

- Règle de variabilisation

$$\square F(a) \wedge F(b) \wedge \dots \rightarrow C \Rightarrow \exists v, F(v) \rightarrow C$$

ex : grand(sommet objet) \wedge grand(base objet) \wedge ... \rightarrow C

$\Rightarrow \exists$ partie, grand(partie) \rightarrow C

- Règle de changement de conjonction en disjonction

$$\square A \wedge B \rightarrow C \Rightarrow A \vee B \rightarrow C$$

ex : grand \wedge rouge \rightarrow coûteux \Rightarrow grand \vee rouge \rightarrow coûteux

- Règle d'extension du domaine de quantification

$$\square \forall v, F(v) \rightarrow C \Rightarrow \exists v, F(v) \rightarrow C$$

ex : \forall partie, grand(partie) \rightarrow C $\Rightarrow \exists$ partie, grand(partie) \rightarrow C

Opérateurs de généralisation

- **Inversion de la résolution**

- $(A \wedge B \rightarrow C) \& (\neg A \wedge D \rightarrow C) \Rightarrow B \vee D \rightarrow C$

- ex : $(\text{vieux} \wedge \text{grand} \rightarrow C) \& (\neg \text{vieux} \wedge \text{rouge} \rightarrow C)$
 $\Rightarrow \text{grand} \vee \text{rouge} \rightarrow C$

- **Règle anti-extension**

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

- **Règle constructive de généralisation (modifiant les descripteurs)**

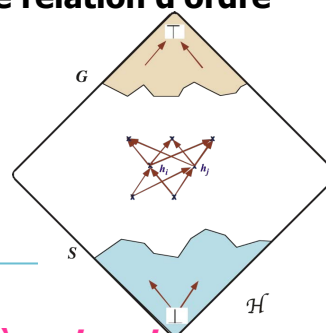
- $(A \wedge B \rightarrow C) \& (D \rightarrow C) \Rightarrow (A \wedge D \rightarrow C)$

Représentation de l'espace des versions

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

- sa borne supérieure : **G-set**

- sa borne inférieure : **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**

Apprentissage

- **Principe : mise à jour en continue de l'espace des versions**
- **Idée : maintenir le S-set et le G-set après chaque nouvel exemple**

⇒ **Algorithme d'élimination des candidats**

Algorithme d'élimination des candidats

- **Initialiser S et G par (respectivement) :**
 - l'ensemble des hypothèses les plus spécifiques (les plus générales) cohérentes avec le 1^{er} 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 $S = G = \emptyset$

Mise à jour de S

- x_i est négatif

- Eliminer les hypothèses de S couvrant 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

Mise à jour de G

- x_i est positif

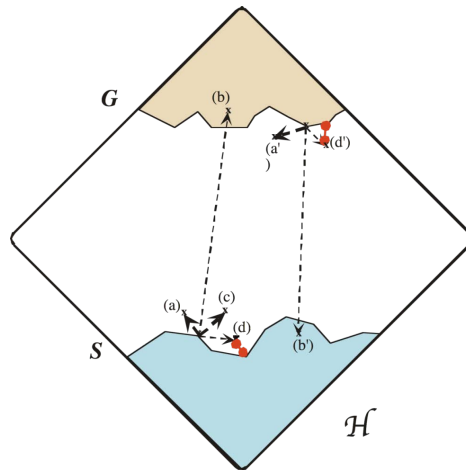
- Eliminer les hypothèses de G ne couvrant pas x_i

- x_i est négatif

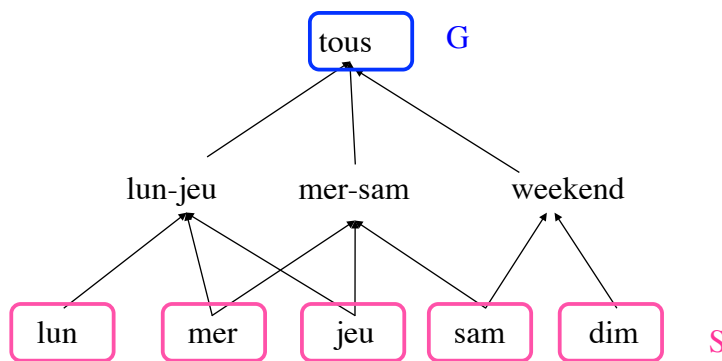
- Spécialiser les hypothèses de G couvrant x_i juste assez pour qu'elles ne le couvrent plus
- Puis éliminer les hypothèses de G
 - n'étant pas plus générales qu'au moins un élément de S

Algorithme d'élimination des candidats

Mise à jour des bornes S et G

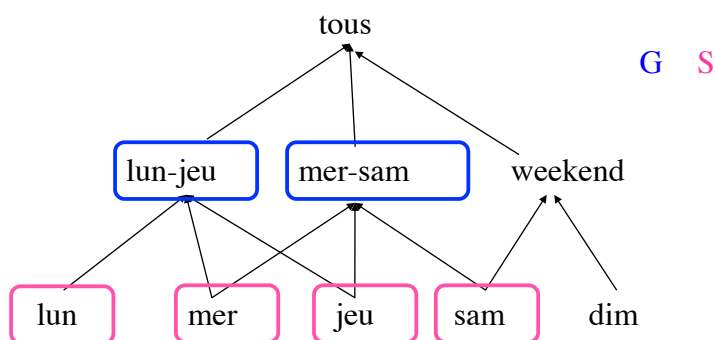


Exemple 1 (1/5)

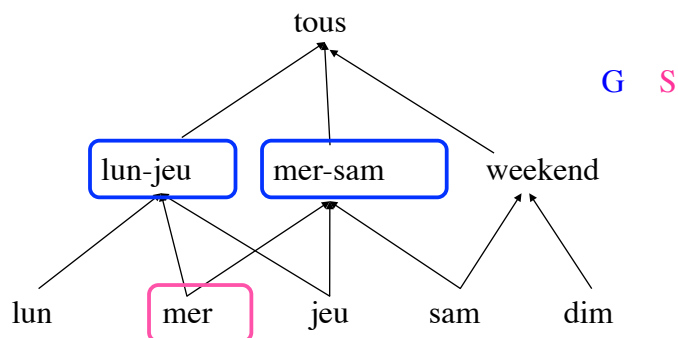


Exemple 1 (2/5)**Exemple négatif : (dim,-)**

- ❖ *dim* est enlevé de *S* (car couvre dim)
- ❖ *Tous* (dans *G*) est spécialisé en *lun-jeu* et *mer-sam*

**Exemple 1 (3/5)****Exemple positif : (mer,+)**

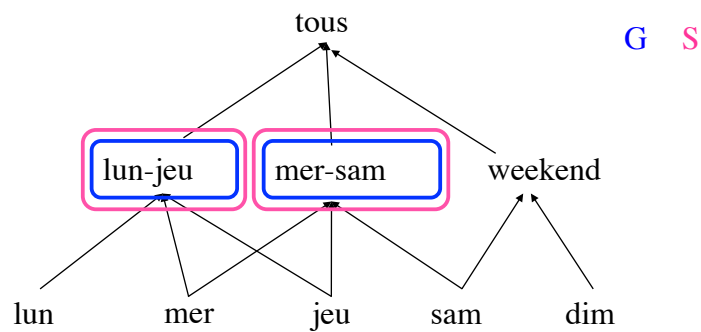
- ❖ *G* n'est pas modifié
- ❖ *Lun, jeu et sam* (dans *S*) sont généralisés :
 - *lun-jeu et mer-sam* : plus généraux que *mer* : Ces généralisations ne sont donc pas gardées dans *S*



Exemple 1 (4/5)

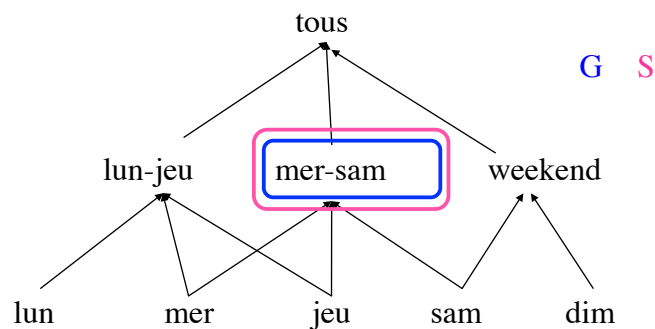
Exemple positif : (jeu,+)

- ❖ *G* n'est pas modifié
- ❖ *mer* (dans *S*) est généralisé :
 - *lun-jeu* et *mer-sam*

**Exemple 1 (5/5)**

Exemple négatif : (mar,-)

- ❖ *Lun-jeu* couvrant *mar* est enlevé de *S*
- ❖ *Lun-jeu* (dans *G*) est spécialisé :
 - Aucune des spécialisations possibles n'est plus générale qu'un élément de *S* : aucune n'est conservée dans *G*



Exemple (2)

objet : taille {grand petit}

forme-géométrique : {triangle, cercle}

couleur : {rouge, bleu}

$S = \{\emptyset\}$ $G = \{\emptyset\}$

E1 positif: (grand triangle rouge) & (petit cercle bleu)

E2 positif: (petit triangle rouge) & (grand cercle bleu)

E3 négatif: { (grand triangle bleu) (petit triangle bleu)}

Exemple 2 (suite)

Exemple positif: {(grand triangle rouge) (petit cercle bleu)}

$S1 = \{(grand\ triangle\ rouge)\ (petit\ cercle\ bleu)\}$

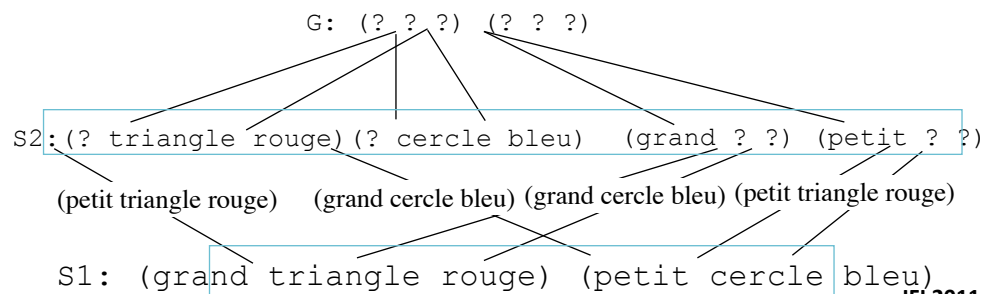
$G1 = \{(???)\ (???)\}$

Exemple positif: {(petit triangle rouge) (grand cercle bleu)}

$S2 : \{(?\ triangle\ rouge)\ (?\ cercle\ bleu),$

(grand ??) (petit ??)\}

$G2 = \{(???)\ (???)\}$

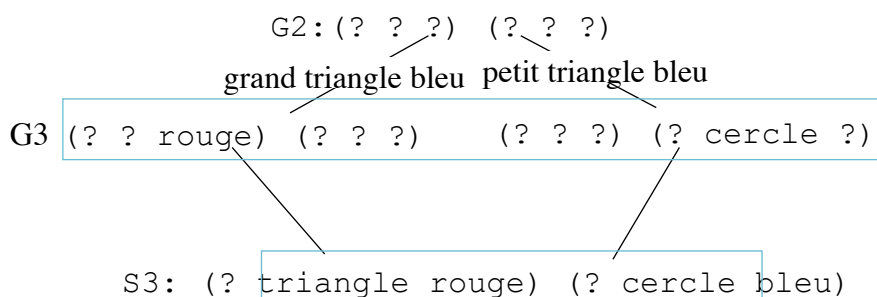


Exemple 2 (fin)

E3 négatif: { (grand triangle bleu) (petit triangle bleu)}

S3 = { (? triangle rouge) (? cercle bleu) }

G3 = {(? ? rouge) (? ? ?), (? ? ?) (? cercle ?) }



Exercices: calculer S et G (pour le 6 Décembre 2011)

Cas 1 (converge)

- E1 positif (grand rouge cercle)
- E2 négatif (petit rouge triangle)
- E3 positif (petit rouge cercle)
- E4 négatif (grand bleu cercle)

Cas 2 (ne converge pas)

- E1 positif (grand rouge cercle)
- E2 négatif (petit bleu triangle)
- E3 positif (petit rouge cercle)
- E4 négatif (moyen vert carré)

Cas 3 (langage insuffisant)

- E1 positif (grand rouge cercle)
- E2 négatif (grand bleu cercle)
- E4 positif (petit bleu cercle)

Exemple 3

Exercice : appliquer l'algorithme d'élimination des candidats pour calculer l'espace des versions de l'échantillon suivant :

| Bec aplati | Grand | Couleur sombre | Canard ? |
|------------|-------|----------------|----------|
| Oui | Non | Oui | Oui |
| Non | Oui | Non | Non |
| Oui | Non | Non | Oui |
| Oui | Oui | Oui | Non |
| Non | Non | Non | Non |

Ce qu'il faut retenir

- 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)**

1.5- Approche actuelle : les limites

• L'état de l'art actuel en apprentissage:

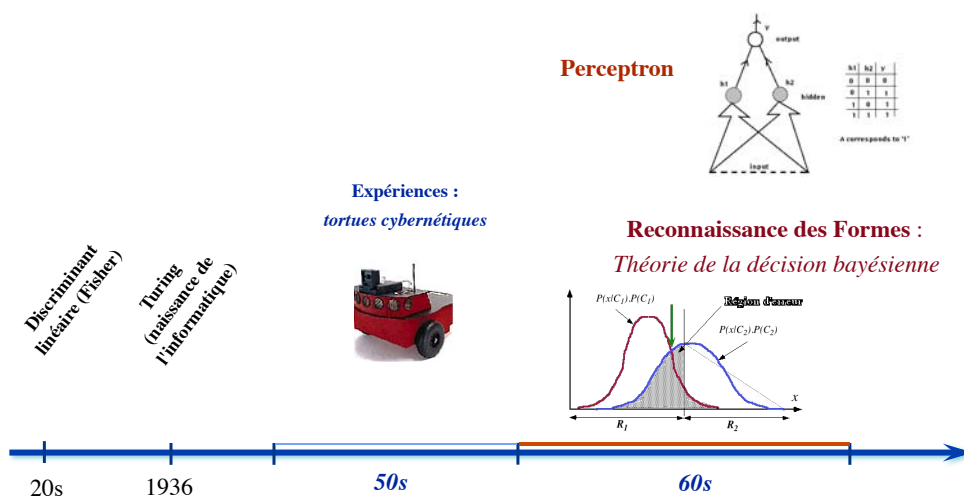
- Données i.i.d. (indépendant et identiquement distribué)
- Distribution statique
- Données étiquetées
- Classes approximativement équilibrées

• Versus e.g. les besoins de la robotique

- Données résultant :
 - De séquences
 - D'un apprentissage actif
- Contexte changeant
- Pauvrement étiquetées

} Données non i.i.d.

1.5- Perspective historique (1)



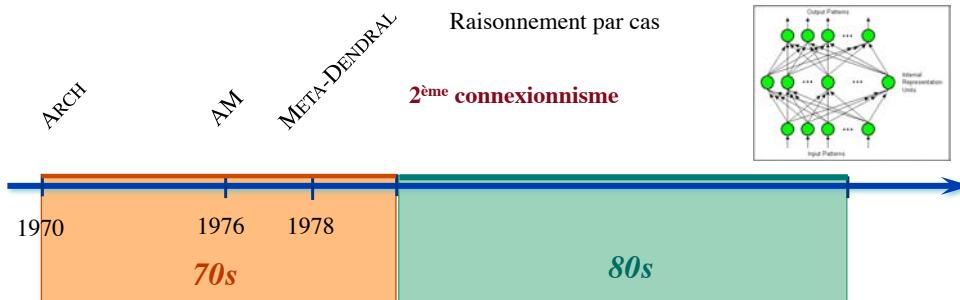
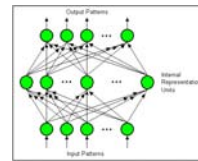
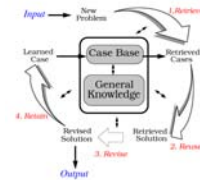
1.5- Perspective historique (2)

Apprentissage artificiel : une explosion

Systèmes dédiés à une tâche :
inspiration psychologique

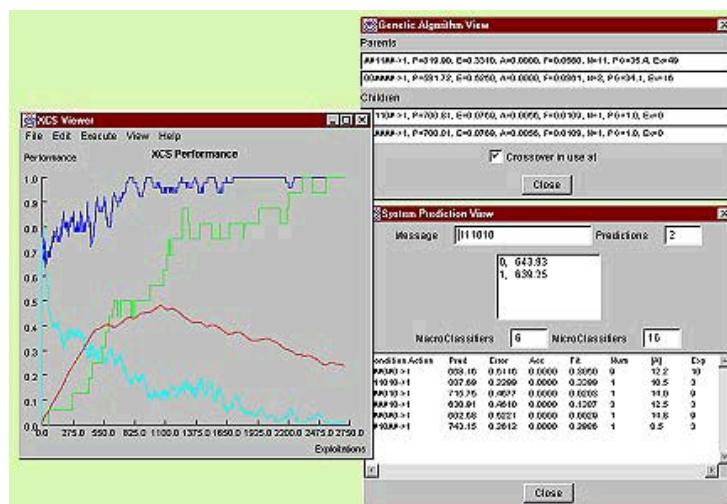
Induction supervisée

- Arbres de décision
- Algorithmes génétiques
- Explanation-Based Learning
- Raisonnement par cas

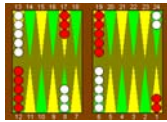


1.5- Perspective historique (2')

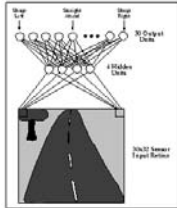
- Systèmes à usage industriels



1.5- Perspective historique (3)



Apprentissage artificiel :
une théorisation
et une mise à l'épreuve



Théorie de Vapnik

Nouvelles méthodes :

- SVMs
- Boosting

Data mining
Text mining



1.5- Perspective historique (4)

• Maintenant

- **Domination sans partage du paradigme dominant**
 - Apprentissage comme estimation / approximation de fonction
 - Données supposées tirées aléatoirement
 - Nouveau principe inductif : toujours prendre en compte l'espace d'hypothèses
- **Nouvelles techniques d'apprentissage issues de la théorie**
 - Séparateurs à Vastes Marges (SVM : *Support Vector Machines*)
 - Boosting
- **Prépondérance des applications de fouille dans les grandes bases de données**
 - Peu structurées
 - Données fournies en vrac
 - ➔ *Nouvelles mesures de performance*

1.5- Perspective historique : l'avenir ?

• Demain ... ?

□ Retour vers des problèmes à données plus structurées

- Exploration automatique de la toile (structure à tous les niveaux : grammatical, séquence, texte, discours, culture)

□ Nouveaux aspects

- Nouvelles demandes :
 - Systèmes à longue durée de vie
 - Aide à l'éducation
 - Apprentissage collectif
- Incrémentalité
- Transferts d'une tâche à une autre, d'un domaine à un autre, d'un agent à un autre

➔ Nouveaux problèmes

➔ Nouvelles techniques

1.5 - Où en est-on ?

| | |
|------------------|--|
| supervisé | Bien maîtrisé mais ... |
| non supervisé | Encore beaucoup à faire |
| d'efficacité | Travaux pionniers, ... |
| par renforcement | Encore beaucoup à faire Etendre le domaine |
| de K. complexes | Quasi rien |
| stades ... | Rien |

1.5- Questions ...

- *Peut-on apprendre n'importe quoi ?*
- *Peut-on apprendre à partir de rien (tabula rasae) ?*
- *Suffit-il d'avoir plus de neurones pour apprendre mieux ?*
- *Quel lien entre généralisation et abstraction ?*
- ...

5- ... et programmes de recherche

- Phénomènes de *transition de phase en induction*
- Vers une *science du dynamique* :
 - *Quels sont les systèmes dépendants de l'ordre des entrées ?*
 - *Pour ceux-là, quel est l'ordre optimal de présentation des données ?*
- *Apprentissage et ... oubli*
 - *L'oubli peut-il être utile ?*
- Y a-t-il des *passages obligés dans l'apprentissage de connaissances complexes ?*
 - *Ex : la notion d'impetus avant celle de force et d'inertie ?*