



# Contributions aux modèles de Markov cachés métaheuristiques d'apprentissage, nouveaux modèles et visualisation de dissimilarité

Sébastien Aupetit

Laboratoire d'Informatique de l'Université François-Rabelais de Tours  
Équipe Handicap et Nouvelles Technologies

30 novembre 2005

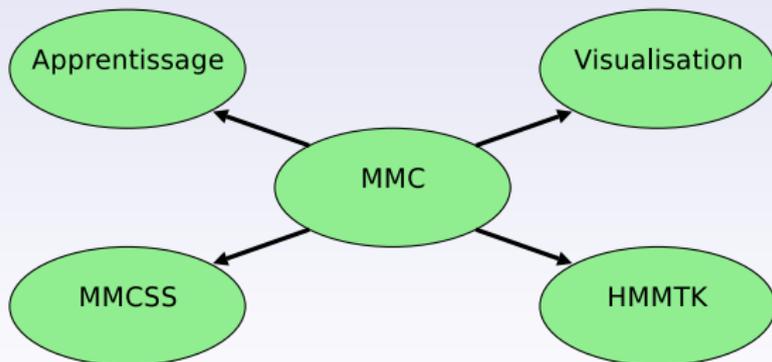
Jury :

M. Jin-Kao	HAO	Examineur	Professeur	Université d'Angers
M. Jacques	LÉVY VEHEL	Examineur	Directeur de recherche	INRIA
M. Nicolas	MONMARCHÉ	Examineur	Maître de conférences	Université François-Rabelais de Tours
M. Philippe	PREUX	Rapporteur	Professeur	Université Lille 3
M. Patrick	SIARRY	Rapporteur	Professeur	Université de Paris 12
M. Mohamed	SLIMANE	Examineur	Professeur	Université François-Rabelais de Tours
M. Pierre	LIARDET	Invité	Professeur	Université de Provence

## Les objectifs de la thèse

- ▶ Étendre les recherches sur l'apprentissage de MMC
- ▶ Évaluer l'intérêt des MMC à substitutions de symboles (MMCSS)
- ▶ Exploiter la visualisation de données pour mieux comprendre les MMC

Nouveaux algorithmes  
Comparaison des performances  
Différents espaces de solutions  
Reconnaissance d'images  
Segmentation d'images



# Plan

## Etat de l'art

- Les modèles de Markov cachés
- L'apprentissage de modèles de Markov cachés
- Fouille de données et visualisation d'informations

## Contributions aux modèles de Markov cachés

- Métaheuristiques pour l'apprentissage de MMC
- Les modèles de Markov cachés à substitutions de symboles
- Matrice de *scatterplots* pseudo-euclidienne
- HMMTK : bibliothèque logicielle de manipulation de MMC

## Conclusions et perspectives

# Plan

## Etat de l'art

- Les modèles de Markov cachés
- L'apprentissage de modèles de Markov cachés
- Fouille de données et visualisation d'informations

## Contributions aux modèles de Markov cachés

- Métaheuristiques pour l'apprentissage de MMC
- Les modèles de Markov cachés à substitutions de symboles
- Matrice de *scatterplots* pseudo-euclidienne
- HMMTK : bibliothèque logicielle de manipulation de MMC

## Conclusions et perspectives

# Plan

## Etat de l'art

### Les modèles de Markov cachés

L'apprentissage de modèles de Markov cachés

Fouille de données et visualisation d'informations

## Contributions aux modèles de Markov cachés

Métaheuristiques pour l'apprentissage de MMC

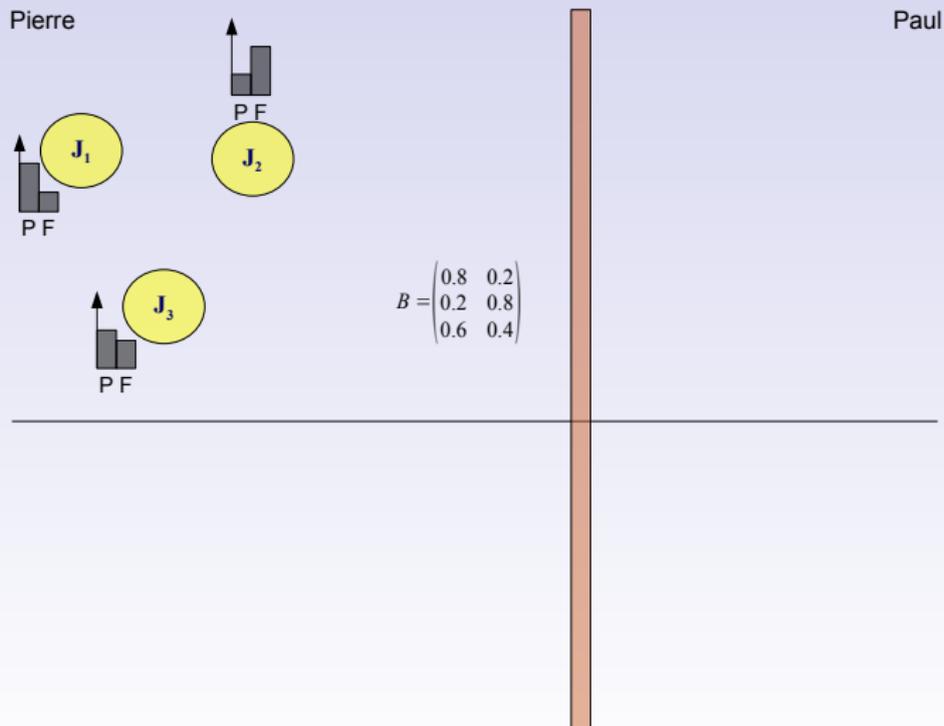
Les modèles de Markov cachés à substitutions de symboles

Matrice de *scatterplots* pseudo-euclidienne

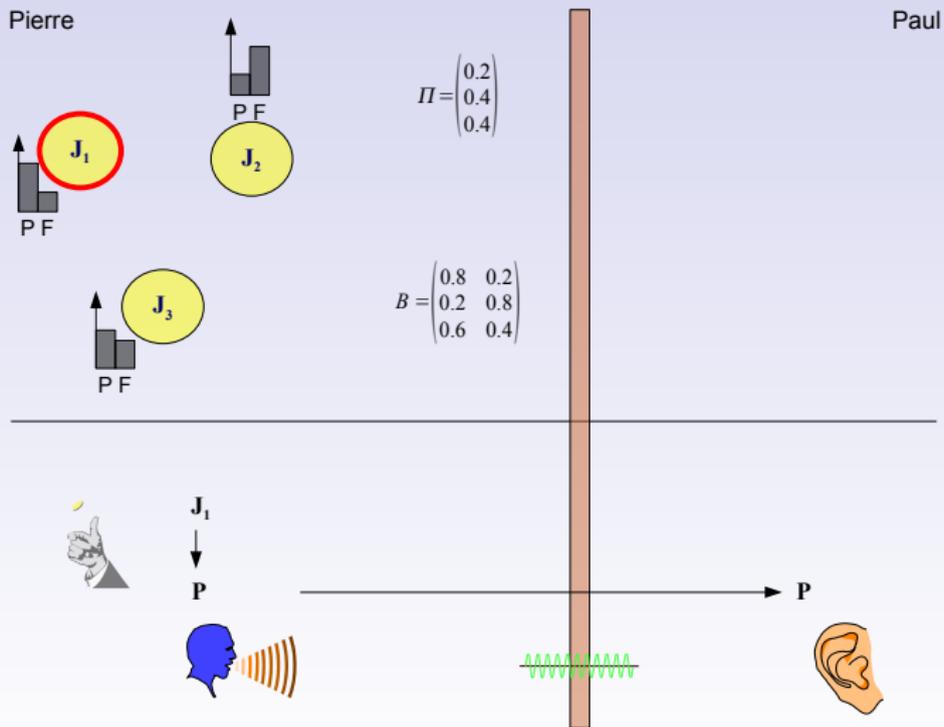
HMMTK : bibliothèque logicielle de manipulation de MMC

## Conclusions et perspectives

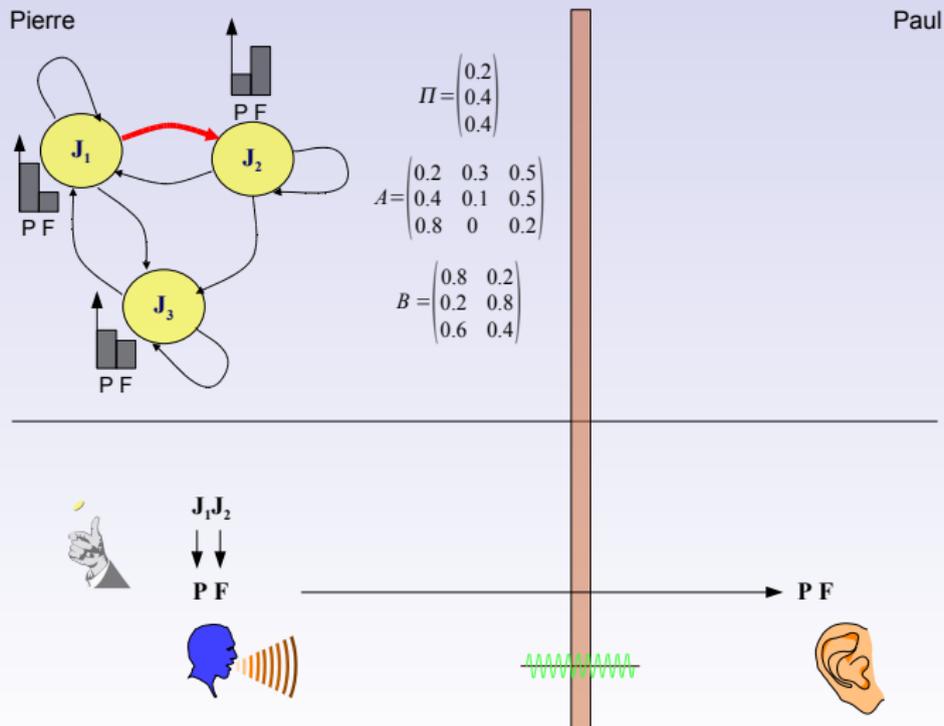
## Les modèles de Markov cachés (MMC) (1/2)



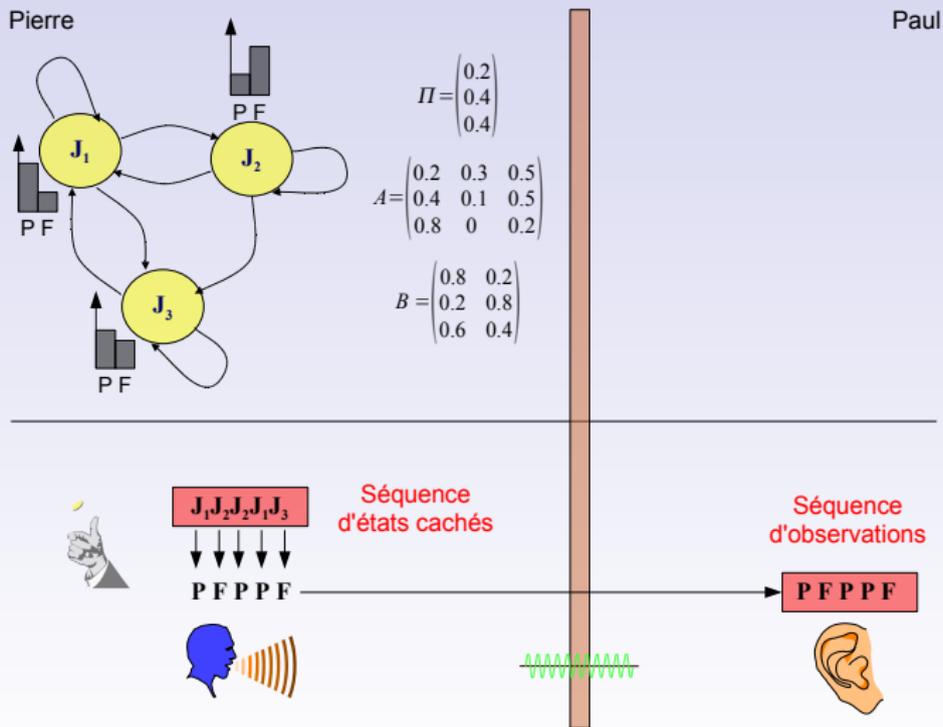
## Les modèles de Markov cachés (MMC) (1/2)



# Les modèles de Markov cachés (MMC) (1/2)



## Les modèles de Markov cachés (MMC) (1/2)



## Les modèles de Markov cachés (MMC) (2/2)

Pour utiliser les MMC, il faut résoudre plusieurs problèmes classiques :

## Les modèles de Markov cachés (MMC) (2/2)

Pour utiliser les MMC, il faut résoudre plusieurs problèmes classiques :

- ▶ Calcul de la vraisemblance :
  - ▶ évaluer l'adéquation entre une observation  $O$  et un modèle  $\lambda$
  - ▶ **algorithme Forward / Backward**

## Les modèles de Markov cachés (MMC) (2/2)

Pour utiliser les MMC, il faut résoudre plusieurs problèmes classiques :

- ▶ Calcul de la vraisemblance :
  - ▶ évaluer l'adéquation entre une observation  $O$  et un modèle  $\lambda$
  - ▶ **algorithme Forward / Backward**
- ▶ Segmentation d'une séquence d'observations :
  - ▶ calculer la séquence cachée la plus probablement suivie
  - ▶ **algorithme de Viterbi**

## Les modèles de Markov cachés (MMC) (2/2)

Pour utiliser les MMC, il faut résoudre plusieurs problèmes classiques :

- ▶ Calcul de la vraisemblance :
  - ▶ évaluer l'adéquation entre une observation  $O$  et un modèle  $\lambda$
  - ▶ **algorithme Forward / Backward**
- ▶ Segmentation d'une séquence d'observations :
  - ▶ calculer la séquence cachée la plus probablement suivie
  - ▶ **algorithme de Viterbi**
- ▶ Ajuster les paramètres d'un modèle :
  - ▶ dépend du critère à maximiser

## Les critères d'apprentissage (1/2)

- Des heuristiques spécialisées pour plusieurs critères d'apprentissage de MMC :

Maximisation du critère	Algorithme
$P(V = O, S = Q/\lambda)$	Apprentissage étiqueté
$P(V = O/\lambda)$	Baum-Welch, Gradient
$\frac{P(V=O/\lambda)}{\prod_{k=1}^K P(V=O_k/\lambda)}$	Information mutuelle
$P(\lambda/V = O)$	Maximum <i>a posteriori</i>
$P(V = O, S = Q^*/\lambda)$	<i>Segmental k-means</i>

## Les critères d'apprentissage (1/2)

- Des heuristiques spécialisées pour plusieurs critères d'apprentissage de MMC :

Maximisation du critère	Algorithme
$P(V = O, S = Q/\lambda)$	Apprentissage étiqueté
$P(V = O/\lambda)$	Baum-Welch, Gradient
$\frac{P(V=O/\lambda)}{\prod_{k=1}^K P(V=O_k/\lambda)}$	Information mutuelle
$P(\lambda/V = O)$	Maximum <i>a posteriori</i>
$P(V = O, S = Q^*/\lambda)$	Segmental <i>k-means</i>

## Les critères d'apprentissage (1/2)

- ▶ Des heuristiques spécialisées pour plusieurs critères d'apprentissage de MMC :

Maximisation du critère	Algorithme
$P(V = O, S = Q/\lambda)$	Apprentissage étiqueté
$P(V = O/\lambda)$	Baum-Welch, Gradient
$\frac{P(V=O/\lambda)}{\prod_{k=1}^K P(V=O_k/\lambda)}$	Information mutuelle
$P(\lambda/V = O)$	Maximum <i>a posteriori</i>
$P(V = O, S = Q^*/\lambda)$	<i>Segmental k-means</i>

- ▶ Remarques :
  - ▶ ces heuristiques convergent vers un **optimum local du critère**
  - ▶ le résultat dépend du modèle initial

## Les critères d'apprentissage (2/2)

- ▶ Pour utiliser efficacement les MMC dans des systèmes d'intelligence artificielle :
  - ▶ les optima locaux obtenus par les heuristiques ne sont pas suffisants
  - ▶ on a besoin de trouver de meilleurs modèles par rapport au critère considéré

## Les critères d'apprentissage (2/2)

- ▶ Pour utiliser efficacement les MMC dans des systèmes d'intelligence artificielle :
  - ▶ les optima locaux obtenus par les heuristiques ne sont pas suffisants
  - ▶ on a besoin de trouver de meilleurs modèles par rapport au critère considéré
- ▶ Une solution : **les métaheuristiques**

# Plan

## Etat de l'art

Les modèles de Markov cachés

**L'apprentissage de modèles de Markov cachés**

Fouille de données et visualisation d'informations

## Contributions aux modèles de Markov cachés

Métaheuristiques pour l'apprentissage de MMC

Les modèles de Markov cachés à substitutions de symboles

Matrice de *scatterplots* pseudo-euclidienne

HMMTK : bibliothèque logicielle de manipulation de MMC

## Conclusions et perspectives

## Métaheuristiques pour l'apprentissage de MMC (1/2)

- ▶ Métaheuristique : procédure générique utilisée pour explorer un espace de solutions

## Métaheuristiques pour l'apprentissage de MMC (1/2)

- ▶ Métaheuristique : procédure générique utilisée pour explorer un espace de solutions
- ▶ Les métaheuristiques qui ont été employées :
  - ▶ Recherche aléatoire
  - ▶ Recuit simulé (Paul, 1985) (Hamam and Al Ani, 1996)
  - ▶ Recherche tabou (Chen et al., 2004)
  - ▶ Les algorithmes génétiques (Kwong and Chau, 1997) (Slimane et al., 1999) (Brouard, 1999) (Thomsen, 2002)
  - ▶ Apprentissage incrémental à base de population (Maxwell and Anderson, 1999)
  - ▶ Colonie de fourmis artificielles (Monmarché, 2000)
  - ▶ Optimisation par essaim particulière (Rasmussen and T. Krink, 2003)

## Métaheuristiques pour l'apprentissage de MMC (1/2)

- ▶ Métaheuristique : procédure générique utilisée pour explorer un espace de solutions
- ▶ Les métaheuristiques qui ont été employées :
  - ▶ Recherche aléatoire
  - ▶ Recuit simulé (Paul, 1985) (Hamam and Al Ani, 1996)
  - ▶ Recherche tabou (Chen et al., 2004)
  - ▶ **Les algorithmes génétiques** (Kwong and Chau, 1997) (**Slimane et al., 1999**) (**Brouard, 1999**) (Thomsen, 2002)
  - ▶ Apprentissage incrémental à base de population (Maxwell and Anderson, 1999)
  - ▶ **Colonie de fourmis artificielles** (Monmarché, 2000)
  - ▶ Optimisation par essaim particulière (Rasmussen and T. Krink, 2003)

## Métaheuristiques pour l'apprentissage de MMC (2/2)

- ▶ Ces travaux posent souvent des problèmes :
  - ▶ Les paramètres des méthodes sont insuffisamment étudiés

## Métaheuristiques pour l'apprentissage de MMC (2/2)

- ▶ Ces travaux posent souvent des problèmes :
  - ▶ Les paramètres des méthodes sont insuffisamment étudiés
  - ▶ Comparaison à des heuristiques spécialisées uniquement sur des jeux de données très particuliers (MMC gauche-droite, . . .)

## Métaheuristiques pour l'apprentissage de MMC (2/2)

- ▶ Ces travaux posent souvent des problèmes :
  - ▶ Les paramètres des méthodes sont insuffisamment étudiés
  - ▶ Comparaison à des heuristiques spécialisées uniquement sur des jeux de données très particuliers (MMC gauche-droite, . . .)
  - ▶ Aucune comparaison des métaheuristiques entre elles

## Métaheuristiques pour l'apprentissage de MMC (2/2)

- ▶ Ces travaux posent souvent des problèmes :
  - ▶ Les paramètres des méthodes sont insuffisamment étudiés
  - ▶ Comparaison à des heuristiques spécialisées uniquement sur des jeux de données très particuliers (MMC gauche-droite, . . .)
  - ▶ Aucune comparaison des métaheuristiques entre elles
  - ▶ Aucun jeu de données standards pour croiser les résultats

## Métaheuristiques pour l'apprentissage de MMC (2/2)

- ▶ Ces travaux posent souvent des problèmes :
  - ▶ Les paramètres des méthodes sont insuffisamment étudiés
  - ▶ Comparaison à des heuristiques spécialisées uniquement sur des jeux de données très particuliers (MMC gauche-droite, ...)
  - ▶ Aucune comparaison des métaheuristiques entre elles
  - ▶ Aucun jeu de données standards pour croiser les résultats
- ▶ Bilan : **On ne sait pas quelle stratégie fonctionne le mieux !**

# Plan

## Etat de l'art

Les modèles de Markov cachés

L'apprentissage de modèles de Markov cachés

**Fouille de données et visualisation d'informations**

Contributions aux modèles de Markov cachés

Métaheuristiques pour l'apprentissage de MMC

Les modèles de Markov cachés à substitutions de symboles

Matrice de *scatterplots* pseudo-euclidienne

HMMTK : bibliothèque logicielle de manipulation de MMC

Conclusions et perspectives

# La fouille de données

- ▶ Fouille de données : extraction de connaissances à partir de données

## La fouille de données

- ▶ Fouille de données : extraction de connaissances à partir de données
- ▶ Plusieurs types de fouille de données :

	Manuelle	Visuelle	Automatique
Concision des règles extraites	+++	++	+
Quantité de données	+	++	+++

## La fouille de données

- ▶ Fouille de données : extraction de connaissances à partir de données
- ▶ Plusieurs types de fouille de données :

	Manuelle	Visuelle	Automatique
Concision des règles extraites	+++	++	+
Quantité de données	+	++	+++

## La visualisation d'informations (1/2)

La fouille visuelle de données ou visualisation d'informations :

- ▶ Dépend principalement de trois facteurs :
  - ▶ les objectifs :
    - ▶ comprendre,
    - ▶ détecter,
    - ▶ confirmer,
    - ▶ expliquer,
    - ▶ ...

## La visualisation d'informations (1/2)

La fouille visuelle de données ou visualisation d'informations :

- ▶ Dépend principalement de trois facteurs :
  - ▶ les objectifs :
    - ▶ comprendre,
    - ▶ détecter,
    - ▶ confirmer,
    - ▶ expliquer,
    - ▶ ...
  - ▶ le type des données :
    - ▶ des données multidimensionnelles
    - ▶ des graphes
    - ▶ des données temporelles
    - ▶ des dissimilarités
    - ▶ ...

## La visualisation d'informations (1/2)

La fouille visuelle de données ou visualisation d'informations :

- ▶ Dépend principalement de trois facteurs :
  - ▶ les objectifs :
    - ▶ comprendre,
    - ▶ détecter,
    - ▶ confirmer,
    - ▶ expliquer,
    - ▶ ...
  - ▶ le type des données :
    - ▶ des données multidimensionnelles
    - ▶ des graphes
    - ▶ des données temporelles
    - ▶ des dissimilarités
    - ▶ ...
  - ▶ la quantité de données

## La visualisation d'informations (2/2)

- ▶ Plusieurs grandes catégories de techniques de visualisation :
  - ▶ matrice de *scatterplots*
  - ▶ les techniques de projection
  - ▶ la visualisation de graphes
  - ▶ les matrices de similarités en niveaux de gris
  - ▶ les *parallel coordinates*
  - ▶ l'empilement de dimensions
  - ▶ les visualisations iconiques
  - ▶ les techniques à base de pixels

## La visualisation d'informations (2/2)

- ▶ Plusieurs grandes catégories de techniques de visualisation :
  - ▶ *matrice de scatterplots*
  - ▶ les techniques de projection
  - ▶ la visualisation de graphes
  - ▶ les matrices de similarités en niveaux de gris
  - ▶ les *parallel coordinates*
  - ▶ l'empilement de dimensions
  - ▶ les visualisations iconiques
  - ▶ les techniques à base de pixels

## La visualisation d'informations (2/2)

- ▶ Plusieurs grandes catégories de techniques de visualisation :
  - ▶ matrice de *scatterplots*
  - ▶ les techniques de projection
  - ▶ la visualisation de graphes
  - ▶ les matrices de similarités en niveaux de gris
  - ▶ les *parallel coordinates*
  - ▶ l'empilement de dimensions
  - ▶ les visualisations iconiques
  - ▶ les techniques à base de pixels
- ▶ Il existe de nombreuses techniques de visualisation d'informations

## La visualisation d'informations (2/2)

- ▶ Plusieurs grandes catégories de techniques de visualisation :
  - ▶ matrice de *scatterplots*
  - ▶ les techniques de projection
  - ▶ la visualisation de graphes
  - ▶ les matrices de similarités en niveaux de gris
  - ▶ les *parallel coordinates*
  - ▶ l'empilement de dimensions
  - ▶ les visualisations iconiques
  - ▶ les techniques à base de pixels
- ▶ Il existe de nombreuses techniques de visualisation d'informations
- ▶ Remarque :
  - ▶ les techniques classiques contiennent tous les composants de base pour la création de visualisation d'informations

## La visualisation d'informations (2/2)

- ▶ Plusieurs grandes catégories de techniques de visualisation :
  - ▶ matrice de *scatterplots*
  - ▶ les techniques de projection
  - ▶ la visualisation de graphes
  - ▶ les matrices de similarités en niveaux de gris
  - ▶ les *parallel coordinates*
  - ▶ l'empilement de dimensions
  - ▶ les visualisations iconiques
  - ▶ les techniques à base de pixels
- ▶ Il existe de nombreuses techniques de visualisation d'informations
- ▶ Remarque :
  - ▶ les techniques classiques contiennent tous les composants de base pour la création de visualisation d'informations
  - ▶ la recherche dans le domaine de la fouille visuelle de données ne fait plus qu'agencer, adapter ou améliorer ces composants de base

# Plan

## Etat de l'art

- Les modèles de Markov cachés
- L'apprentissage de modèles de Markov cachés
- Fouille de données et visualisation d'informations

## Contributions aux modèles de Markov cachés

- Métaheuristiques pour l'apprentissage de MMC
- Les modèles de Markov cachés à substitutions de symboles
- Matrice de *scatterplots* pseudo-euclidienne
- HMMTK : bibliothèque logicielle de manipulation de MMC

## Conclusions et perspectives

# Plan

## Etat de l'art

- Les modèles de Markov cachés
- L'apprentissage de modèles de Markov cachés
- Fouille de données et visualisation d'informations

## Contributions aux modèles de Markov cachés

### Métaheuristiques pour l'apprentissage de MMC

- Les modèles de Markov cachés à substitutions de symboles
- Matrice de *scatterplots* pseudo-euclidienne
- HMMTK : bibliothèque logicielle de manipulation de MMC

## Conclusions et perspectives

## L'apprentissage de MMC (1/2)

- ▶ Difficultés de l'apprentissage de MMC :
  - ▶ fonction objectif complexe
  - ▶ de nombreux optima locaux
  - ▶ présence de « symétries » dans l'espace des solutions
  - ▶ souvent considérée comme un problème d'optimisation continue

## L'apprentissage de MMC (1/2)

- ▶ Difficultés de l'apprentissage de MMC :
  - ▶ fonction objectif complexe
  - ▶ de nombreux optima locaux
  - ▶ présence de « symétries » dans l'espace des solutions
  - ▶ souvent considérée comme un problème d'optimisation continue
- ▶ L'apprentissage de MMC :
  - ▶ la méthode la plus souvent utilisée :
    - ▶ optimiser sur un critère (ex : la vraisemblance)
    - ▶ évaluer sur un autre critère (ex : taux de bonne classification)

## L'apprentissage de MMC (1/2)

- ▶ Difficultés de l'apprentissage de MMC :
    - ▶ fonction objectif complexe
    - ▶ de nombreux optima locaux
    - ▶ présence de « symétries » dans l'espace des solutions
    - ▶ souvent considérée comme un problème d'optimisation continue
  - ▶ L'apprentissage de MMC :
    - ▶ la méthode la plus souvent utilisée :
      - ▶ optimiser sur un critère (ex : la vraisemblance)
      - ▶ évaluer sur un autre critère (ex : taux de bonne classification)
- ⇒ risque d'avoir des objectifs contradictoires

## L'apprentissage de MMC (2/2)

- ▶ La démarche que nous avons suivie :
  - ▶ **Optimiser et évaluer sur le même critère !**
    - ▶ on suppose que le critère est choisi en fonction de la tâche à réaliser

## L'apprentissage de MMC (2/2)

- ▶ La démarche que nous avons suivie :
  - ▶ **Optimiser et évaluer sur le même critère !**
    - ▶ on suppose que le critère est choisi en fonction de la tâche à réaliser
  - ▶ Critère du maximum de vraisemblance :
    - ▶ maximisation de  $P(V = O/\lambda)$

## L'apprentissage de MMC (2/2)

- ▶ La démarche que nous avons suivie :
  - ▶ **Optimiser et évaluer sur le même critère !**
    - ▶ on suppose que le critère est choisi en fonction de la tâche à réaliser
  - ▶ Critère du maximum de vraisemblance :
    - ▶ maximisation de  $P(V = O/\lambda)$
  - ▶ Expérimentation :
    - ▶ sur différents espaces de solutions
    - ▶ de plusieurs métaheuristiques

## L'apprentissage de MMC (2/2)

- ▶ La démarche que nous avons suivie :
  - ▶ **Optimiser et évaluer sur le même critère !**
    - ▶ on suppose que le critère est choisi en fonction de la tâche à réaliser
  - ▶ Critère du maximum de vraisemblance :
    - ▶ maximisation de  $P(V = O/\lambda)$
  - ▶ Expérimentation :
    - ▶ sur différents espaces de solutions
    - ▶ de plusieurs métaheuristiques
  - ▶ Hybridation avec l'heuristique de Baum-Welch :
    - ▶ on applique plusieurs itérations de Baum-Welch à chaque solution considérée

## Les espaces de solutions (1/2)

- ▶  $\Lambda$  : espace des triplets de matrices stochastiques
  - ▶ espace continu et convexe
  - ▶ les algorithmes *Forward* et de Baum-Welch sont définis sur  $\Lambda$

## Les espaces de solutions (1/2)

- ▶  $\Lambda$  : espace des triplets de matrices stochastiques
  - ▶ espace continu et convexe
  - ▶ les algorithmes *Forward* et de Baum-Welch sont définis sur  $\Lambda$
- ▶  $\mathcal{S}^T$  : espace des séquences d'états cachés de longueur  $T$ 
  - ▶ espace discret

## Les espaces de solutions (1/2)

- ▶  $\Lambda$  : espace des triplets de matrices stochastiques
  - ▶ espace continu et convexe
  - ▶ les algorithmes *Forward* et de Baum-Welch sont définis sur  $\Lambda$
- ▶  $\mathbb{S}^T$  : espace des séquences d'états cachés de longueur  $T$ 
  - ▶ espace discret
  - ▶ apprentissage étiqueté :  $\mathbb{S}^T \mapsto \Lambda$   
mais cette transformation n'est **ni injective, ni surjective**

## Les espaces de solutions (1/2)

- ▶  $\Lambda$  : espace des triplets de matrices stochastiques
  - ▶ espace continu et convexe
  - ▶ les algorithmes *Forward* et de Baum-Welch sont définis sur  $\Lambda$
- ▶  $\mathbb{S}^T$  : espace des séquences d'états cachés de longueur  $T$ 
  - ▶ espace discret
  - ▶ apprentissage étiqueté :  $\mathbb{S}^T \mapsto \Lambda$   
mais cette transformation n'est **ni injective, ni surjective**  
 $\implies$  on ne peut pas utiliser Baum-Welch sans perdre en cohérence

## Les espaces de solutions (2/2)

- ▶  $(\Omega, \oplus, \odot)$  (Aupetit et al., 2005c) :
  - ▶  $\Lambda^* \subset \Lambda$  : triplet de matrices stochastiques dont aucun coefficient n'est nul

## Les espaces de solutions (2/2)

- ▶  $(\Omega, \oplus, \odot)$  (Aupetit et al., 2005c) :
  - ▶  $\Lambda^* \subset \Lambda$  : triplet de matrices stochastiques dont aucun coefficient n'est nul
  - ▶  $\Lambda^* \begin{matrix} \xrightarrow{\psi} \\ \xleftarrow{\phi} \end{matrix} (\Omega, \oplus, \odot)$
  - ▶  $\psi$  et  $\phi$  sont **bijectives**
  - ▶  $(\Omega, \oplus, \odot)$  est un espace vectoriel

## Les espaces de solutions (2/2)

- ▶  $(\Omega, \oplus, \odot)$  (Aupetit et al., 2005c) :
  - ▶  $\Lambda^* \subset \Lambda$  : triplet de matrices stochastiques dont aucun coefficient n'est nul
  - ▶  $\Lambda^* \begin{array}{c} \xrightarrow{\psi} \\ \xleftarrow{\phi} \end{array} (\Omega, \oplus, \odot)$
  - ▶  $\psi$  et  $\phi$  sont **bijectives**
  - ▶  $(\Omega, \oplus, \odot)$  est un espace vectoriel
- ▶ Les algorithmes existants :

## Les espaces de solutions (2/2)

- ▶  $(\Omega, \oplus, \odot)$  (Aupetit et al., 2005c) :
  - ▶  $\Lambda^* \subset \Lambda$  : triplet de matrices stochastiques dont aucun coefficient n'est nul
  - ▶  $\Lambda^* \begin{matrix} \xrightarrow{\psi} \\ \xleftarrow{\phi} \end{matrix} (\Omega, \oplus, \odot)$
  - ▶  $\psi$  et  $\phi$  sont **bijectives**
  - ▶  $(\Omega, \oplus, \odot)$  est un espace vectoriel
- ▶ Les algorithmes existants :
  - ▶ utilisent souvent l'espace de solutions  $\Lambda$

## Les espaces de solutions (2/2)

- ▶  $(\Omega, \oplus, \odot)$  (Aupetit et al., 2005c) :
  - ▶  $\Lambda^* \subset \Lambda$  : triplet de matrices stochastiques dont aucun coefficient n'est nul
  - ▶  $\Lambda^* \begin{matrix} \xrightarrow{\psi} \\ \xleftarrow{\phi} \end{matrix} (\Omega, \oplus, \odot)$
  - ▶  $\psi$  et  $\phi$  sont **bijectives**
  - ▶  $(\Omega, \oplus, \odot)$  est un espace vectoriel
- ▶ Les algorithmes existants :
  - ▶ utilisent souvent l'espace de solutions  $\Lambda$
  - ▶ utilisent presque jamais l'espace de solutions  $\mathbb{S}^T$

## Les espaces de solutions (2/2)

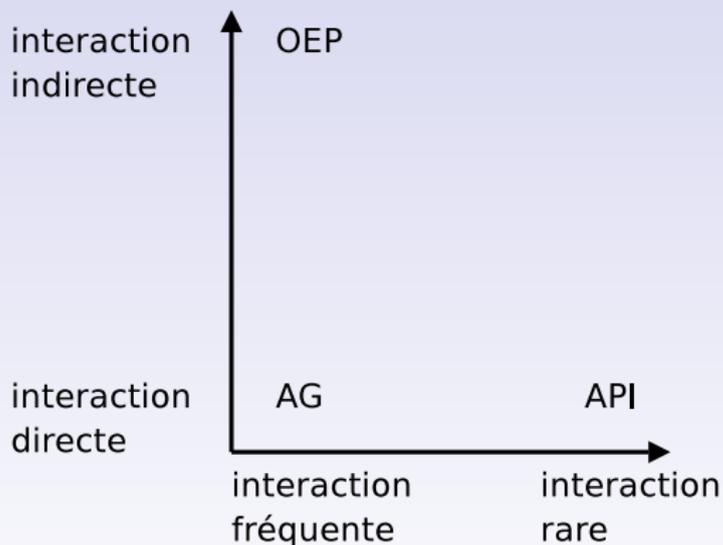
- ▶  $(\Omega, \oplus, \odot)$  (Aupetit et al., 2005c) :
  - ▶  $\Lambda^* \subset \Lambda$  : triplet de matrices stochastiques dont aucun coefficient n'est nul
  - ▶  $\Lambda^* \begin{matrix} \xrightarrow{\psi} \\ \xleftarrow{\phi} \end{matrix} (\Omega, \oplus, \odot)$
  - ▶  $\psi$  et  $\phi$  sont **bijectives**
  - ▶  $(\Omega, \oplus, \odot)$  est un espace vectoriel
- ▶ Les algorithmes existants :
  - ▶ utilisent souvent l'espace de solutions  $\Lambda$
  - ▶ utilisent presque jamais l'espace de solutions  $\mathbb{S}^T$
  - ▶ n'ont jamais utilisé la structure vectorielle  $(\Omega, \oplus, \odot)$

## Trois métaheuristiques génériques

Trois métaheuristiques à base de population :

- ▶ Les algorithmes génétiques (Holland, 1975)
  - ▶ modélisation des principes de l'évolution génétique
- ▶ L'algorithme API, fourmis artificielles (Monmarché, 2000)
  - ▶ modélisation du comportement de fourrageage des fourmis *Pachycondyla apicalis*
- ▶ L'optimisation par essaim particulaire (Kennedy and Eberhart, 1995)
  - ▶ modélisation du comportement de déplacement d'un essaim

## Les différences comportementales principales



## Nos adaptations

► Nos adaptations :

Algorithmes	$\Lambda$	$\mathcal{S}^T$	$\Omega$
Algorithmes génétiques	AG	AGDiscret	API* OEPDistance / OEPSocial
API (fourmis artificielles)	API	APIDiscret	
OEP			

## Nos adaptations

► Nos adaptations :

Algorithmes	$\Lambda$	$\mathcal{S}^T$	$\Omega$
Algorithmes génétiques	AG	AGDiscret	API* OEPDistance / OEPSocial
API (fourmis artificielles)	API	APIDiscret	
OEP			

► Hybridation avec l'algorithme de Baum-Welch

## Études comparatives (1/3)

- ▶ Expérimentation sur des images de visages de la base ORL  
(Samaria and Harter, 1994)



## Études comparatives (1/3)

- ▶ Expérimentation sur des images de visages de la base ORL  
(Samaria and Harter, 1994)



- ▶ Deux étapes :
  - ▶ Déterminer une bonne configuration de paramètres pour chaque métaheuristique
  - ▶ Comparer les métaheuristiques

## Études comparatives (2/3)

- ▶ Déterminer une bonne configuration de paramètres par métaheuristique
  - ▶ visualisation de l'impact de chaque paramètre sur la performance de la méthode (graphes PAPFP)
  - ▶ estimation lissée des profils de performances
  - ▶ sélection itérative des valeurs des paramètres

## Études comparatives (2/3)

- ▶ Déterminer une bonne configuration de paramètres par métaheuristique
  - ▶ visualisation de l'impact de chaque paramètre sur la performance de la méthode (graphes PAPFP)
  - ▶ estimation lissée des profils de performances
  - ▶ sélection itérative des valeurs des paramètres
- ▶ Conditions de la comparaison des métaheuristiques :
  - ▶ Sans Baum-Welch : 30000 évaluations par *Forward*
  - ▶ Avec Baum-Welch : 1000 itérations de Baum-Welch

## Études comparatives (3/3)

- ▶ Comparer les métaheuristiques : tendances générales

$\mathcal{N}_{\text{BW}}$	0	2	5
Performance	+	+++	+++
Rapidité de convergence	++	+++	++
Nombre d'opérations	+++	+	+

## Études comparatives (3/3)

- ▶ Comparer les métaheuristiques : tendances générales

$\mathcal{N}_{\text{BW}}$	0	2	5
Performance	+	+++	+++
Rapidité de convergence	++	+++	++
Nombre d'opérations	+++	+	+

## Études comparatives (3/3)

- ▶ Comparer les métaheuristiques : tendances générales

$\mathcal{N}_{BW}$	0	2	5
Performance	+	+++	+++
Rapidité de convergence	++	+++	++
Nombre d'opérations	+++	+	+

- ▶ Il faut utiliser l'hybridation avec Baum-Welch

## Études comparatives (3/3)

- ▶ Comparer les métaheuristiques : tendances générales

$\mathcal{N}_{\text{BW}}$	0	2	5
Performance	+	+++	+++
Rapidité de convergence	++	+++	++
Nombre d'opérations	+++	+	+

- ▶ Il faut utiliser l'hybridation avec Baum-Welch
- ▶ Il ne faut pas trop d'itérations de Baum-Welch

## Études comparatives (3/3)

- ▶ Comparer les métaheuristiques : tendances générales

$\mathcal{N}_{\text{BW}}$	0	2	5
Performance	+	+++	+++
Rapidité de convergence	++	+++	++
Nombre d'opérations	+++	+	+

- ▶ Il faut utiliser l'hybridation avec Baum-Welch
- ▶ Il ne faut pas trop d'itérations de Baum-Welch

## Études comparatives (3/3)

- ▶ Comparer les métaheuristiques : tendances générales

$\mathcal{N}_{BW}$	0	2	5
Performance	+	+++	+++
Rapidité de convergence	++	+++	++
Nombre d'opérations	+++	+	+

- ▶ Il faut utiliser l'hybridation avec Baum-Welch
- ▶ Il ne faut pas trop d'itérations de Baum-Welch
- ▶ Cas  $\mathcal{N}_{BW} > 0$  :
  - ▶ Meilleur : API avec un paramétrage hétérogène et AG sur  $\Lambda$
  - ▶ API sur  $\Lambda$  avec un paramétrage hétérogène  $\geq$  AG sur  $\Lambda$

## Études comparatives (3/3)

- ▶ Comparer les métaheuristiques : tendances générales

$\mathcal{N}_{BW}$	0	2	5
Performance	+	+++	+++
Rapidité de convergence	++	+++	++
Nombre d'opérations	+++	+	+

- ▶ Il faut utiliser l'hybridation avec Baum-Welch
- ▶ Il ne faut pas trop d'itérations de Baum-Welch
- ▶ Cas  $\mathcal{N}_{BW} > 0$  :
  - ▶ Meilleur : API avec un paramétrage hétérogène et AG sur  $\Lambda$
  - ▶ API sur  $\Lambda$  avec un paramétrage hétérogène  $\geq$  AG sur  $\Lambda$
- ▶ Les solutions initiales des algorithmes sur l'espace  $\mathbb{S}^T$  sont de meilleure qualité que celles obtenues sur  $\Lambda$

# Plan

## Etat de l'art

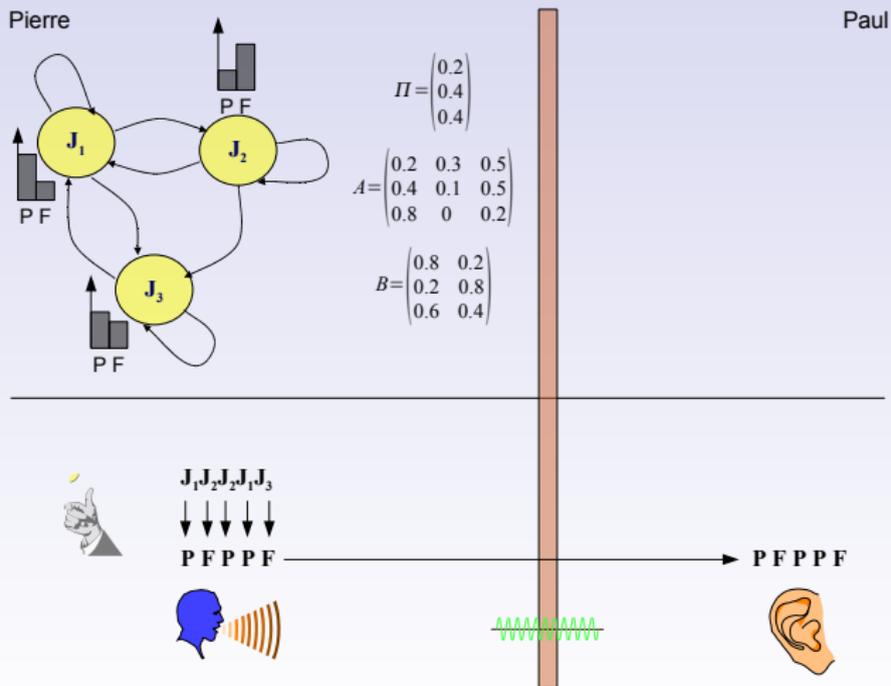
- Les modèles de Markov cachés
- L'apprentissage de modèles de Markov cachés
- Fouille de données et visualisation d'informations

## Contributions aux modèles de Markov cachés

- Métaheuristiques pour l'apprentissage de MMC
- Les modèles de Markov cachés à substitutions de symboles**
- Matrice de *scatterplots* pseudo-euclidienne
- HMMTK : bibliothèque logicielle de manipulation de MMC

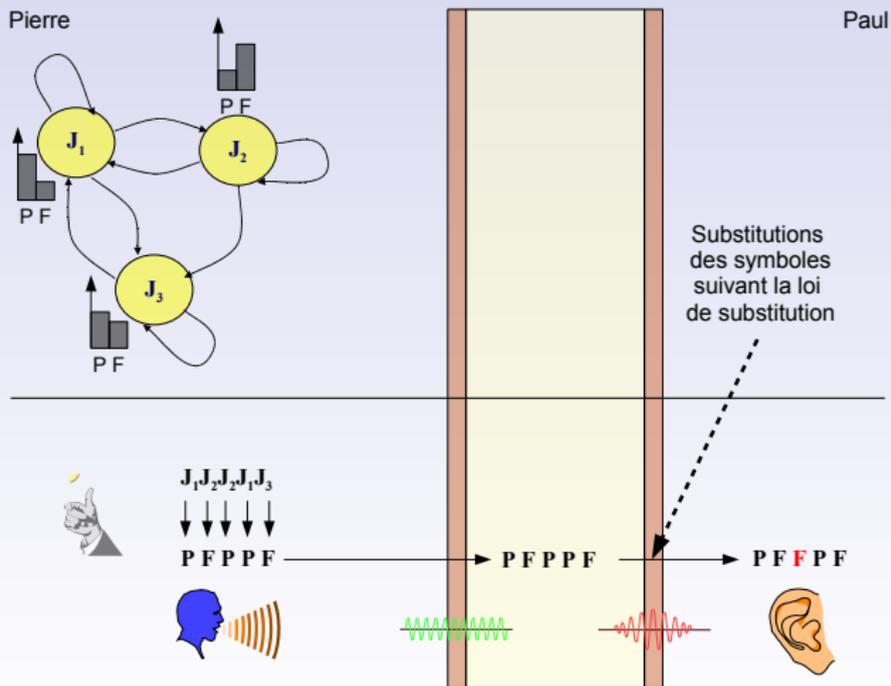
## Conclusions et perspectives

# Les MMC à Substitutions de Symboles (MMCSS) (1/3)



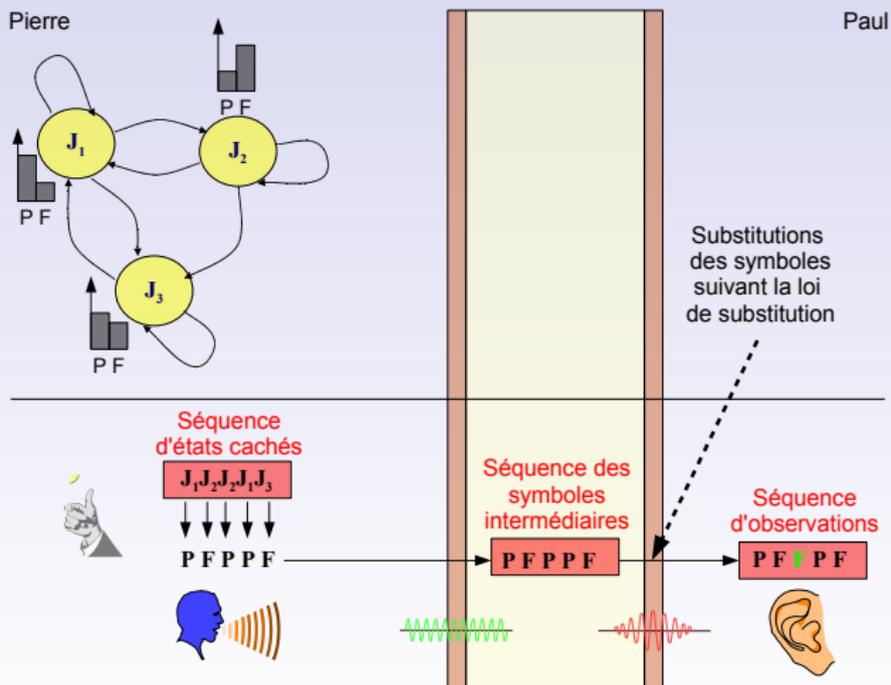
# Les MMC à Substitutions de Symboles (MMCSS) (1/3)

(Aupetit et al., 2002)



# Les MMC à Substitutions de Symboles (MMCSS) (1/3)

(Aupetit et al., 2002)



## Les MMC à Substitutions de Symboles (MMCSS) (2/3)

Les algorithmes indispensables aux MMC ont dû être adaptés : (Aupetit et al., 2002) :

## Les MMC à Substitutions de Symboles (MMCSS) (2/3)

Les algorithmes indispensables aux MMC ont dû être adaptés : (Aupetit et al., 2002) :

- ▶ Calcul de la vraisemblance :
  - ▶ **algorithme Forward / Backward**

## Les MMC à Substitutions de Symboles (MMCSS) (2/3)

Les algorithmes indispensables aux MMC ont dû être adaptés : (Aupetit et al., 2002) :

- ▶ Calcul de la vraisemblance :
  - ▶ **algorithme Forward / Backward**
- ▶ Segmentation d'une séquence d'observations :

## Les MMC à Substitutions de Symboles (MMCSS) (2/3)

Les algorithmes indispensables aux MMC ont dû être adaptés : (Aupetit et al., 2002) :

- ▶ Calcul de la vraisemblance :
  - ▶ **algorithme Forward / Backward**
- ▶ Segmentation d'une séquence d'observations :
  - ▶ calculer la séquence cachée la plus probablement suivie :
    - ▶ **algorithme de Viterbi**

## Les MMC à Substitutions de Symboles (MMCSS) (2/3)

Les algorithmes indispensables aux MMC ont dû être adaptés : (Aupetit et al., 2002) :

- ▶ Calcul de la vraisemblance :
  - ▶ **algorithme Forward / Backward**
- ▶ Segmentation d'une séquence d'observations :
  - ▶ calculer la séquence cachée la plus probablement suivie :
    - ▶ **algorithme de Viterbi**
  - ▶ calculer la séquence cachée et la séquence intermédiaire les plus probablement suivies :
    - ▶ **algorithme de Viterbi « étendu »**

## Les MMC à Substitutions de Symboles (MMCSS) (3/3)

- ▶ Ajustement des paramètres d'un modèle :

## Les MMC à Substitutions de Symboles (MMCSS) (3/3)

- ▶ Ajustement des paramètres d'un modèle :
  - ▶ le cas de la maximisation de la vraisemblance :
    - ▶ **algorithme de Baum-Welch**

## Les MMC à Substitutions de Symboles (MMCSS) (3/3)

- ▶ Ajustement des paramètres d'un modèle :
  - ▶ le cas de la maximisation de la vraisemblance :
    - ▶ **algorithme de Baum-Welch**
  - ▶ le cas général :
    - ▶ **calcul du gradient**

## Les MMC à Substitutions de Symboles (MMCSS) (3/3)

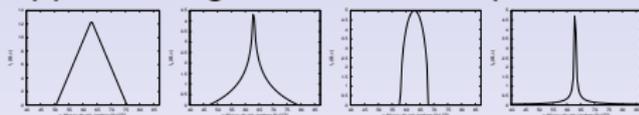
- ▶ Ajustement des paramètres d'un modèle :
  - ▶ le cas de la maximisation de la vraisemblance :
    - ▶ **algorithme de Baum-Welch**
  - ▶ le cas général :
    - ▶ **calcul du gradient**
- ▶ Propriété remarquable :
  - ▶ Les algorithmes MMCSS ont des complexités du même ordre que celles des algorithmes MMC du point de vue de la longueur de la séquence d'observations

## Applications des MMCSS (Aupetit et al., 2004) (1/2)

- ▶ Classification d'images sur la base ORL (Samaria and Harter, 1994) :

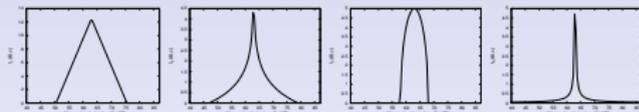
## Applications des MMCSS (Aupetit et al., 2004) (1/2)

- ▶ Classification d'images sur la base ORL (Samaria and Harter, 1994) :
  - ▶ Apprentissage de MMCSS à partir d'images avec différentes lois



## Applications des MMCSS (Aupetit et al., 2004) (1/2)

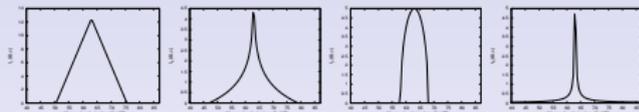
- ▶ Classification d'images sur la base ORL (Samaria and Harter, 1994) :
  - ▶ Apprentissage de MMCSS à partir d'images avec différentes lois



- ▶ Méthode :
  - ▶ Image  $O$ , apprise en  $\lambda$
  - ▶ Classement des 400 images par vraisemblance décroissante

## Applications des MMCSS (Aupetit et al., 2004) (1/2)

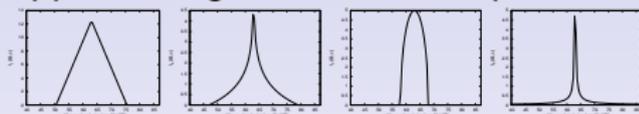
- ▶ Classification d'images sur la base ORL (Samaria and Harter, 1994) :
  - ▶ Apprentissage de MMCSS à partir d'images avec différentes lois



- ▶ Méthode :
  - ▶ Image  $O$ , apprise en  $\lambda$
  - ▶ Classement des 400 images par vraisemblance décroissante
- ▶ Objectifs :
  - ▶ Les images de la même classe que  $O$  doivent se trouver au début du classement

## Applications des MMCSS (Aupetit et al., 2004) (1/2)

- ▶ Classification d'images sur la base ORL (Samaria and Harter, 1994) :
  - ▶ Apprentissage de MMCSS à partir d'images avec différentes lois



- ▶ Méthode :
  - ▶ Image  $O$ , apprise en  $\lambda$
  - ▶ Classement des 400 images par vraisemblance décroissante
- ▶ Objectifs :
  - ▶ Les images de la même classe que  $O$  doivent se trouver au début du classement
- ▶ Résultats (pour une loi bien choisie) :
  - ▶ Réduction des rejets catégoriques
  - ▶ Les images de la classe de  $O$  remontent dans le classement
  - ▶ Il n'est plus nécessaire de ré-échelonner les niveaux de gris grâce à l'incorporation d'informations

## Applications des MMCSS (2/2)

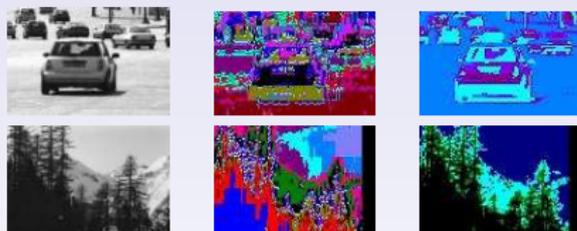
- ▶ Segmentation d'images :

## Applications des MMCSS (2/2)

- ▶ Segmentation d'images :
  - ▶ Comparaison de la segmentation d'images avec des lois de formes différentes (triangle, triangle non nul, gaussienne)

## Applications des MMCSS (2/2)

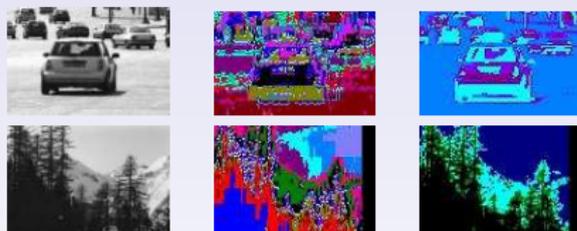
- ▶ Segmentation d'images :
  - ▶ Comparaison de la segmentation d'images avec des lois de formes différentes (triangle, triangle non nul, gaussienne)
  - ▶ Résultats : (Original, MMC, MMCSS loi gaussienne)



	MMC	MMCSS
Homogénéité / taille des zones	+	+++
Nombre d'états cachés après segmentation	+++	+

## Applications des MMCSS (2/2)

- ▶ Segmentation d'images :
  - ▶ Comparaison de la segmentation d'images avec des lois de formes différentes (triangle, triangle non nul, gaussienne)
  - ▶ Résultats : (Original, MMC, MMCSS loi gaussienne)



	MMC	MMCSS
Homogénéité / taille des zones	+	+++
Nombre d'états cachés après segmentation	+++	+

# Plan

## Etat de l'art

- Les modèles de Markov cachés
- L'apprentissage de modèles de Markov cachés
- Fouille de données et visualisation d'informations

## Contributions aux modèles de Markov cachés

- Métaheuristiques pour l'apprentissage de MMC
- Les modèles de Markov cachés à substitutions de symboles
- Matrice de *scatterplots* pseudo-euclidienne**
- HMMTK : bibliothèque logicielle de manipulation de MMC

## Conclusions et perspectives

## Pour mieux comprendre ce que des MMC modélisent

Plusieurs possibilités :

## Pour mieux comprendre ce que des MMC modélisent

Plusieurs possibilités :

- ▶ Avec un MMC :

## Pour mieux comprendre ce que des MMC modélisent

Plusieurs possibilités :

- ▶ Avec un MMC :
  - ▶ étude de ses distributions de probabilités

## Pour mieux comprendre ce que des MMC modélisent

Plusieurs possibilités :

- ▶ Avec un MMC :
  - ▶ étude de ses distributions de probabilités
  - ▶ étude de l'impact d'un MMC sur plusieurs données :
    - ▶ vraisemblance
    - ▶ segmentation
    - ▶ ...

## Pour mieux comprendre ce que des MMC modélisent

Plusieurs possibilités :

- ▶ Avec un MMC :
  - ▶ étude de ses distributions de probabilités
  - ▶ étude de l'impact d'un MMC sur plusieurs données :
    - ▶ vraisemblance
    - ▶ segmentation
    - ▶ ...
- ▶ Avec un ensemble de MMC :

## Pour mieux comprendre ce que des MMC modélisent

Plusieurs possibilités :

- ▶ Avec un MMC :
  - ▶ étude de ses distributions de probabilités
  - ▶ étude de l'impact d'un MMC sur plusieurs données :
    - ▶ vraisemblance
    - ▶ segmentation
    - ▶ ...
- ▶ Avec un ensemble de MMC :
  - ▶ étude de leurs ressemblances / relations :
    - ▶ Approche multidimensionnelle
    - ▶ Approche par dissimilarité

## L'approche multidimensionnelle

- ▶ On considère les MMC comme des points dans un espace  $\mathbb{R}^k$

## L'approche multidimensionnelle

- ▶ On considère les MMC comme des points dans un espace  $\mathbb{R}^k$
- ▶ De nombreuses techniques de visualisation sont disponibles

## L'approche multidimensionnelle

- ▶ On considère les MMC comme des points dans un espace  $\mathbb{R}^k$
- ▶ De nombreuses techniques de visualisation sont disponibles
- ▶ Si l'espace de représentation est  $\Lambda$  :

## L'approche multidimensionnelle

- ▶ On considère les MMC comme des points dans un espace  $\mathbb{R}^K$
- ▶ De nombreuses techniques de visualisation sont disponibles
- ▶ Si l'espace de représentation est  $\Lambda$  :
  - ▶ il y a trop de dimensions
    - ▶ exemple : pour une image en 256 NdG et 10 états cachés, on a  $K = 2670$

## L'approche multidimensionnelle

- ▶ On considère les MMC comme des points dans un espace  $\mathbb{R}^K$
- ▶ De nombreuses techniques de visualisation sont disponibles
- ▶ Si l'espace de représentation est  $\Lambda$  :
  - ▶ il y a trop de dimensions
    - ▶ exemple : pour une image en 256 NdG et 10 états cachés, on a  $K = 2670$
  - ▶ espace trop contraint :
    - ▶ les motifs prédominants risquent d'être ceux dus aux contraintes

## L'approche multidimensionnelle

- ▶ On considère les MMC comme des points dans un espace  $\mathbb{R}^K$
- ▶ De nombreuses techniques de visualisation sont disponibles
- ▶ Si l'espace de représentation est  $\Lambda$  :
  - ▶ il y a trop de dimensions
    - ▶ exemple : pour une image en 256 NdG et 10 états cachés, on a  $K = 2670$
  - ▶ espace trop contraint :
    - ▶ les motifs prédominants risquent d'être ceux dus aux contraintes
- ▶ Si l'espace de représentation est  $(\Omega^*, \oplus, \odot)$  :
  - ▶ il y a toujours trop de dimensions

## L'approche multidimensionnelle

- ▶ On considère les MMC comme des points dans un espace  $\mathbb{R}^K$
- ▶ De nombreuses techniques de visualisation sont disponibles
- ▶ Si l'espace de représentation est  $\Lambda$  :
  - ▶ il y a trop de dimensions
    - ▶ exemple : pour une image en 256 NdG et 10 états cachés, on a  $K = 2670$
  - ▶ espace trop contraint :
    - ▶ les motifs prédominants risquent d'être ceux dus aux contraintes
- ▶ Si l'espace de représentation est  $(\Omega^*, \oplus, \odot)$  :
  - ▶ il y a toujours trop de dimensions
- ▶ Les états d'un MMC peuvent être renumérotés :
  - ▶ les ressemblances ne se traduisent pas forcément par une proximité des coordonnées dans  $\mathbb{R}^K$

## L'approche par dissimilarité

### ► Propriétés :

- symétrie :  $w(x, y) = w(y, x)$
- positivité :  $w(x, y) \geq 0$
- auto-similarité :  $w(x, x) = 0$

## L'approche par dissimilarité

- ▶ Propriétés :
  - ▶ symétrie :  $w(x, y) = w(y, x)$
  - ▶ positivité :  $w(x, y) \geq 0$
  - ▶ auto-similarité :  $w(x, x) = 0$
- ▶ Grande flexibilité : peu de contraintes sur le choix de la mesure

## L'approche par dissimilarité

- ▶ Propriétés :
  - ▶ symétrie :  $w(x, y) = w(y, x)$
  - ▶ positivité :  $w(x, y) \geq 0$
  - ▶ auto-similarité :  $w(x, x) = 0$
- ▶ Grande flexibilité : peu de contraintes sur le choix de la mesure
- ▶ Généralisable : on peut appliquer la méthode à d'autres types d'objets

## L'approche par dissimilarité

- ▶ Propriétés :
  - ▶ symétrie :  $w(x, y) = w(y, x)$
  - ▶ positivité :  $w(x, y) \geq 0$
  - ▶ auto-similarité :  $w(x, x) = 0$
- ▶ Grande flexibilité : peu de contraintes sur le choix de la mesure
- ▶ Généralisable : on peut appliquer la méthode à d'autres types d'objets
- ▶ Visualisation :

## L'approche par dissimilarité

- ▶ Propriétés :
  - ▶ symétrie :  $w(x, y) = w(y, x)$
  - ▶ positivité :  $w(x, y) \geq 0$
  - ▶ auto-similarité :  $w(x, x) = 0$
- ▶ Grande flexibilité : peu de contraintes sur le choix de la mesure
- ▶ Généralisable : on peut appliquer la méthode à d'autres types d'objets
- ▶ Visualisation :
  - ▶ de graphes : rend les groupes visibles

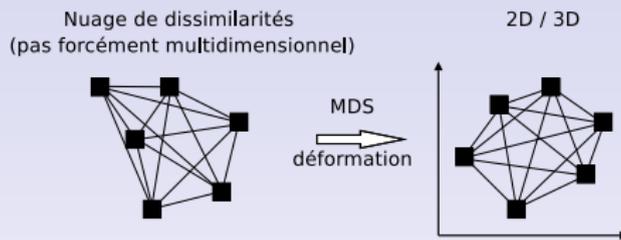
## L'approche par dissimilarité

- ▶ Propriétés :
  - ▶ symétrie :  $w(x, y) = w(y, x)$
  - ▶ positivité :  $w(x, y) \geq 0$
  - ▶ auto-similarité :  $w(x, x) = 0$
- ▶ Grande flexibilité : peu de contraintes sur le choix de la mesure
- ▶ Généralisable : on peut appliquer la méthode à d'autres types d'objets
- ▶ Visualisation :
  - ▶ de graphes : rend les groupes visibles
  - ▶ d'un nuage de points : recherche de représentants des MMC dans un espace selon un certain critère

## L'approche par dissimilarité

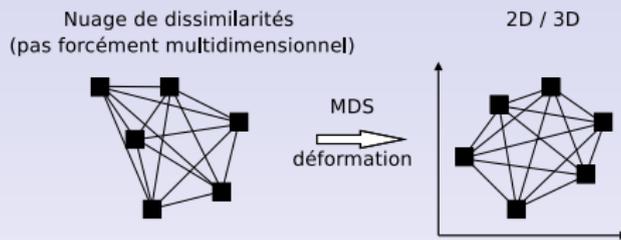
- ▶ Propriétés :
  - ▶ symétrie :  $w(x, y) = w(y, x)$
  - ▶ positivité :  $w(x, y) \geq 0$
  - ▶ auto-similarité :  $w(x, x) = 0$
- ▶ Grande flexibilité : peu de contraintes sur le choix de la mesure
- ▶ Généralisable : on peut appliquer la méthode à d'autres types d'objets
- ▶ Visualisation :
  - ▶ de graphes : rend les groupes visibles
  - ▶ **d'un nuage de points** : recherche de représentants des MMC dans un espace selon un certain critère

## Plongement de dissimilarités dans un espace réel (1/3)



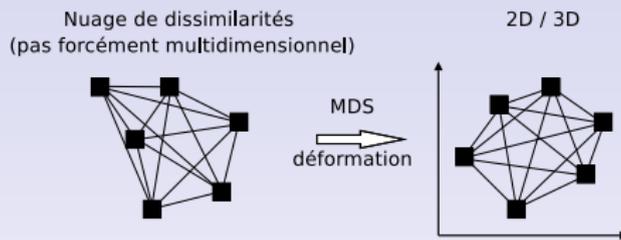
- ▶ L'approche classique par MDS :
  - ▶ plongement dans un espace de faibles dimensions (2D ou 3D)

## Plongement de dissimilarités dans un espace réel (1/3)



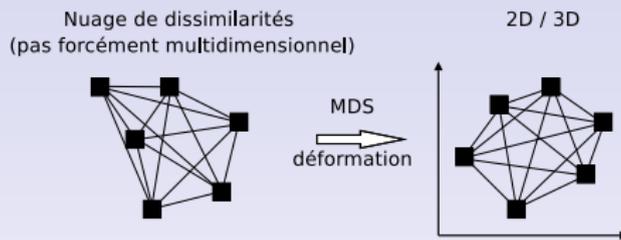
- ▶ L'approche classique par MDS :
  - ▶ plongement dans un espace de faibles dimensions (2D ou 3D)
  - ▶ visualisation du nuage 2D ou 3D

## Plongement de dissimilarités dans un espace réel (1/3)



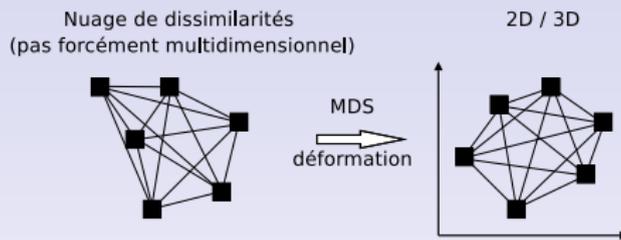
- ▶ L'approche classique par MDS :
  - ▶ plongement dans un espace de faibles dimensions (2D ou 3D)
  - ▶ visualisation du nuage 2D ou 3D
  - ▶ nombreux critères de plongement

# Plongement de dissimilarités dans un espace réel (1/3)



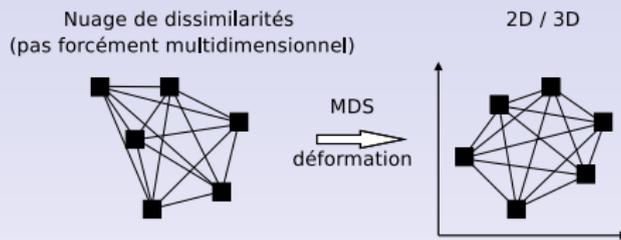
- ▶ L'approche classique par MDS :
  - ▶ plongement dans un espace de faibles dimensions (2D ou 3D)
  - ▶ visualisation du nuage 2D ou 3D
  - ▶ nombreux critères de plongement
  - ▶ les limites :

# Plongement de dissimilarités dans un espace réel (1/3)



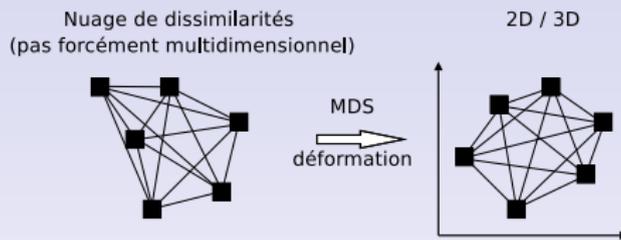
- ▶ L'approche classique par MDS :
  - ▶ plongement dans un espace de faibles dimensions (2D ou 3D)
  - ▶ visualisation du nuage 2D ou 3D
  - ▶ nombreux critères de plongement
  - ▶ les limites :
    - ▶ la dissimilarité est « encodée » visuellement

# Plongement de dissimilarités dans un espace réel (1/3)



- ▶ L'approche classique par MDS :
  - ▶ plongement dans un espace de faibles dimensions (2D ou 3D)
  - ▶ visualisation du nuage 2D ou 3D
  - ▶ nombreux critères de plongement
  - ▶ les limites :
    - ▶ la dissimilarité est « encodée » visuellement
    - ▶ le plongement est obtenu par optimisation d'un critère : le principe du meilleur compromis

# Plongement de dissimilarités dans un espace réel (1/3)



- ▶ L'approche classique par MDS :
  - ▶ plongement dans un espace de faibles dimensions (2D ou 3D)
  - ▶ visualisation du nuage 2D ou 3D
  - ▶ nombreux critères de plongement
  - ▶ les limites :
    - ▶ la dissimilarité est « encodée » visuellement
    - ▶ le plongement est obtenu par optimisation d'un critère : le principe du meilleur compromis
  - ▶ parfois, on déforme trop  $\implies$  devient inutilisable !!

# Plongement de dissimilarités dans un espace réel

(2/3) (Aupetit et al., 2005a) (Aupetit et al., 2005b)

- ▶ Plongement dans un espace pseudo-euclidien  $\mathbb{R}^K$  :
  - ▶  $\mathbb{R}^K$  est muni d'un produit scalaire pseudo-euclidien :  $\langle \cdot | \cdot \rangle$

## Plongement de dissimilarités dans un espace réel

(2/3) (Aupetit et al., 2005a) (Aupetit et al., 2005b)

- ▶ Plongement dans un espace pseudo-euclidien  $\mathbb{R}^K$  :
  - ▶  $\mathbb{R}^K$  est muni d'un produit scalaire pseudo-euclidien :  $\langle \cdot | \cdot \rangle$
  - ▶  $\langle \cdot | \cdot \rangle$  n'est pas forcément défini positif

## Plongement de dissimilarités dans un espace réel

(2/3) (Aupetit et al., 2005a) (Aupetit et al., 2005b)

- ▶ Plongement dans un espace pseudo-euclidien  $\mathbb{R}^K$  :
  - ▶  $\mathbb{R}^K$  est muni d'un produit scalaire pseudo-euclidien :  $\langle \cdot | \cdot \rangle$
  - ▶  $\langle \cdot | \cdot \rangle$  n'est pas forcément défini positif
  - ▶  $q$  est la forme quadratique associée

## Plongement de dissimilarités dans un espace réel

(2/3) (Aupetit et al., 2005a) (Aupetit et al., 2005b)

- ▶ Plongement dans un espace pseudo-euclidien  $\mathbb{R}^K$  :
  - ▶  $\mathbb{R}^K$  est muni d'un produit scalaire pseudo-euclidien :  $\langle \cdot | \cdot \rangle$
  - ▶  $\langle \cdot | \cdot \rangle$  n'est pas forcément défini positif
  - ▶  $q$  est la forme quadratique associée
  - ▶  $\psi : E \mapsto \mathbb{R}^K$  la fonction de plongement des données

## Plongement de dissimilarités dans un espace réel

(2/3) (Aupetit et al., 2005a) (Aupetit et al., 2005b)

- ▶ Plongement dans un espace pseudo-euclidien  $\mathbb{R}^K$  :
  - ▶  $\mathbb{R}^K$  est muni d'un produit scalaire pseudo-euclidien :  $\langle \cdot | \cdot \rangle$
  - ▶  $\langle \cdot | \cdot \rangle$  n'est pas forcément défini positif
  - ▶  $q$  est la forme quadratique associée
  - ▶  $\psi : E \mapsto \mathbb{R}^K$  la fonction de plongement des données
  - ▶  $\psi$  et  $\langle \cdot | \cdot \rangle$  sont construits à partir de la dissimilarité  $w$  :

$$w(\mathbf{x}, \mathbf{y}) = q(\psi(\mathbf{x}) - \psi(\mathbf{y})) = \langle \psi(\mathbf{x}) - \psi(\mathbf{y}) | \psi(\mathbf{x}) - \psi(\mathbf{y}) \rangle$$

# Plongement de dissimilarités dans un espace réel

(3/3) (Aupetit et al., 2005a) (Aupetit et al., 2005b)

- ▶ Plongement dans un espace pseudo-euclidien
  - ▶ Avantages :

# Plongement de dissimilarités dans un espace réel

(3/3) (Aupetit et al., 2005a) (Aupetit et al., 2005b)

- ▶ Plongement dans un espace pseudo-euclidien
  - ▶ Avantages :
    - ▶ la dissimilarité n'est pas « encodée » visuellement mais structurellement via  $\psi$  et  $\langle \cdot | \cdot \rangle$

# Plongement de dissimilarités dans un espace réel

(3/3) (Aupetit et al., 2005a) (Aupetit et al., 2005b)

- ▶ Plongement dans un espace pseudo-euclidien
  - ▶ Avantages :
    - ▶ la dissimilarité n'est pas « encodée » visuellement mais structurellement via  $\psi$  et  $\langle \cdot | \cdot \rangle$
    - ▶ le plongement s'effectue sans déformation

# Plongement de dissimilarités dans un espace réel

(3/3) (Aupetit et al., 2005a) (Aupetit et al., 2005b)

- ▶ Plongement dans un espace pseudo-euclidien
  - ▶ Avantages :
    - ▶ la dissimilarité n'est pas « encodée » visuellement mais structurellement via  $\psi$  et  $\langle \cdot | \cdot \rangle$
    - ▶ le plongement s'effectue sans déformation
    - ▶ la forme quadratique  $q$  associée traduit les valeurs de  $w$

# Plongement de dissimilarités dans un espace réel

(3/3) (Aupetit et al., 2005a) (Aupetit et al., 2005b)

- ▶ Plongement dans un espace pseudo-euclidien
  - ▶ Avantages :
    - ▶ la dissimilarité n'est pas « encodée » visuellement mais structurellement via  $\psi$  et  $\langle \cdot | \cdot \rangle$
    - ▶ le plongement s'effectue sans déformation
    - ▶ la forme quadratique  $q$  associée traduit les valeurs de  $w$
  - ▶ Inconvénients :

# Plongement de dissimilarités dans un espace réel

(3/3) (Aupetit et al., 2005a) (Aupetit et al., 2005b)

- ▶ Plongement dans un espace pseudo-euclidien
  - ▶ Avantages :
    - ▶ la dissimilarité n'est pas « encodée » visuellement mais structurellement via  $\psi$  et  $\langle \cdot | \cdot \rangle$
    - ▶ le plongement s'effectue sans déformation
    - ▶ la forme quadratique  $q$  associée traduit les valeurs de  $w$
  - ▶ Inconvénients :
    - ▶ la proximité visuelle ne traduit pas forcément la dissimilarité

# Plongement de dissimilarités dans un espace réel

(3/3) (Aupetit et al., 2005a) (Aupetit et al., 2005b)

- ▶ Plongement dans un espace pseudo-euclidien
  - ▶ Avantages :
    - ▶ la dissimilarité n'est pas « encodée » visuellement mais structurellement via  $\psi$  et  $\langle \cdot | \cdot \rangle$
    - ▶ le plongement s'effectue sans déformation
    - ▶ la forme quadratique  $q$  associée traduit les valeurs de  $w$
  - ▶ Inconvénients :
    - ▶ la proximité visuelle ne traduit pas forcément la dissimilarité
    - ▶ on a besoin d'outils supplémentaires pour l'interprétation

# Plongement de dissimilarités dans un espace réel

(3/3) (Aupetit et al., 2005a) (Aupetit et al., 2005b)

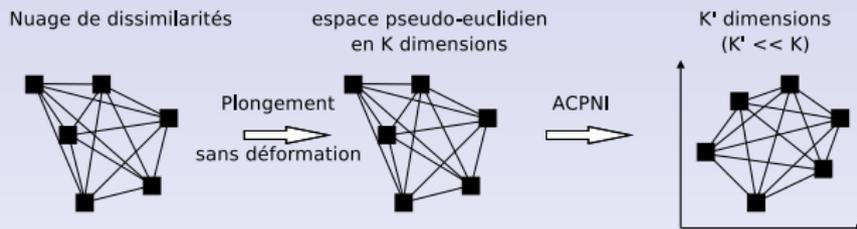
- ▶ Plongement dans un espace pseudo-euclidien
  - ▶ Avantages :
    - ▶ la dissimilarité n'est pas « encodée » visuellement mais structurellement via  $\psi$  et  $\langle \cdot | \cdot \rangle$
    - ▶ le plongement s'effectue sans déformation
    - ▶ la forme quadratique  $q$  associée traduit les valeurs de  $w$
  - ▶ Inconvénients :
    - ▶ la proximité visuelle ne traduit pas forcément la dissimilarité
    - ▶ on a besoin d'outils supplémentaires pour l'interprétation
    - ▶ l'espace dans lequel on plonge le nuage a beaucoup de dimensions :  $K$  objets  $\implies$  jusqu'à  $K - 1$  dimensions !!

# Plongement de dissimilarités dans un espace réel

(3/3) (Aupetit et al., 2005a) (Aupetit et al., 2005b)

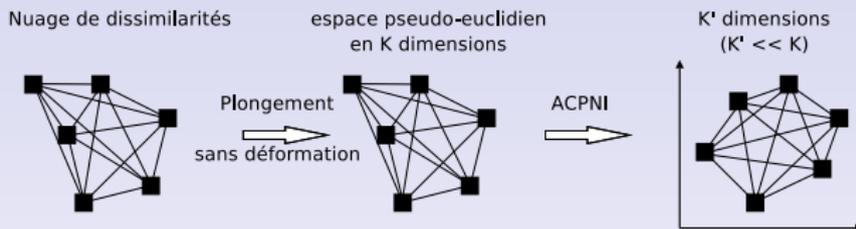
- ▶ Plongement dans un espace pseudo-euclidien
  - ▶ Avantages :
    - ▶ la dissimilarité n'est pas « encodée » visuellement mais structurellement via  $\psi$  et  $\langle \cdot | \cdot \rangle$
    - ▶ le plongement s'effectue sans déformation
    - ▶ la forme quadratique  $q$  associée traduit les valeurs de  $w$
  - ▶ Inconvénients :
    - ▶ la proximité visuelle ne traduit pas forcément la dissimilarité
    - ▶ on a besoin d'outils supplémentaires pour l'interprétation
    - ▶ l'espace dans lequel on plonge le nuage a beaucoup de dimensions :  $K$  objets  $\implies$  jusqu'à  $K - 1$  dimensions !!
  - ▶ Une solution : l'ACPNI

# ACPNI : Analyse en Composantes Principales à Noyau Indéfini



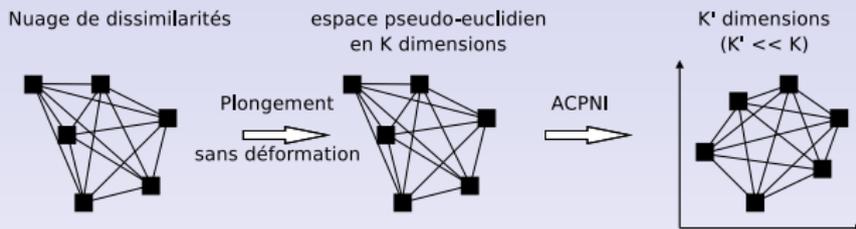
- ▶ Extraction des directions principales du nuage :

# ACPNI : Analyse en Composantes Principales à Noyau Indéfini



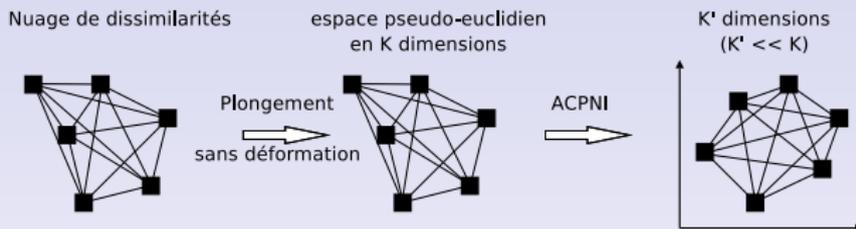
- ▶ Extraction des directions principales du nuage :
  - ▶ généralisation pseudo-euclidienne de l'ACP à noyau

# ACPNI : Analyse en Composantes Principales à Noyau Indéfini



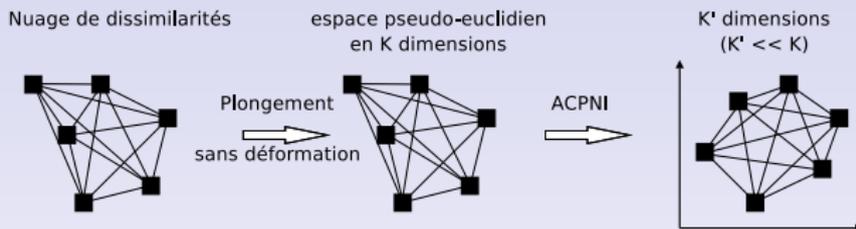
- ▶ Extraction des directions principales du nuage :
  - ▶ généralisation pseudo-euclidienne de l'ACP à noyau
  - ▶ centrage du nuage pour être indépendant de l'origine de l'espace

# ACPNI : Analyse en Composantes Principales à Noyau Indéfini



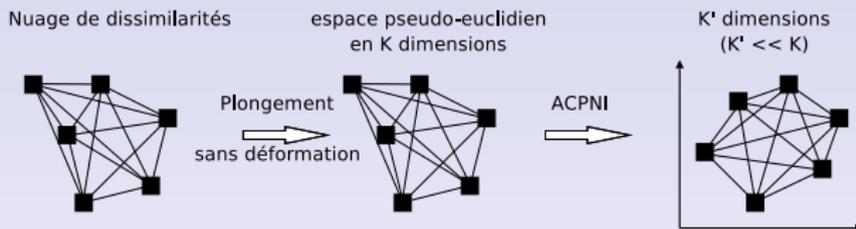
- ▶ Extraction des directions principales du nuage :
  - ▶ généralisation pseudo-euclidienne de l'ACP à noyau
  - ▶ centrage du nuage pour être indépendant de l'origine de l'espace
- ▶ Propriété : le plongement est unique à un isomorphisme près

## ACPNI : Analyse en Composantes Principales à Noyau Indéfini



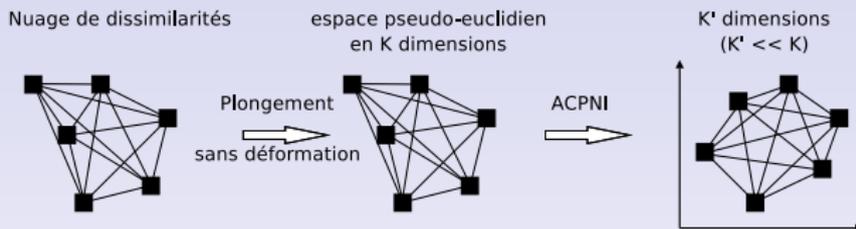
- ▶ Extraction des directions principales du nuage :
  - ▶ généralisation pseudo-euclidienne de l'ACP à noyau
  - ▶ centrage du nuage pour être indépendant de l'origine de l'espace
- ▶ Propriété : le plongement est unique à un isomorphisme près
- ▶ Intérêt de l'ACPNI :

# ACPNI : Analyse en Composantes Principales à Noyau Indéfini



- ▶ Extraction des directions principales du nuage :
  - ▶ généralisation pseudo-euclidienne de l'ACP à noyau
  - ▶ centrage du nuage pour être indépendant de l'origine de l'espace
- ▶ Propriété : le plongement est unique à un isomorphisme près
- ▶ Intérêt de l'ACPNI :
  - ▶ principe similaire à l'ACP classique (recherche des axes factoriels et projection sur les axes factoriels)

# ACPNI : Analyse en Composantes Principales à Noyau Indéfini



- ▶ Extraction des directions principales du nuage :
  - ▶ généralisation pseudo-euclidienne de l'ACP à noyau
  - ▶ centrage du nuage pour être indépendant de l'origine de l'espace
- ▶ Propriété : le plongement est unique à un isomorphisme près
- ▶ Intérêt de l'ACPNI :
  - ▶ principe similaire à l'ACP classique (recherche des axes factoriels et projection sur les axes factoriels)
  - ▶ sélection des axes résumant le mieux le nuage en fonction de l'amplitude des valeurs propres

## ACPNI et espaces pseudo-euclidiens en deux dimensions

- ▶ ACPNI : les axes factoriels forment une base orthogonale donc par un changement de base, on a :

$$q(\mathbf{x}) = \langle \mathbf{x} | \mathbf{x} \rangle = x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_{p+n}^2$$

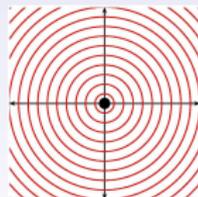
## ACPNI et espaces pseudo-euclidiens en deux dimensions

- ▶ ACPNI : les axes factoriels forment une base orthogonale donc par un changement de base, on a :

$$q(x) = \langle x|x \rangle = x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_{p+n}^2$$

- ▶ Propriétés des espaces pseudo-euclidiens en deux dimensions :

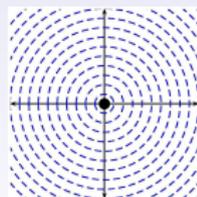
$$\mathcal{B}_c = \{x \in \mathbb{R}^K / q(x) = c\}$$



euclidien

(2,0)

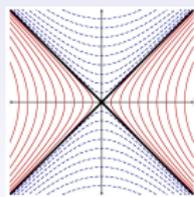
$$q_{i,j}(x) = x_i^2 + x_j^2$$



anti-euclidien

(0,2)

$$q_{i,j}(x) = -x_i^2 - x_j^2$$



hyperbolique

(1,1)

$$q_{i,j}(x) = x_i^2 - x_j^2$$

## La MSPE : matrice de *scatterplots* pseudo-euclidienne

- ▶ Représente le nuage selon tous les couples de dimensions après ACPNI  $\implies$  constitution d'une matrice

## La MSPE : matrice de *scatterplots* pseudo-euclidienne

- ▶ Représente le nuage selon tous les couples de dimensions après ACPNI  $\implies$  constitution d'une matrice
- ▶ Aides à l'interprétation :
  - ▶ les graphes sont proportionnels à leur contribution à la dissimilarité
  - ▶ affiche les isovaleurs (les hyper sphères) associés à un graphe
  - ▶ permet le marquage des points

## La MSPE : matrice de *scatterplots* pseudo-euclidienne

- ▶ Représente le nuage selon tous les couples de dimensions après ACPNI  $\implies$  constitution d'une matrice
- ▶ Aides à l'interprétation :
  - ▶ les graphes sont proportionnels à leur contribution à la dissimilarité
  - ▶ affiche les isovaleurs (les hyper sphères) associés à un graphe
  - ▶ permet le marquage des points
- ▶ Chaque point est représenté par un carré qui peut être rempli :
  - ▶ en fonction d'une valeur associée à la donnée (ex : la classe)
  - ▶ par une image

## Application de la MSPE aux MMC (1/4)

- ▶ Comparaison de 4 dissimilarités sur les 50 MMC

## Application de la MSPE aux MMC (1/4)

- ▶ Comparaison de 4 dissimilarités sur les 50 MMC
- ▶ 50 MMC obtenus par apprentissage de 50 visages répartis en 5 classes

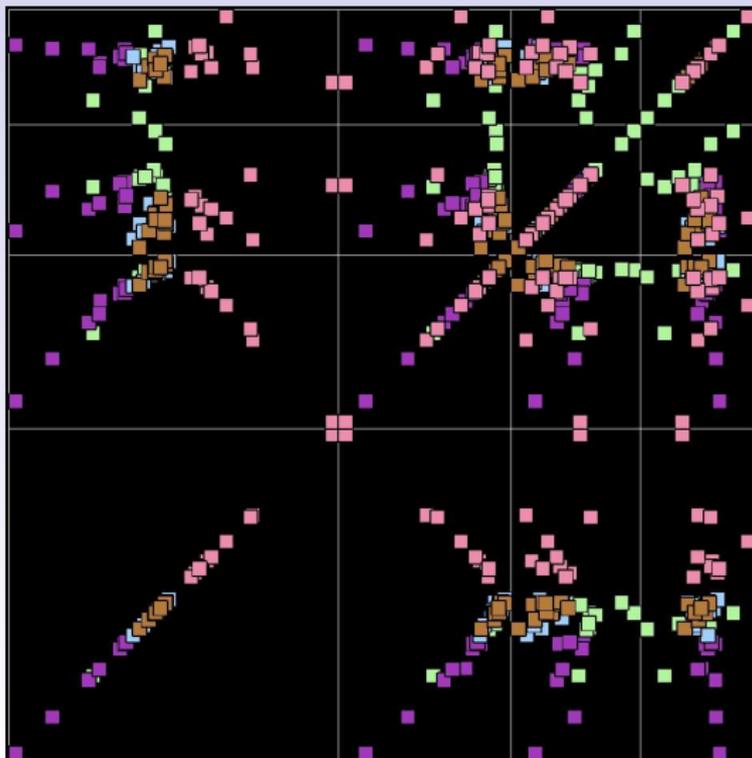
## Application de la MSPE aux MMC (1/4)

- ▶ Comparaison de 4 dissimilarités sur les 50 MMC
- ▶ 50 MMC obtenus par apprentissage de 50 visages répartis en 5 classes
- ▶ Les dissimilarités mesurent la ressemblance entre les distributions de niveaux de gris émis par les MMC

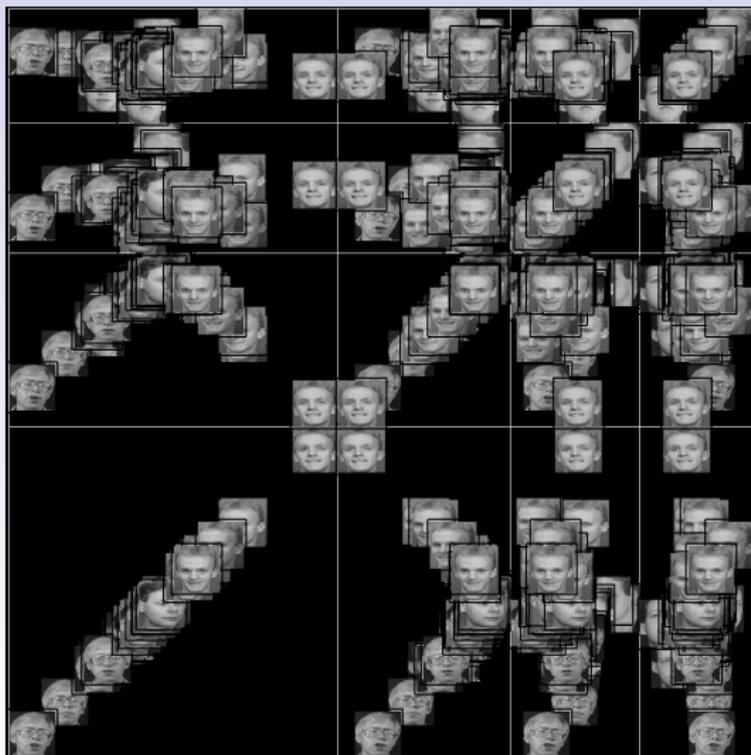
## Application de la MSPE aux MMC (1/4)

- ▶ Comparaison de 4 dissimilarités sur les 50 MMC
- ▶ 50 MMC obtenus par apprentissage de 50 visages répartis en 5 classes
- ▶ Les dissimilarités mesurent la ressemblance entre les distributions de niveaux de gris émis par les MMC
- ▶ Objectifs :
  - ▶ Que modélisent réellement les MMC ?
  - ▶ Quelles sont les différences entre les dissimilarités ?
  - ▶ Y-a-t-il une mesure de dissimilarités meilleure qu'une autre ?

## Application de la MSPE aux MMC (2/4)



## Application de la MSPE aux MMC (3/4)



## Application de la MSPE aux MMC (4/4)

Ce que nous avons pu montrer :

- ▶ une approche par MDS impliquerait de très grosses déformations

## Application de la MSPE aux MMC (4/4)

Ce que nous avons pu montrer :

- ▶ une approche par MDS impliquerait de très grosses déformations
- ▶ les quatre dissimilarités ont des échelles très différentes :
  - ▶ une analyse numérique serait lourde à mener
  - ▶ une analyse par MDS ne permettrait pas de conclure facilement

## Application de la MSPE aux MMC (4/4)

Ce que nous avons pu montrer :

- ▶ une approche par MDS impliquerait de très grosses déformations
- ▶ les quatre dissimilarités ont des échelles très différentes :
  - ▶ une analyse numérique serait lourde à mener
  - ▶ une analyse par MDS ne permettrait pas de conclure facilement
- ▶ les quatre dissimilarités sont globalement équivalentes :
  - ▶ les axes factoriels sont globalement dans les mêmes directions
  - ▶ il vaut mieux prendre la plus rapide à calculer car on ne perd rien
  - ▶ certaines dissimilarités sont plus sensibles au « bruit » que d'autres

## Application de la MSPE aux MMC (4/4)

Ce que nous avons pu montrer :

- ▶ une approche par MDS impliquerait de très grosses déformations
- ▶ les quatre dissimilarités ont des échelles très différentes :
  - ▶ une analyse numérique serait lourde à mener
  - ▶ une analyse par MDS ne permettrait pas de conclure facilement
- ▶ les quatre dissimilarités sont globalement équivalentes :
  - ▶ les axes factoriels sont globalement dans les mêmes directions
  - ▶ il vaut mieux prendre la plus rapide à calculer car on ne perd rien
  - ▶ certaines dissimilarités sont plus sensibles au « bruit » que d'autres
- ▶ les classes sont regroupées dans le système de coordonnées pseudo-euclidiens

## Application de la MSPE aux MMC (4/4)

Ce que nous avons pu montrer :

- ▶ une approche par MDS impliquerait de très grosses déformations
- ▶ les quatre dissimilarités ont des échelles très différentes :
  - ▶ une analyse numérique serait lourde à mener
  - ▶ une analyse par MDS ne permettrait pas de conclure facilement
- ▶ les quatre dissimilarités sont globalement équivalentes :
  - ▶ les axes factoriels sont globalement dans les mêmes directions
  - ▶ il vaut mieux prendre la plus rapide à calculer car on ne perd rien
  - ▶ certaines dissimilarités sont plus sensibles au « bruit » que d'autres
- ▶ les classes sont regroupées dans le système de coordonnées pseudo-euclidiens
- ▶ le premier axe factoriel correspond à la luminosité globale de l'image  $\implies$  la luminosité d'une image est modélisée par le MMC

# Plan

## Etat de l'art

- Les modèles de Markov cachés
- L'apprentissage de modèles de Markov cachés
- Fouille de données et visualisation d'informations

## Contributions aux modèles de Markov cachés

- Métaheuristiques pour l'apprentissage de MMC
- Les modèles de Markov cachés à substitutions de symboles
- Matrice de *scatterplots* pseudo-euclidienne
- HMMTK : bibliothèque logicielle de manipulation de MMC

## Conclusions et perspectives

## HMMTK : Hidden Markov Model Tool Kit

- ▶ Nous avons développé une bibliothèque de fonctions pour manipuler des MMC :
  - ▶ licence libre CeCILL v2 (compatible GPL)
  - ▶ généraliste
  - ▶ modulaire et extensible
  - ▶ capable de paralléliser les calculs coûteux de façon quasi-transparente

# Plan

## Etat de l'art

- Les modèles de Markov cachés
- L'apprentissage de modèles de Markov cachés
- Fouille de données et visualisation d'informations

## Contributions aux modèles de Markov cachés

- Métaheuristiques pour l'apprentissage de MMC
- Les modèles de Markov cachés à substitutions de symboles
- Matrice de *scatterplots* pseudo-euclidienne
- HMMTK : bibliothèque logicielle de manipulation de MMC

## Conclusions et perspectives

## Conclusions générales (1/4)

Nous avons pu contribuer aux modèles de Markov cachés dans trois domaines :

- ▶ L'apprentissage de MMC (pour la maximisation de la vraisemblance) :
  - ▶ Création et expérimentation d'une structure vectorielle pour les MMC

## Conclusions générales (1/4)

Nous avons pu contribuer aux modèles de Markov cachés dans trois domaines :

- ▶ L'apprentissage de MMC (pour la maximisation de la vraisemblance) :
  - ▶ Création et expérimentation d'une structure vectorielle pour les MMC
  - ▶ Adaptation, réglage et évaluation de plusieurs métaheuristiques sur différents types d'espace de solutions

## Conclusions générales (1/4)

Nous avons pu contribuer aux modèles de Markov cachés dans trois domaines :

- ▶ L'apprentissage de MMC (pour la maximisation de la vraisemblance) :
  - ▶ Création et expérimentation d'une structure vectorielle pour les MMC
  - ▶ Adaptation, réglage et évaluation de plusieurs métaheuristiques sur différents types d'espace de solutions
  - ▶ Comparaison des performances de plusieurs algorithmes



## Conclusions générales (3/4)

- ▶ La création et l'évaluation des MMC à substitutions de symboles (MMCSS)
  - ▶ Un MMCSS permet l'incorporation de connaissances expertes

## Conclusions générales (3/4)

- ▶ La création et l'évaluation des MMC à substitutions de symboles (MMCSS)
  - ▶ Un MMCSS permet l'incorporation de connaissances expertes
  - ▶ Les algorithmes standards ont pu être adaptés pour ces modèles

## Conclusions générales (3/4)

- ▶ La création et l'évaluation des MMC à substitutions de symboles (MMCSS)
  - ▶ Un MMCSS permet l'incorporation de connaissances expertes
  - ▶ Les algorithmes standards ont pu être adaptés pour ces modèles
  - ▶ Les propriétés apparues lors des expérimentations montrent leur intérêt pour la classification et la segmentation d'images

## Conclusions générales (4/4)

- ▶ La visualisation de dissimilarité pour la compréhension d'un ensemble de MMC :
  - ▶ Méthode générique d'analyse de dissimilarité

## Conclusions générales (4/4)

- ▶ La visualisation de dissimilarité pour la compréhension d'un ensemble de MMC :
  - ▶ Méthode générique d'analyse de dissimilarité
  - ▶ S'appuie sur les espaces pseudo-euclidiens et la notion de produit scalaire pseudo-euclidien

## Conclusions générales (4/4)

- ▶ La visualisation de dissimilarité pour la compréhension d'un ensemble de MMC :
  - ▶ Méthode générique d'analyse de dissimilarité
  - ▶ S'appuie sur les espaces pseudo-euclidiens et la notion de produit scalaire pseudo-euclidien
  - ▶ On effectue un plongement à but structurel plutôt qu'à but visuel

## Conclusions générales (4/4)

- ▶ La visualisation de dissimilarité pour la compréhension d'un ensemble de MMC :
  - ▶ Méthode générique d'analyse de dissimilarité
  - ▶ S'appuie sur les espaces pseudo-euclidiens et la notion de produit scalaire pseudo-euclidien
  - ▶ On effectue un plongement à but structurel plutôt qu'à but visuel
  - ▶ Deux techniques : ACPNI et MSPE

## Conclusions générales (4/4)

- ▶ La visualisation de dissimilarité pour la compréhension d'un ensemble de MMC :
  - ▶ Méthode générique d'analyse de dissimilarité
  - ▶ S'appuie sur les espaces pseudo-euclidiens et la notion de produit scalaire pseudo-euclidien
  - ▶ On effectue un plongement à but structurel plutôt qu'à but visuel
  - ▶ Deux techniques : ACPNI et MSPE
  - ▶ Nous avons montré l'intérêt de cette approche pour la comparaison de dissimilarité sur des MMC

## Conclusions générales (4/4)

- ▶ La visualisation de dissimilarité pour la compréhension d'un ensemble de MMC :
  - ▶ Méthode générique d'analyse de dissimilarité
  - ▶ S'appuie sur les espaces pseudo-euclidiens et la notion de produit scalaire pseudo-euclidien
  - ▶ On effectue un plongement à but structurel plutôt qu'à but visuel
  - ▶ Deux techniques : ACPNI et MSPE
  - ▶ Nous avons montré l'intérêt de cette approche pour la comparaison de dissimilarité sur des MMC
- ▶ Nous avons pu contribuer également aux MMC sous la forme d'une bibliothèque logiciel

## De nombreuses perspectives (1/4)

- ▶ Apprentissage de MMC :
  - ▶ Exploiter l'apprentissage étiqueté pour fournir une meilleure initialisation des autres métaheuristiques

## De nombreuses perspectives (1/4)

- ▶ Apprentissage de MMC :
  - ▶ Exploiter l'apprentissage étiqueté pour fournir une meilleure initialisation des autres métaheuristiques
  - ▶ Confirmer ou infirmer nos résultats sur :
    - ▶ D'autres données
    - ▶ D'autres critères
    - ▶ D'autres types de MMC

## De nombreuses perspectives (1/4)

- ▶ Apprentissage de MMC :
  - ▶ Exploiter l'apprentissage étiqueté pour fournir une meilleure initialisation des autres métaheuristiques
  - ▶ Confirmer ou infirmer nos résultats sur :
    - ▶ D'autres données
    - ▶ D'autres critères
    - ▶ D'autres types de MMC
  - ▶ Envisager d'autres pistes :

## De nombreuses perspectives (2/4)

Algorithme	$\Lambda$	$S^T$	$\Omega, \Lambda^*$
Apprentissage étiqueté Baum-Welch Information mutuelle  <i>Segmental k-means</i>	(Baum and Eagon, 1967) (Rabiner, 1989) (Schluter et al., 1997) (Vertanen, 2004)  (Juang and Rabiner, 1990)	X	
Recuit simulé	(Paul, 1985)	(Hamam and Al Ani, 1996)	
Recherche tabou	(Chen et al., 2004)		
Algorithme génétique	(Slimane et al., 1999) (Brouard, 1999)* (Thomsen, 2002)	(Aupetit, 2005) +Baum-Welch	X
PBIL	(Maxwell and Anderson, 1999)		
API (fourmis artificielles)	(Monmarché, 2000)*	(Aupetit, 2005) +Baum-Welch	(Aupetit, 2005)
OEP	(Rasmussen and T. Krink, 2003)	X +Baum-Welch	(Aupetit, 2005) X

## De nombreuses perspectives (3/4)

- ▶ MMCSS :
  - ▶ Une fois la loi de substitution définie, un MMCSS peut s'utiliser en lieu et place d'un MMC de façon transparente

## De nombreuses perspectives (3/4)

### ▶ MMCSS :

- ▶ Une fois la loi de substitution définie, un MMCSS peut s'utiliser en lieu et place d'un MMC de façon transparente
- ▶ Nous prévoyons d'appliquer les MMCSS là où les MMC sont actuellement utilisés :
  - ▶ Pour inclure des connaissances expertes là où on n'en exploite pas encore
  - ▶ Pour tenter d'améliorer les performances des SIA

## De nombreuses perspectives (3/4)

### ▶ MMCSS :

- ▶ Une fois la loi de substitution définie, un MMCSS peut s'utiliser en lieu et place d'un MMC de façon transparente
- ▶ Nous prévoyons d'appliquer les MMCSS là où les MMC sont actuellement utilisés :
  - ▶ Pour inclure des connaissances expertes là où on n'en exploite pas encore
  - ▶ Pour tenter d'améliorer les performances des SIA
- ▶ Étendre les applications au traitement d'images

## De nombreuses perspectives (3/4)

### ▶ MMCSS :

- ▶ Une fois la loi de substitution définie, un MMCSS peut s'utiliser en lieu et place d'un MMC de façon transparente
- ▶ Nous prévoyons d'appliquer les MMCSS là où les MMC sont actuellement utilisés :
  - ▶ Pour inclure des connaissances expertes là où on n'en exploite pas encore
  - ▶ Pour tenter d'améliorer les performances des SIA
- ▶ Étendre les applications au traitement d'images
- ▶ Appliquer les MMCSS pour le traitement du texte :
  - ▶ L'indexation automatique,
  - ▶ La comparaison de documents,
  - ▶ La détection de plagiat,
  - ▶ Le résumé automatique,
  - ▶ ...

## De nombreuses perspectives (4/4)

- ▶ La visualisation d'informations :
  - ▶ Pour des MMC :
    - ▶ de l'impact d'un ou plusieurs MMC sur des données
    - ▶ de l'espace des solutions pour les MMC
    - ▶ des populations des métaheuristiques AG, API et OEP dans l'espace des MMC

## De nombreuses perspectives (4/4)

- ▶ La visualisation d'informations :
  - ▶ Pour des MMC :
    - ▶ de l'impact d'un ou plusieurs MMC sur des données
    - ▶ de l'espace des solutions pour les MMC
    - ▶ des populations des métaheuristiques AG, API et OEP dans l'espace des MMC
  - ▶ La visualisation de dissimilarité :
    - ▶ Exploiter le produit scalaire pseudo-euclidien différemment

## De nombreuses perspectives (4/4)

- ▶ La visualisation d'informations :
  - ▶ Pour des MMC :
    - ▶ de l'impact d'un ou plusieurs MMC sur des données
    - ▶ de l'espace des solutions pour les MMC
    - ▶ des populations des métaheuristiques AG, API et OEP dans l'espace des MMC
  - ▶ La visualisation de dissimilarité :
    - ▶ Exploiter le produit scalaire pseudo-euclidien différemment
  - ▶ La visualisation pour l'aide au choix :
    - ▶ Travaux en cours pour le choix de plan de construction pour le projet HM2PH (thèse de A. Puret)

Merci de votre attention

















# La structure vectorielle $(\Omega, \oplus, \odot)$

## ▶ le cas particulier :

- ▶ soit  $\mathbb{G}_K^* \subset \mathbb{G}_K$  tel que toutes les composantes sont  $> 0$
- ▶ une fonction de régularisation :  $r_K(x)_i = x_i - \max_{j=1..K} x_j$
- ▶ soit  $\Omega_K = r_K(\mathbb{R}^K)$

- ▶ deux fonctions bijectives :  $\mathbb{G}_K^* \begin{array}{c} \xrightarrow{\psi_K} \\ \xleftarrow{\phi_K} \end{array} \Omega_K$

- ▶ deux opérateurs :

$$x \oplus_K y = y \oplus_K x = r_K(x + y)$$

$$c \odot_K x = x \odot_K c = r_K(c \cdot x)$$

- ▶  $(\Omega_K, \oplus_K, \odot_K)$  forme un espace vectoriel

## ▶ le cas général :

- ▶  $\Lambda^* \equiv \mathbb{G}_N^* \mathbb{G}_N^* \mathbb{G}_M^* \mathbb{G}_M^* \mathbb{G}_N^* \mathbb{G}_N^*$ ,  $\Omega \equiv \Omega_N \Omega_N^N \Omega_M^N \Omega_M^N$

- ▶  $\Lambda^* \begin{array}{c} \xrightarrow{\psi} \\ \xleftarrow{\phi} \end{array} \Omega$

- ▶  $(\Omega, \oplus, \odot)$  est un espace vectoriel







































Aupetit, S. (2005).

*Contributions aux modèles de Markov cachés : métaheuristiques d'apprentissage, nouveaux modèles et visualisation de dissimilarité.*

Thèse de doctorat, Laboratoire d'Informatique de l'Université François-Rabelais de Tours, Tours, France.

Aupetit, S., Monmarché, N., and Slimane, M. (2004).

Utilisation des chaînes de Markov cachées à substitutions de symboles pour l'apprentissage et la reconnaissance robuste d'images.

In *Manifestation des jeunes chercheurs en STIC (MajecSTIC)*, Calais.

Aupetit, S., Monmarché, N., and Slimane, M. (2005a).

Fouille visuelle de dissimilarités à l'aide de matrices de scatterplots pseudo-euclidiennes.

In Makarenkov, V., Cucumel, G., and Lapointe, F.-J., editors, *Comptes rendus des 12-ème rencontres de la Société Francophone de Classification (SFC'05)*, pages 43–46, UQAM, Quebec, Canada.

Aupetit, S., Monmarché, N., and Slimane, M. (2005b).

Visualisation de dissimilarités pour des Modèles de Markov Cachés.

In *6ème congrès de la Société Française de Recherche Opérationnelle et d'Aide à la Décision (ROADEF'05)*, pages 70–71, Tours, France.

Aupetit, S., Monmarché, N., Slimane, M., and Liardet, S. (2005c).

An Exponential Representation in the API Algorithm for Hidden Markov Models Training.

In *Proceedings of the 7th International Conference on Artificial Evolution (EA'05)*.  
to appear.

Aupetit, S., Slimane, M., and Monmarché, N. (2002).

Les modèles de Markov cachés à substitutions de symboles.

Technical Report 262, Laboratoire d'Informatique de l'Université François-Rabelais de Tours, Tours, France.

Baum, L. E. and Eagon, J. A. (1967).

An inequality with applications to statistical estimation for probabilistic functions of markov processes to a model for ecology.

*Bull American Mathematical Society*, 73 :360–363.

Brouard, T. (1999).

*Algorithmes hybrides d'apprentissage de chaînes de Markov cachées : conception et applications à la reconnaissance des formes.*

PhD thesis, Laboratoire d'Informatique de l'Université François-Rabelais de Tours.

Chen, T.-Y., Mei, X.-D., Pan, J.-S., and Sun, S.-H. (2004).

Optimization of hmm by the tabu search algorithm.