

Hidden Markov Model

RODON Jérémie

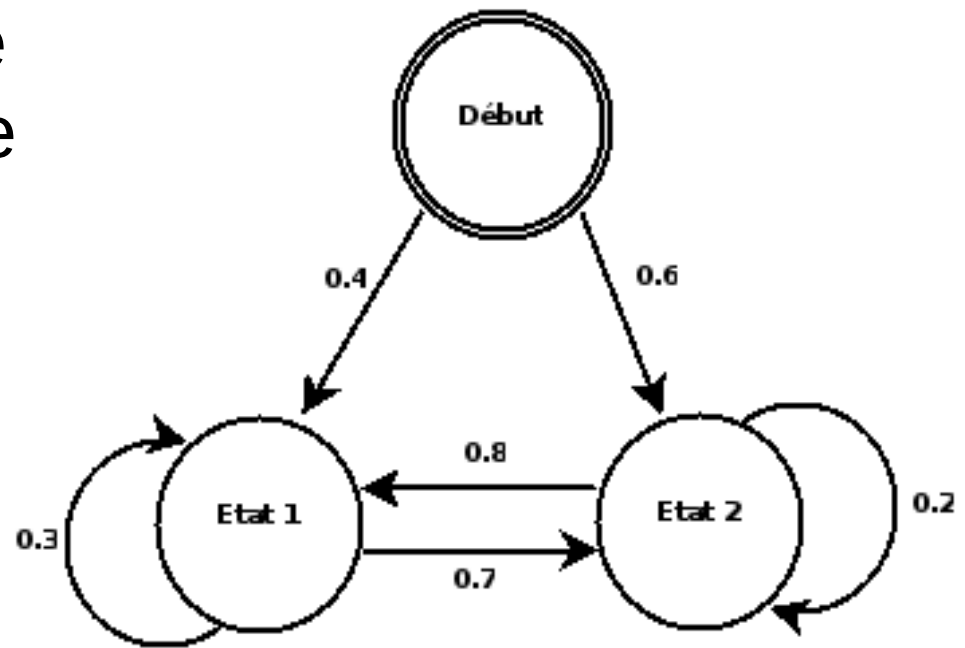
08/01/2013

Plan de la présentation

- Qu'est-ce qu'un HMM ?
 - Pour commencer : Markov Model
 - Hidden Markov Model
- Les trois questions majeures du HMM
 - Séquence d'états la plus probable
 - Probabilité d'une séquence d'observations
 - Maximiser la probabilité d'une séquence d'observations
- Modélisation et scalabilité
- A quoi peut servir un HMM ?

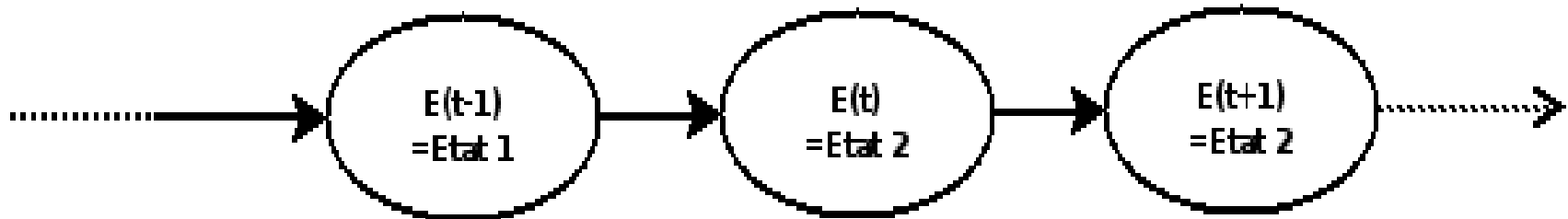
Pour commencer : Markov Model

- Un modèle (ou chaîne) de Markov se compose d'états et de transitions
- Chaque transition est associée à sa probabilité d'être empruntée
- On ajoute à cela les probabilités de départ



Pour commencer : Markov Model

- A chaque unité de temps, on opère une transition
- L'évolution dans le modèle génère une séquence d'états

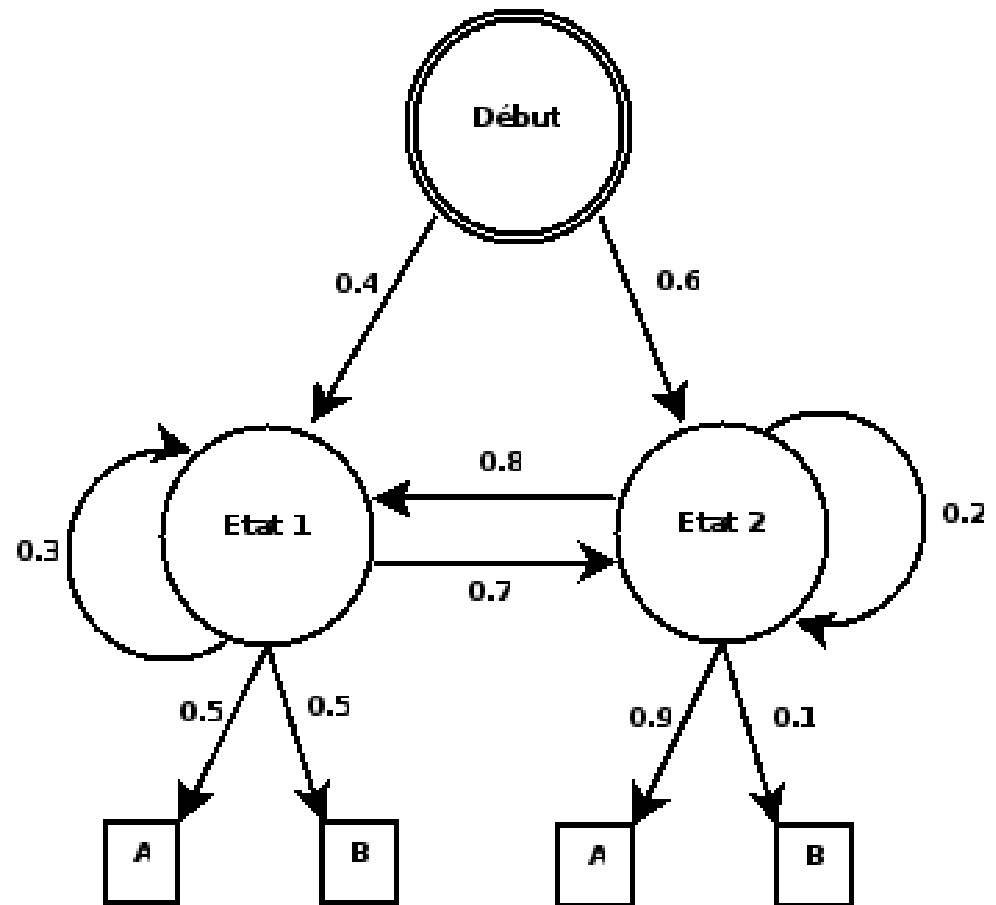


Hidden Markov Model

- Un modèle mathématique statistique
- Un modèle de Markov dont les états ne sont pas observables : il sont cachés.
- Les états du modèle émettent des «observations» sur lesquels nous pouvons travailler

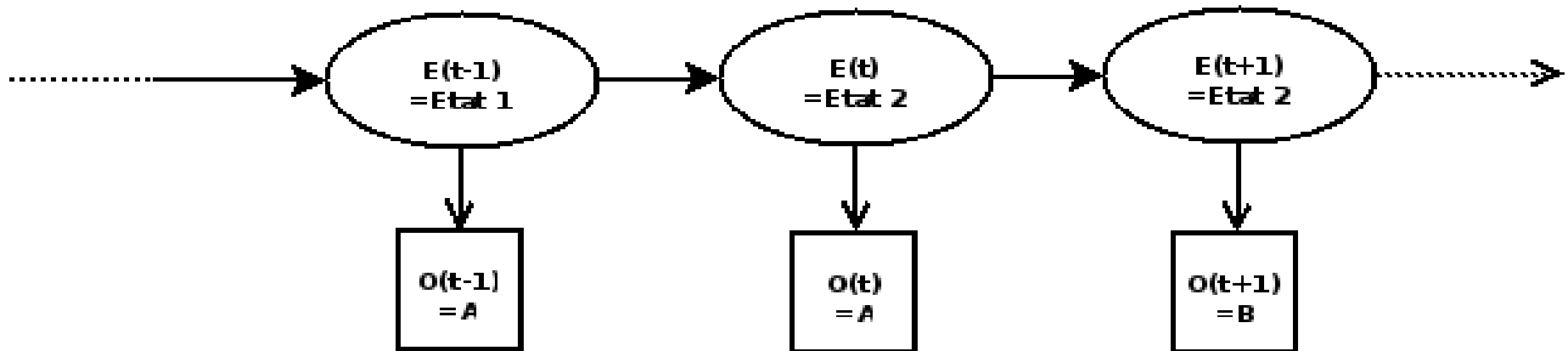
Hidden Markov Model

- Nous avons un modèle de Markov
- Chaque état se voit doté de probabilités d'émission de symboles
- Seuls ces symboles sont observables, les états sont cachés



Hidden Markov Model

- A chaque unité de temps, on opère une transition
- L'évolution dans le modèle génère une séquence d'observations émises par les états



Les trois questions majeures du HMM

Étant donné un HMM et une séquence de symboles (d'observations) :

- ☐ Quelle est la séquence d'état la plus probable qui a généré cette séquence de symboles ?
- ☐ Quelle est la probabilité de voir apparaître cette séquence de symboles ?
- ☐ Comment paramétrer le HMM pour que la probabilité d'apparition de cette séquence soit maximale ?

Séquence d'états la plus probable

- On a une séquence d'observations, on connaît le HMM qui l'a générée et on souhaite retrouver la séquence d'état génératrice
- Il n'y a pas qu'une seule réponse, mais il y en a une plus probable que les autres
- L'*algorithme de Viterbi* permet de trouver cette séquence en $O(N^2T)$ (N états, T observations)

Séquence d'états la plus probable

- Cette question a deux interprétations possibles :
 - Soit on cherche la séquence d'états qui présente la plus haute probabilité de génération de la séquence d'observations
 - Soit on cherche la séquence d'états qui maximise le nombre d'états corrects espérés

Les séquences répondant à chacune de ces interprétations ne sont pas forcément les mêmes

- L'*algorithme de Viterbi* peut être adapté pour ces deux sous-questions

Probabilité d'une séquence d'observations

- On a une séquence d'observation et on souhaite connaître la probabilité d'apparition de cette séquence en fonction d'un HMM donné
- Cette question à une réponse unique
- Le *forward-algorithm* permet de trouver cette probabilité en $O(N^2T)$ (N états, T observations)

Maximiser la probabilité d'une séquence d'observations

- On a une séquence d'observations et on veut obtenir le HMM qui a la meilleure probabilité d'émission de cette séquence
- Il n'existe pas d'algorithme permettant de résoudre ce problème de manière sûre et optimale
- Mais on peut s'approcher de la solution grâce à l'*algorithme de Baum-Welch*

Maximiser la probabilité d'une séquence d'observations

- L'algorithme est itératif, chaque tour de boucle améliorant la probabilité d'émission de la séquence
 - On commence par initialiser le HMM
 - On réévalue le HMM en fonction de la séquence d'observation
 - On calcule la probabilité d'émission de la séquence
 - Si elle augmente, on refait un tour
- La complexité de chaque tour est $O(N^2T)$

Maximiser la probabilité d'une séquence d'observations

- L'algorithme est itératif, chaque tour de boucle améliorant la probabilité d'émission de la séquence
 - On commence par initialiser le HMM
 - On réévalue le HMM en fonction de la séquence d'observation
 - On calcule la probabilité d'émission de la séquence
 - Si elle augmente, on refait un tour
- La complexité de chaque tour est $O(N^2T)$

Modélisation

Pour modéliser un HMM avec N états et M observations, on définit 3 matrices :

- Start, de taille $1 \times N$, contenant les probabilités de départ
- Transitions, de taille $N \times N$, contenant les probabilités de transition entre les états
- Emissions, de taille $N \times M$, contenant les probabilités d'émission de chacun des états

Ces trois matrices ont des lignes « stochastiques », i.e. Chaque élément est une probabilité et la somme des éléments est 1

Scalabilité

- Les algorithmes traitant les HMM impliquent de multiplier des probabilités entre elles
- On peut très vite obtenir des nombres très petits, impossibles à représenter par des flottants
- Le contournement classique consiste à utiliser les logarithmes : $a*b = \exp(\ln(a)+\ln(b))$
- Mais l'imprécision des flottants produit tout de même des erreurs
- Idéalement il faut du matériel spécialisé ou beaucoup de temps...

A quoi peut servir un HMM ?

- Tout ce qui implique de retrouver une séquence « cachée » génératrice d'une séquence observable
 - Exemple : La séquence observable est l'enregistrement de paroles, la séquence cachée les mots qui l'ont produite
- On peut l'entraîner à reconnaître « quelque chose », puis l'utiliser pour vérifier si « autre chose » y ressemble
 - Exemple : Vérifier si un texte est en français ou pas

Webographie

Les documents ayant servi à élaborer cet exposé (en anglais) :

- http://en.wikipedia.org/wiki/Hidden_Markov_model
- <http://www.cs.sjsu.edu/~stamp/RUA/HMM.pdf>
- http://en.wikipedia.org/wiki/Viterbi_algorithm

Merci !

Des questions ?