Algorithme de VITERBI

Problème de plus long chemin

Sébastien Lagrange
LARIS - Universite d'Angers - France
http://perso-laris.univ-angers.fr/~lagrange/

Polytech Angers

24 mai 2022

Plan

Rappels sur les probabilités conditionnelles

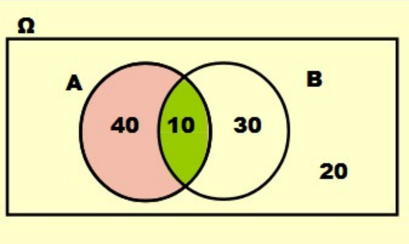
Définition

- $P(A|B) = \frac{P(A \cap B)}{P(B)}$

Rappels sur les probabilités conditionnelles

Définition

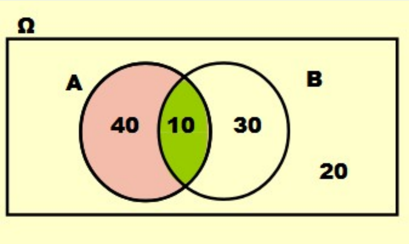
- $P(A|B) = \frac{P(A \cap B)}{P(B)}$



Rappels sur les probabilités conditionnelles

Définition

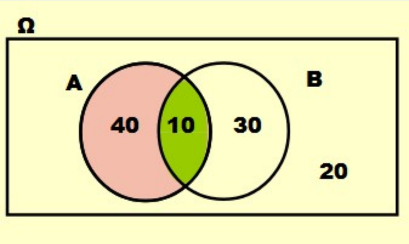
- $P(A|B) = \frac{P(A \cap B)}{P(B)}$
- $P(A \cap B) = P(B)P(A|B) = P(A)P(B|A)$



Rappels sur les probabilités conditionnelles

Définition

- $P(A|B) = \frac{P(A \cap B)}{P(B)}$
- $P(A \cap B) = P(B)P(A|B) = P(A)P(B|A)$
- on note $P(A, B) = P(A \cap B)$



Rappels sur les chaînes de Markov

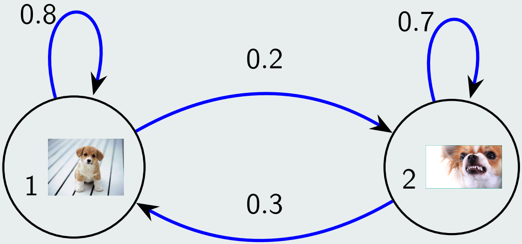
Définition

- $X = \{1, \ldots, N\}$ l'ensemble des N états du modèle
- $\Pi = (\pi_1, \ldots, \pi_N)$ un vecteur avec π_i la probabilité de commencer dans l'état i
- $P = \{p_{ij}\}_{1 \leq i,j \leq N}$ matrice de Markov. (p_{ij} est la probabilité de passer de l'état i à l'état j notée aussi $Pr(i, j)$)

Rappels sur les chaînes de Markov

Exemple

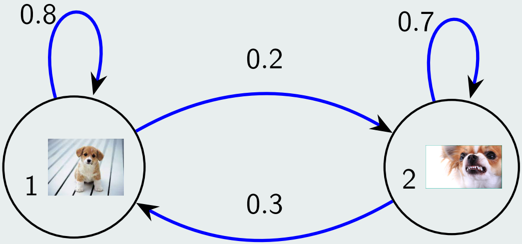
$X = \{1, 2\}$,
$\Pi = (0.5, 0.5)$,
$P = \begin{pmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{pmatrix}$



Rappels sur les chaînes de Markov

Exemple

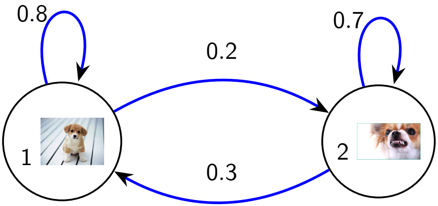
$X = \{1, 2\}$,
$\Pi = (0.5, 0.5)$,
$$P = \begin{pmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{pmatrix}$$



Propriétés d'une chaîne de Markov

- $Pr(x_k = j | x_{k-1} = i, x_{k-2} = l, \ldots) = Pr(x_k = j | x_{k-1} = i)$ (ordre 1),
- $Pr(x_k = j | x_{k-1} = i) = Pr(x_{k+1} = j | x_k = i)$ (stationnarité).

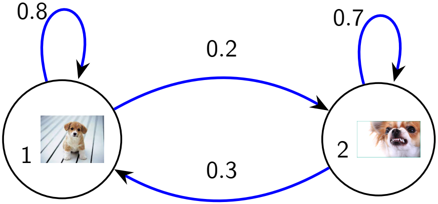
Les chaînes de Markov



Question classique

Calculer la probabilité d'être dans chaque état après $K = 1$ itérations.

Les chaînes de Markov

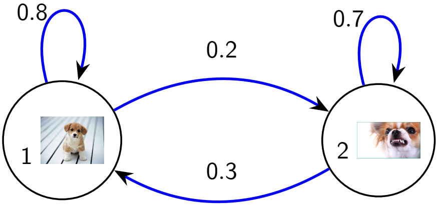


Question classique

Calculer la probabilité d'être dans chaque état après $K = 1$ itérations.

$$Pr(x_1 = 1) = \pi_1 \cdot p_{1,1} + \pi_2 \cdot p_{2,1} = 0.5 \times 0.8 + 0.5 \times 0.3$$

Les chaînes de Markov

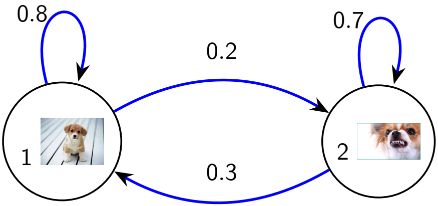


Question classique

Calculer la probabilité d'être dans chaque état après $K = 1$ itérations.

$$Pr(x_1 = 1) = \pi_1 \cdot p_{1,1} + \pi_2 \cdot p_{2,1} = 0.5 \times 0.8 + 0.5 \times 0.3$$
$$Pr(x_1 = 2) = \pi_1 \cdot p_{1,2} + \pi_2 \cdot p_{2,2} = 0.5 \times 0.2 + 0.5 \times 0.7$$

Les chaînes de Markov



Question classique

Calculer la probabilité d'être dans chaque état après $K = 1$ itérations.

$$Pr(x_1 = 1) = \pi_1 \cdot p_{1,1} + \pi_2 \cdot p_{2,1} = 0.5 \times 0.8 + 0.5 \times 0.3$$

$$Pr(x_1 = 2) = \pi_1 \cdot p_{1,2} + \pi_2 \cdot p_{2,2} = 0.5 \times 0.2 + 0.5 \times 0.7$$

$$\begin{pmatrix} Pr(x_1 = 1) & Pr(x_1 = 2) \end{pmatrix} = \underbrace{\begin{pmatrix} \pi_1 & \pi_2 \end{pmatrix}}_{\Pi} \underbrace{\begin{pmatrix} p_{1,1} & p_{1,2} \\ p_{2,1} & p_{2,2} \end{pmatrix}}$$

Les chaînes de Markov

Exemple

Calculer la probabilité d'être chaque état après $K = 2$ itérations.

Les chaînes de Markov

Exemple

Calculer la probabilité d'être chaque état après $K = 2$ itérations.

$$\begin{pmatrix} Pr(x_2 = 1) & Pr(x_2 = 2) \end{pmatrix} = \begin{pmatrix} \pi_1 & \pi_2 \end{pmatrix} \begin{pmatrix} p_{1,1} & p_{1,2} \\ p_{2,1} & p_{2,2} \end{pmatrix} \begin{pmatrix} p_{1,1} & p_{1,2} \\ p_{2,1} & p_{2,2} \end{pmatrix}$$

Les chaînes de Markov

Exemple

Calculer la probabilité d'être chaque état après $K = 2$ itérations.

$$\begin{pmatrix} Pr(x_2 = 1) & Pr(x_2 = 2) \end{pmatrix} = \begin{pmatrix} \pi_1 & \pi_2 \end{pmatrix} \begin{pmatrix} p_{1,1} & p_{1,2} \\ p_{2,1} & p_{2,2} \end{pmatrix} \begin{pmatrix} p_{1,1} & p_{1,2} \\ p_{2,1} & p_{2,2} \end{pmatrix}$$

$$= \begin{pmatrix} \pi_1 & \pi_2 \end{pmatrix} \begin{pmatrix} p_{1,1}p_{1,1} + p_{1,2}p_{2,1} & p_{1,2}p_{1,1} + p_{1,2}p_{2,2} \\ p_{2,1}p_{1,1} + p_{2,2}p_{2,1} & p_{2,2}p_{1,1} + p_{2,2}p_{2,2} \end{pmatrix}$$

Les chaînes de Markov

Exemple

Calculer la probabilité d'être chaque état après $K = 2$ itérations.

$$\begin{pmatrix} Pr(x_2 = 1) & Pr(x_2 = 2) \end{pmatrix} = \begin{pmatrix} \pi_1 & \pi_2 \end{pmatrix} \begin{pmatrix} p_{1,1} & p_{1,2} \\ p_{2,1} & p_{2,2} \end{pmatrix} \begin{pmatrix} p_{1,1} & p_{1,2} \\ p_{2,1} & p_{2,2} \end{pmatrix}$$

$$= \begin{pmatrix} \pi_1 & \pi_2 \end{pmatrix} \begin{pmatrix} p_{1,1}p_{1,1} + p_{1,2}p_{2,1} & p_{1,2}p_{1,1} + p_{1,2}p_{2,2} \\ p_{2,1}p_{1,1} + p_{2,2}p_{2,1} & p_{2,2}p_{1,1} + p_{2,2}p_{2,2} \end{pmatrix}$$

Remarque

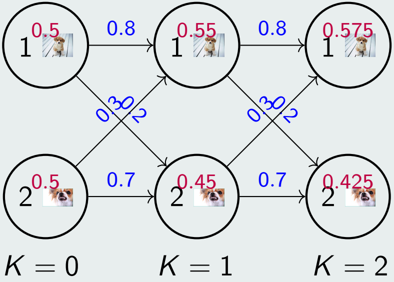
$P =$   $\begin{pmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{pmatrix}$,

$P^2 =$   $\begin{pmatrix} 0.7 & 0.3 \\ 0.45 & 0.55 \end{pmatrix} \dots$

Les chaînes de Markov

Vision Treillis



$$\Pi P = \begin{pmatrix} 0.5 & 0.5 \end{pmatrix} \begin{pmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{pmatrix} = \begin{pmatrix} 0.55 & 0.45 \end{pmatrix}$$

$$\Pi P^2 = \begin{pmatrix} 0.5 & 0.5 \end{pmatrix} \begin{pmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{pmatrix}^2 = \begin{pmatrix} 0.575 & 0.425 \end{pmatrix}$$

Les chaînes de Markov

Proposition

Le vecteur contenant la probabilité d'être dans l'état i après K itérations observations est

$$\begin{pmatrix} Pr(x_k = 1) & \dots & Pr(x_k = N) \end{pmatrix} = \Pi P^K$$

Les chaînes de Markov

Proposition

Le vecteur contenant la probabilité d'être dans l'état i après K itérations observations est

$$\begin{pmatrix} Pr(x_k = 1) & \dots & Pr(x_k = N) \end{pmatrix} = \Pi P^K$$

Théorème

Pour $K \to \infty$, ΠP^K converge vers un vecteur propre associé à la valeur propre 1.

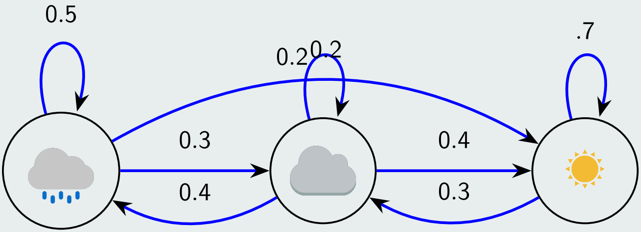
Modèle de Markov Caché

Définition

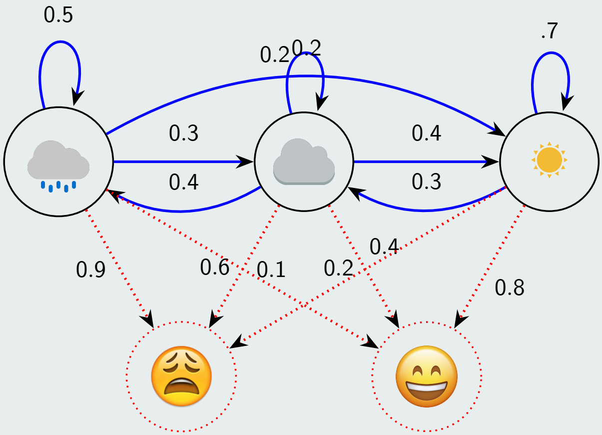
un Modèle de Markov Caché est défini par

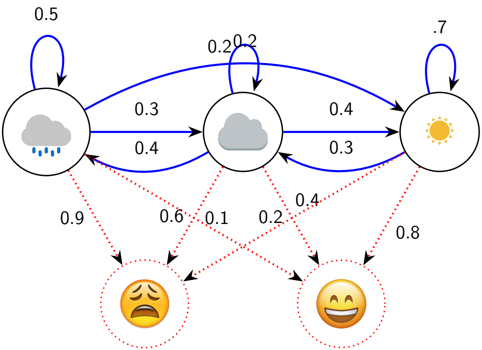
- $X = \{x_1, \ldots, s_N\}$ l'ensemble des N états du modèle
- $Y = \{y_1, \ldots, y_M\}$ l'ensemble des M mesures.
- $\Pi = (\pi_1, \ldots, \pi_N)$ un vecteur avec π_i la probabilité de commencer dans l'état i
- $P = \{p_{ij}\}_{1 \leq i,j \leq N}$ matrice de transmission.
- $G = \{g_{jm}\}_{1 \leq j \leq N, 1 \leq m \leq K}$ matrice d'émission.(g_{jm} est la probabilité de mesurer m dans l'état j)

Exemple



Exemple





Exemple

$$P = \begin{pmatrix} 0.5 & 0.3 & 0.2 \\ 0.4 & 0.2 & 0.4 \\ 0.0 & 0.3 & 0.7 \end{pmatrix} \quad G = \begin{pmatrix} 0.9 & 0.1 \\ 0.6 & 0.4 \\ 0.2 & 0.8 \end{pmatrix} \quad \Pi = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

Matrice de transmission — Matrice d'emission

Objectif

Connaissant une suite d'observations, trouver la séquence d'états la plus probable.

- Méthode naïve
- L'algorithme de Viterbi

Problématique

?	⟶	?	⟶	?
↓		↓		↓
😄		😄		😩

	🌧	☁	☀
🌧	0.5	0.3	0.2
☁	0.4	0.2	0.4
☀	0.0	0.3	0.7

	😩	😄
🌧	0.9	0.1
☁	0.6	0.4
☀	0.2	0.8

Problématique

$$\begin{array}{ccccc} ? & \longrightarrow & ? & \longrightarrow & ? \\ \downarrow & & \downarrow & & \downarrow \\ 😄 & & 😄 & & 😩 \end{array}$$

$$\begin{array}{c} \\ 🌧 \\ ☁ \\ ☀ \end{array}\begin{array}{c} \begin{array}{ccc} 🌧 & ☁ & ☀ \end{array} \\ \begin{pmatrix} 0.5 & 0.3 & 0.2 \\ 0.4 & 0.2 & 0.4 \\ 0.0 & 0.3 & 0.7 \end{pmatrix} \end{array} \begin{array}{c} \\ 🌧 \\ ☁ \\ ☀ \end{array}\begin{array}{c} \begin{array}{cc} 😩 & 😄 \end{array} \\ \begin{pmatrix} 0.9 & 0.1 \\ 0.6 & 0.4 \\ 0.2 & 0.8 \end{pmatrix} \end{array}$$

Trouver la séquence d'états la plus probable

$$\arg\max_{x_0,x_1,x_2} Pr(X_0 = x_0, X_1 = x_1, X_2 = x_2 | Y_0 = 😄, Y_1 = 😄, Y_2 = 😩)$$

Calcul de la séquence la plus probable

Simplifions $Pr(x_0, x_1, x_2 | y_0 =$ 😄$, y_1 =$ 😄$, y_2 =$ 😩$)$

Calcul de la séquence la plus probable

Simplifions $Pr(x_0, x_1, x_2 | y_0 = 😄, y_1 = 😄, y_2 = 😩)$

- $Pr(x_0, x_1, x_2 | y_0, y_1, y_2) = \frac{Pr(y_0, y_1, y_2, x_0, x_1, x_2)}{Pr(y_0, y_1, y_2)}$

Calcul de la séquence la plus probable

Simplifions $Pr(x_0, x_1, x_2 | y_0 =$ 😄, $y_1 =$ 😄, $y_2 =$ 😩$)$

- $Pr(x_0, x_1, x_2 | y_0, y_1, y_2) = \frac{Pr(y_0, y_1, y_2, x_0, x_1, x_2)}{Pr(y_0, y_1, y_2)}$
- $Pr(x_0, x_1, x_2, y_0, y_1, y_2) =$

Calcul de la séquence la plus probable

Simplifions $Pr(x_0, x_1, x_2 | y_0 =$ 😄$, y_1 =$ 😄$, y_2 =$ 😩$)$

- $Pr(x_0, x_1, x_2 | y_0, y_1, y_2) = \frac{Pr(y_0, y_1, y_2, x_0, x_1, x_2)}{Pr(y_0, y_1, y_2)}$
- $Pr(x_0, x_1, x_2, y_0, y_1, y_2) =$
  $Pr(x_0) Pr(x_1, x_2, y_0, y_1, y_2 | x_0) =$

Calcul de la séquence la plus probable

Simplifions $Pr(x_0, x_1, x_2 | y_0 = 😁, y_1 = 😁, y_2 = 😩)$

- $Pr(x_0, x_1, x_2 | y_0, y_1, y_2) = \frac{Pr(y_0, y_1, y_2, x_0, x_1, x_2)}{Pr(y_0, y_1, y_2)}$
- $Pr(x_0, x_1, x_2, y_0, y_1, y_2) =$
  $Pr(x_0) Pr(x_1, x_2, y_0, y_1, y_2 | x_0) =$
  $Pr(x_0) Pr(x_1 | x_0) Pr(x_2, y_0, y_1, y_2 | x_0, x_1) =$

Calcul de la séquence la plus probable

Simplifions $Pr(x_0, x_1, x_2 | y_0 = \text{😄}, y_1 = \text{😄}, y_2 = \text{😩})$

- $Pr(x_0, x_1, x_2 | y_0, y_1, y_2) = \frac{Pr(y_0, y_1, y_2, x_0, x_1, x_2)}{Pr(y_0, y_1, y_2)}$
- $Pr(x_0, x_1, x_2, y_0, y_1, y_2) =$
  $Pr(x_0)Pr(x_1, x_2, y_0, y_1, y_2 | x_0) =$
  $Pr(x_0)Pr(x_1 | x_0)Pr(x_2, y_0, y_1, y_2 | x_0, x_1) =$
  $Pr(x_0)Pr(x_1 | x_0)Pr(x_2 | x_0, x_1)Pr(y_0, y_1, y_2 | x_0, x_1, x_2) =$

Calcul de la séquence la plus probable

Simplifions $Pr(x_0, x_1, x_2 | y_0 = 😄, y_1 = 😄, y_2 = 😩)$

- $Pr(x_0, x_1, x_2 | y_0, y_1, y_2) = \frac{Pr(y_0, y_1, y_2, x_0, x_1, x_2)}{Pr(y_0, y_1, y_2)}$
- $Pr(x_0, x_1, x_2, y_0, y_1, y_2) =$
  $Pr(x_0)Pr(x_1, x_2, y_0, y_1, y_2 | x_0) =$
  $Pr(x_0)Pr(x_1|x_0)Pr(x_2, y_0, y_1, y_2 | x_0, x_1) =$
  $Pr(x_0)Pr(x_1|x_0)Pr(x_2|x_0, x_1)Pr(y_0, y_1, y_2 | x_0, x_1, x_2) =$

  $Pr(x_0)Pr(x_1|x_0)Pr(x_2|x_1)Pr(y_0|x_0, x_1, x_2)Pr(y_1, y_2 | x_0, x_1, x_2, y_0) =$

Calcul de la séquence la plus probable

Simplifions $Pr(x_0, x_1, x_2 | y_0 = 😄, y_1 = 😄, y_2 = 😩)$

- $Pr(x_0, x_1, x_2 | y_0, y_1, y_2) = \frac{Pr(y_0, y_1, y_2, x_0, x_1, x_2)}{Pr(y_0, y_1, y_2)}$
- $Pr(x_0, x_1, x_2, y_0, y_1, y_2) =$

  $Pr(x_0)Pr(x_1, x_2, y_0, y_1, y_2 | x_0) =$

  $Pr(x_0)Pr(x_1 | x_0)Pr(x_2, y_0, y_1, y_2 | x_0, x_1) =$

  $Pr(x_0)Pr(x_1 | x_0)Pr(x_2 | x_0, x_1)Pr(y_0, y_1, y_2 | x_0, x_1, x_2) =$

  $Pr(x_0)Pr(x_1 | x_0)Pr(x_2 | x_1)Pr(y_0 | x_0, x_1, x_2)Pr(y_1, y_2 | x_0, x_1, x_2, y_0) =$

  $Pr(x_0)Pr(x_1 | x_0)Pr(x_2 | x_1)Pr(y_0 | x_0)Pr(y_1 | x_0, x_1, x_2, y_0)Pr(y_2 | x_0, x_1, x_2$

Calcul de la séquence la plus probable

Simplifions $Pr(x_0, x_1, x_2 | y_0 =$ 😁$, y_1 =$ 😁$, y_2 =$ 😩$)$

- $Pr(x_0, x_1, x_2 | y_0, y_1, y_2) = \frac{Pr(y_0, y_1, y_2, x_0, x_1, x_2)}{Pr(y_0, y_1, y_2)}$
- $Pr(x_0, x_1, x_2, y_0, y_1, y_2) =$
  $Pr(x_0)Pr(x_1, x_2, y_0, y_1, y_2 | x_0) =$
  $Pr(x_0)Pr(x_1|x_0)Pr(x_2, y_0, y_1, y_2 | x_0, x_1) =$
  $Pr(x_0)Pr(x_1|x_0)Pr(x_2|x_0, x_1)Pr(y_0, y_1, y_2 | x_0, x_1, x_2) =$

  $Pr(x_0)Pr(x_1|x_0)Pr(x_2|x_1)Pr(y_0|x_0, x_1, x_2)Pr(y_1, y_2 | x_0, x_1, x_2, y_0) =$

  $Pr(x_0)Pr(x_1|x_0)Pr(x_2|x_1)Pr(y_0|x_0)Pr(y_1|x_0, x_1, x_2, y_0)Pr(y_2|x_0, x_1, x_2$
  $Pr(x_0)Pr(x_1|x_0)Pr(x_2|x_1)Pr(y_0|x_0)Pr(y_1|x_1)Pr(y_2|x_2)$

Calcul de la séquence la plus probable

Simplifions $Pr(x_0, x_1, x_2 | y_0 =$ 😄$, y_1 =$ 😄$, y_2 =$ 😩$)$

- $Pr(x_0, x_1, x_2 | y_0, y_1, y_2) = \frac{Pr(y_0, y_1, y_2, x_0, x_1, x_2)}{Pr(y_0, y_1, y_2)}$
- $Pr(x_0, x_1, x_2, y_0, y_1, y_2) =$
  $Pr(x_0)Pr(x_1, x_2, y_0, y_1, y_2 | x_0) =$
  $Pr(x_0)Pr(x_1|x_0)Pr(x_2, y_0, y_1, y_2 | x_0, x_1) =$
  $Pr(x_0)Pr(x_1|x_0)Pr(x_2|x_0, x_1)Pr(y_0, y_1, y_2 | x_0, x_1, x_2) =$

  $Pr(x_0)Pr(x_1|x_0)Pr(x_2|x_1)Pr(y_0|x_0, x_1, x_2)Pr(y_1, y_2 | x_0, x_1, x_2, y_0) =$

  $Pr(x_0)Pr(x_1|x_0)Pr(x_2|x_1)Pr(y_0|x_0)Pr(y_1|x_0, x_1, x_2, y_0)Pr(y_2|x_0, x_1, x_2$
  $Pr(x_0)Pr(x_1|x_0)Pr(x_2|x_1)Pr(y_0|x_0)Pr(y_1|x_1)Pr(y_2|x_2)$

$$\arg\max_{x_0, x_1, x_2} \prod_i \underbrace{Pr(x_{i+1}|x_i)}_{transmission} \underbrace{Pr(y_i|x_i)}_{emission}$$

Exemple : Calcul de la probabilité d'une séquence donnée

☀ ⟶ ☁ ⟶ ☀
↓ ↓ ↓
😄 😄 😩

	🌧	☁	☀
🌧	0.5	0.3	0.2
☁	0.4	0.2	0.4
☀	0.0	0.3	0.7

	😩	😄
🌧	0.9	0.1
☁	0.6	0.4
☀	0.2	0.8

$Pr(x_0 =$ ☀, $x_1 =$ ☁, $x_2 =$ ☀, $y_0 =$ 😄, $y_1 =$ 😄, $y_2 =$ 😩$)$

$=$ $Pr($☀$)$ $Pr($😄$| =$ ☀$)$ $Pr($☁$|$☀$)$ $Pr($😄$|$☀$)$ $Pr($☀$|$☁$)$ $Pr($😩$|$☀$)$

$= \frac{1}{3} \cdot 0.8 \cdot 0.3 \cdot 0.4 \cdot 0.4 \cdot 0.2$

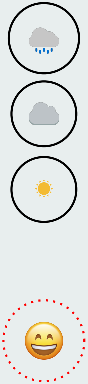
On sait calculer la probabilité d'une séquence d'états donnés :

- $Pr(x_0 =$ ☀️, $x_1 =$ ☀️, $x_2 =$ ☀️, $y_0 =$ 😄, $y_1 =$ 😄, $y_2 =$ 😩$)$
- $Pr(x_0 =$ ☁️, $x_1 =$ ☀️, $x_2 =$ ☀️, $y_0 =$ 😄, $y_1 =$ 😄, $y_2 =$ 😩$)$
- $Pr(x_0 =$ 🌧️, $x_1 =$ ☀️, $x_2 =$ ☀️, $y_0 =$ 😄, $y_1 =$ 😄, $y_2 =$ 😩$)$
- $Pr(x_0 =$ ☀️, $x_1 =$ ☁️, $x_2 =$ ☀️, $y_0 =$ 😄, $y_1 =$ 😄, $y_2 =$ 😩$)$
- $Pr(x_0 =$ ☁️, $x_1 =$ ☁️, $x_2 =$ ☀️, $y_0 =$ 😄, $y_1 =$ 😄, $y_2 =$ 😩$)$
- $Pr(x_0 =$ 🌧️, $x_1 =$ ☁️, $x_2 =$ ☀️, $y_0 =$ 😄, $y_1 =$ 😄, $y_2 =$ 😩$)$
- ...

Calcul de la séquence la plus probable

Complexité $O(N^K)$

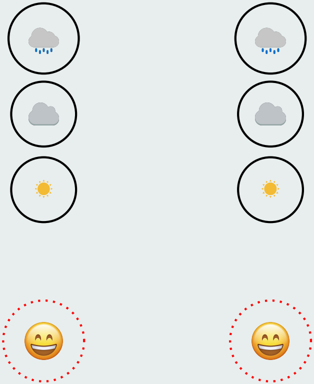
Vision Treillis



	🌧️	☁️	☀️
🌧️	0.5	0.3	0.2
☁️	0.4	0.2	0.4
☀️	0.0	0.3	0.7

	😩	😄
🌧️	0.9	0.1
☁️	0.6	0.4
☀️	0.2	0.8

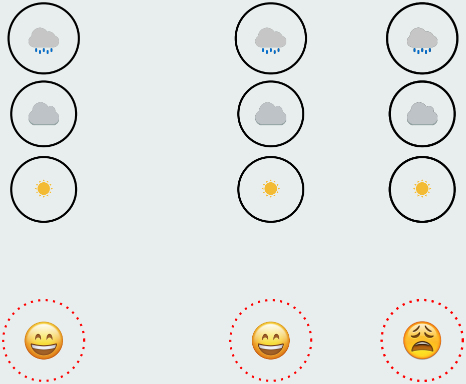
Vision Treillis



	🌧️	☁️	☀️
🌧️	0.5	0.3	0.2
☁️	0.4	0.2	0.4
☀️	0.0	0.3	0.7

	😩	😄
🌧️	0.9	0.1
☁️	0.6	0.4
☀️	0.2	0.8

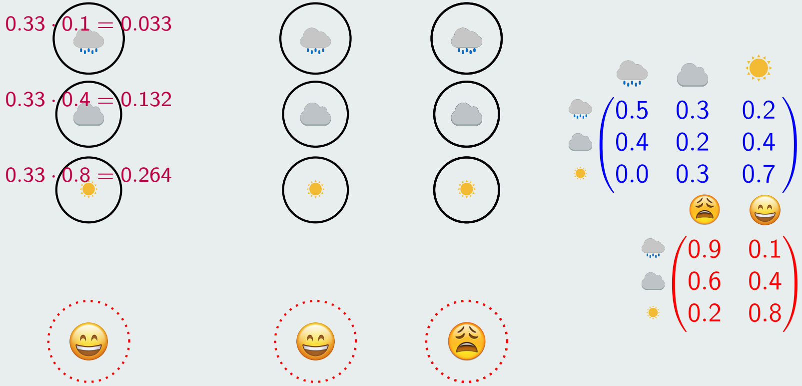
Vision Treillis



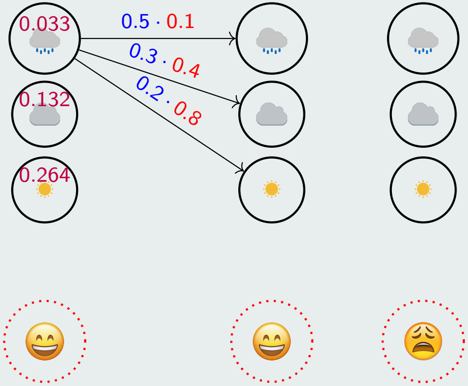
	🌧	☁	☀
🌧	0.5	0.3	0.2
☁	0.4	0.2	0.4
☀	0.0	0.3	0.7

	😩	😄
🌧	0.9	0.1
☁	0.6	0.4
☀	0.2	0.8

Vision Treillis



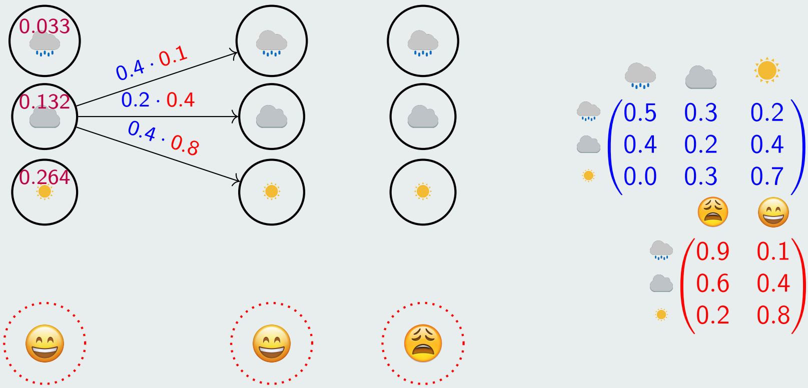
Vision Treillis



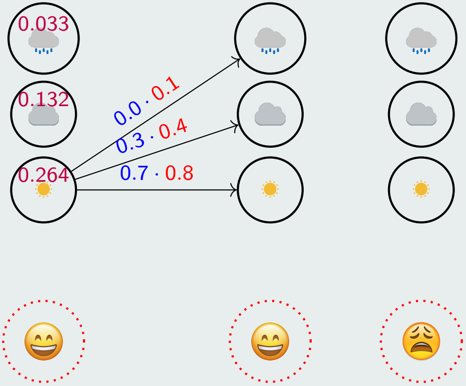
$$\begin{pmatrix} 0.5 & 0.3 & 0.2 \\ 0.4 & 0.2 & 0.4 \\ 0.0 & 0.3 & 0.7 \end{pmatrix}$$

$$\begin{pmatrix} 0.9 & 0.1 \\ 0.6 & 0.4 \\ 0.2 & 0.8 \end{pmatrix}$$

Vision Treillis



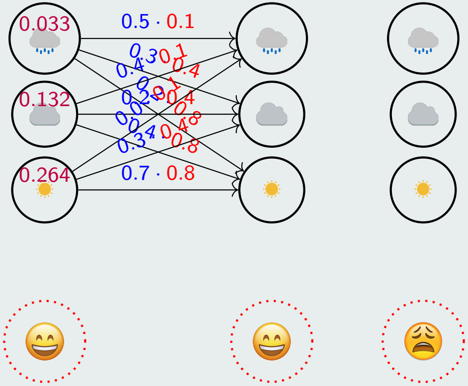
Vision Treillis



$$\begin{pmatrix} 0.5 & 0.3 & 0.2 \\ 0.4 & 0.2 & 0.4 \\ 0.0 & 0.3 & 0.7 \end{pmatrix}$$

$$\begin{pmatrix} 0.9 & 0.1 \\ 0.6 & 0.4 \\ 0.2 & 0.8 \end{pmatrix}$$

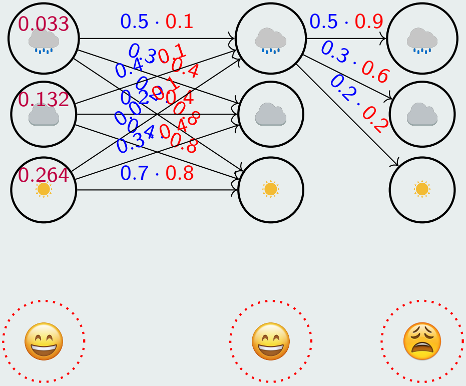
Vision Treillis



$$\begin{pmatrix} 0.5 & 0.3 & 0.2 \\ 0.4 & 0.2 & 0.4 \\ 0.0 & 0.3 & 0.7 \end{pmatrix}$$

$$\begin{pmatrix} 0.9 & 0.1 \\ 0.6 & 0.4 \\ 0.2 & 0.8 \end{pmatrix}$$

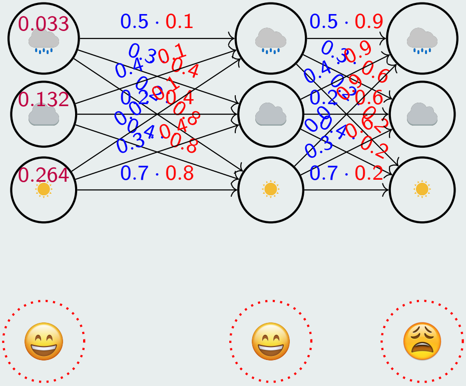
Vision Treillis



	🌧	☁	☀
🌧	0.5	0.3	0.2
☁	0.4	0.2	0.4
☀	0.0	0.3	0.7

	😩	😄
🌧	0.9	0.1
☁	0.6	0.4
☀	0.2	0.8

Vision Treillis

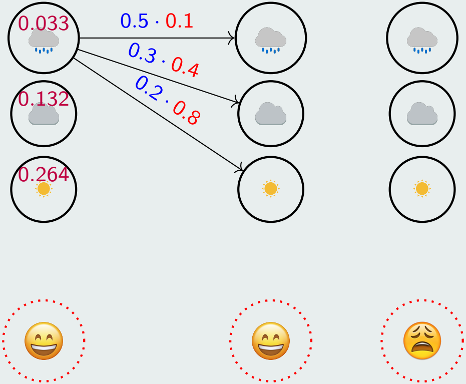


$$\begin{pmatrix} 0.5 & 0.3 & 0.2 \\ 0.4 & 0.2 & 0.4 \\ 0.0 & 0.3 & 0.7 \end{pmatrix}$$

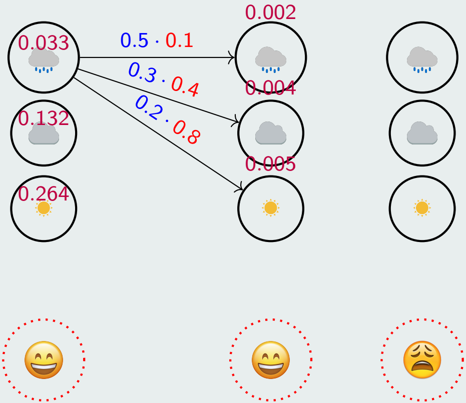
$$\begin{pmatrix} 0.9 & 0.1 \\ 0.6 & 0.4 \\ 0.2 & 0.8 \end{pmatrix}$$

Algorithme de Viterbi

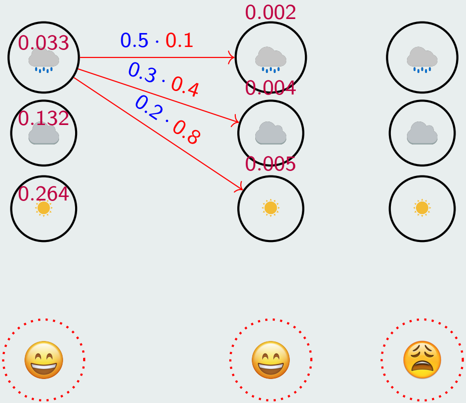
Algorithme de VITERBI



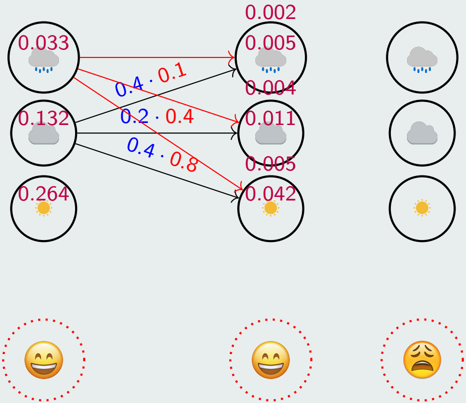
Algorithme de VITERBI



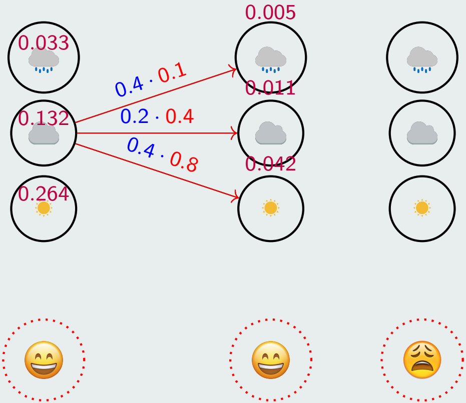
Algorithme de VITERBI



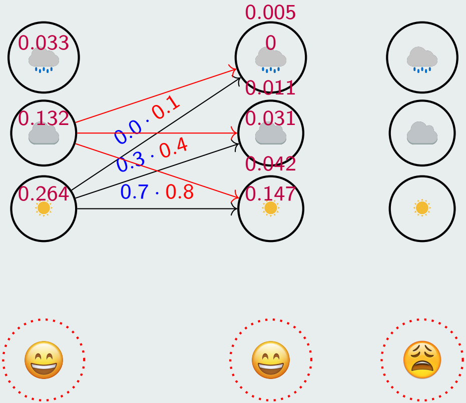
Algorithme de VITERBI



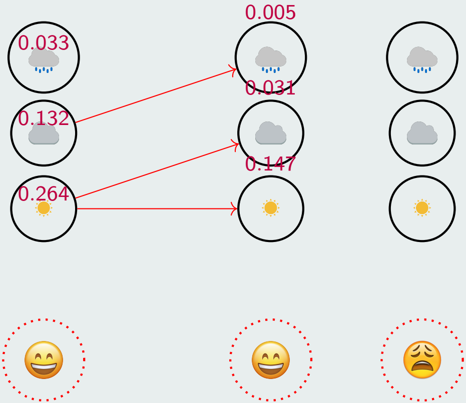
Algorithme de VITERBI



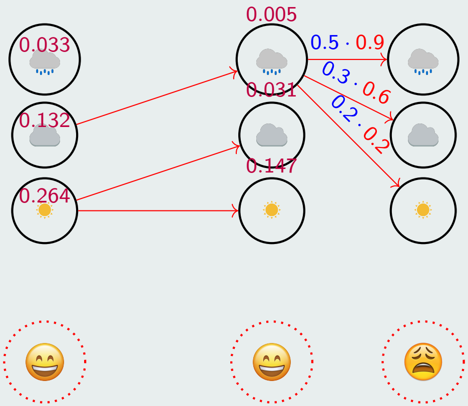
Algorithme de VITERBI



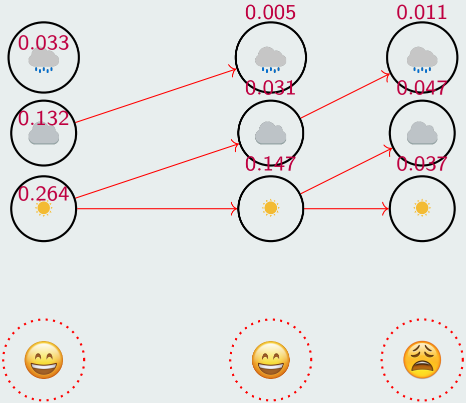
Algorithme de VITERBI



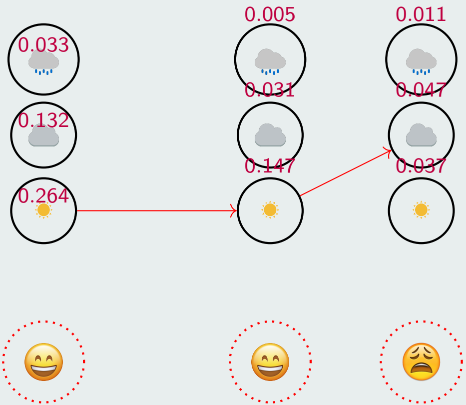
Algorithme de VITERBI



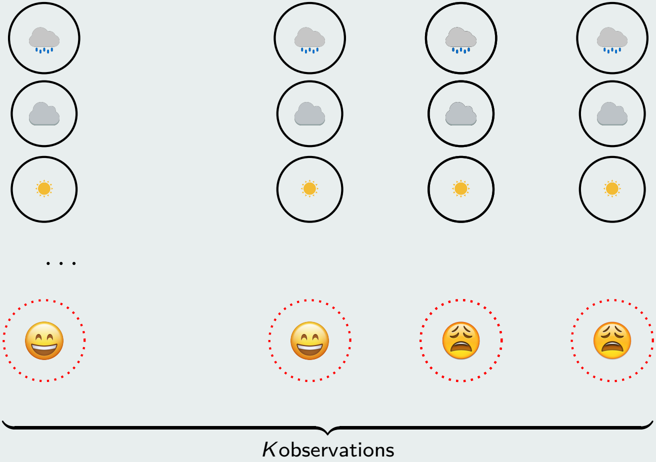
Algorithme de VITERBI



Algorithme de VITERBI



Complexité de l'algorithme de Viterbi $O(K \cdot N^2) << O(N^K)$



Introduction algèbre $(\max, +)$

Algèbre classique $(+, \times)$

- $\pi_1 \times p_{11} + \pi_2 \times p_{12}$

Introduction algèbre $(\max, +)$

Algèbre classique $(+, \times)$

- $\pi_1 \times p_{11} + \pi_2 \times p_{12}$

Algèbre $(\oplus = \max, \otimes = +)$

- $\pi_1 \otimes p_{11} \oplus \pi_2 \otimes p_{12} = \max(\pi_1 + p_{11}, \pi_2 + p_{12})$

Introduction algèbre (max, +)

Algèbre classique (+, ×)

- $\pi_1 \times p_{11} + \pi_2 \times p_{12}$

Algèbre ($\oplus = \max, \otimes = +$)

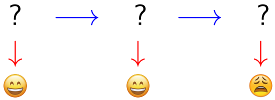
- $\pi_1 \otimes p_{11} \oplus \pi_2 \otimes p_{12} = \max(\pi_1 + p_{11}, \pi_2 + p_{12})$

Exemple en algèbre ($\oplus = \max, \otimes = +$)

$$2 \otimes 5 = 7$$

$$1 \oplus 1 = 1$$

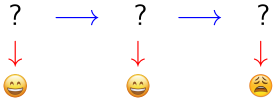
Algèbre (max, +)



Objectif en algèbre (Max, +)

$$\arg\max_{x_0, x_1, x_2} \prod_i Pr(y_i | x_i) Pr(x_{i+1} | x_i)$$

Algèbre (max, +)

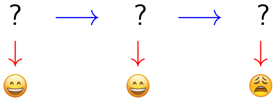


Objectif en algèbre (Max, +)

$$\arg\max_{x_0,x_1,x_2} \prod_i Pr(y_i|x_i)Pr(x_{i+1}|x_i)$$

$$\Leftrightarrow \arg\max_{x_0,x_1,x_2} \log\left(\prod_i Pr(y_i|x_i)Pr(x_{i+1}|x_i)\right)$$

Algèbre (max, +)

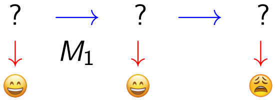


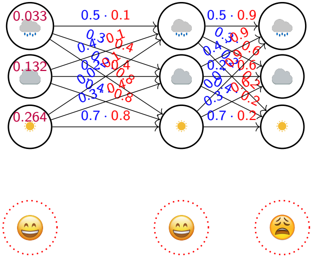
Objectif en algèbre (Max, +)

$$\arg\max_{x_0,x_1,x_2} \prod_i Pr(y_i|x_i)Pr(x_{i+1}|x_i)$$

$$\Leftrightarrow \arg\max_{x_0,x_1,x_2} \log\left(\prod_i Pr(y_i|x_i)Pr(x_{i+1}|x_i)\right)$$

$$\Leftrightarrow \arg\max_{x_0,x_1,x_2} \sum_i \log\left(Pr(x_{i+1}|x_i)\right) + \log(Pr(y_i|x_i))$$





Algèbre classique

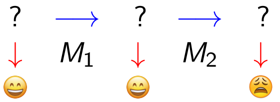
Matrice de probabilité sachant la mesure $Y_1 =$ 😄.

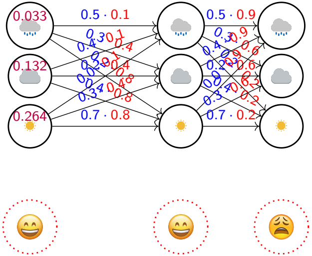
$$M_1 = \begin{pmatrix} 0.5 & 0.3 & 0.2 \\ 0.4 & 0.2 & 0.4 \\ 0.0 & 0.3 & 0.7 \end{pmatrix} \begin{pmatrix} 0.1 & 0 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0 & 0.8 \end{pmatrix}$$

$$M_1 = \begin{pmatrix} 0.5 \cdot 0.1 & 0.3 \cdot 0.4 & 0.2 \cdot 0.8 \\ 0.4 \cdot 0.1 & 0.2 \cdot 0.4 & 0.4 \cdot 0.8 \\ 0.0 \cdot 0.1 & 0.3 \cdot 0.4 & 0.7 \cdot 0.8 \end{pmatrix}$$

	🌧	☁	☀
🌧	0.5	0.3	0.2
☁	0.4	0.2	0.4
☀	0.0	0.3	0.7

	😩	😄
🌧	0.9	0.1
☁	0.6	0.4
☀	0.2	0.8





Algèbre classique

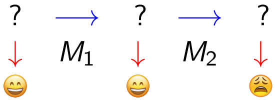
Matrice de probabilité sachant la mesure $Y_2 =$ 😄.

$$M_2 = \begin{pmatrix} 0.5 & 0.3 & 0.2 \\ 0.4 & 0.2 & 0.4 \\ 0.0 & 0.3 & 0.7 \end{pmatrix} \begin{pmatrix} 0.9 & 0 & 0 \\ 0 & 0.6 & 0 \\ 0 & 0 & 0.2 \end{pmatrix}$$

$$M_2 = \begin{pmatrix} 0.5 \cdot 0.9 & 0.3 \cdot 0.6 & 0.2 \cdot 0.2 \\ 0.4 \cdot 0.9 & 0.2 \cdot 0.6 & 0.4 \cdot 0.2 \\ 0.0 \cdot 0.9 & 0.3 \cdot 0.6 & 0.7 \cdot 0.2 \end{pmatrix}$$

	🌧	☁	☀
🌧	0.5	0.3	0.2
☁	0.4	0.2	0.4
☀	0.0	0.3	0.7

	😩	😄
🌧	0.9	0.1
☁	0.6	0.4
☀	0.2	0.8



Remarque

- En algèbre classique :

$$P^2 = \begin{pmatrix} p_{11}p_{11} + p_{12}p_{21} & p_{12}p_{11} + p_{12}p_{22} \\ p_{21}p_{1,} + p_{22}p_{21} & p_{22}p_{11} + p_{22}p_{22} \end{pmatrix}$$

- En algèbre $(\max, +)$:

$$\tilde{P}^2 = \begin{pmatrix} l_{11} \otimes l_{11} \oplus l_{12} \otimes l_{21} & l_{1,2} \otimes l_{11} \oplus l_{12} \otimes l_{22} \\ l_{21} \otimes l_{11} \oplus l_{22} \otimes l_{21} & l_{22} \otimes l_{11} \oplus l_{22} \otimes l_{22} \end{pmatrix}$$

$$\tilde{P}^2 = \begin{pmatrix} \max(l_{11} + l_{11}, l_{12} + l_{21}) & \max(l_{12} + l_{11}, l_{12} + l_{22}) \\ \max(l_{21} + l_{11}, l_{22} + l_{21}) & \max(l_{22} + l_{11}, l_{22} + l_{22}) \end{pmatrix}$$

$$\begin{array}{ccccc} ? & \xrightarrow{} & ? & \xrightarrow{} & ? \\ \downarrow & M_1 & \downarrow & M_2 & \downarrow \end{array}$$

😄 😄 😩

Remarque

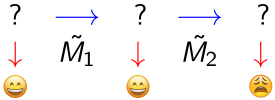
- En algèbre classique :

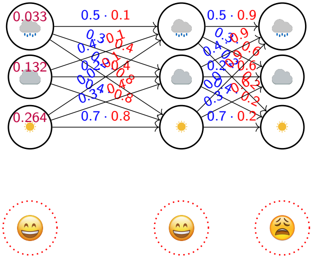
$$P^2 = \begin{pmatrix} p_{11}p_{11} + p_{12}p_{21} & p_{12}p_{11} + p_{12}p_{22} \\ p_{21}p_{11} + p_{22}p_{21} & p_{22}p_{11} + p_{22}p_{22} \end{pmatrix}$$

- En algèbre $(\max, +)$:

$$\tilde{P}^2 = \begin{pmatrix} l_{11} \otimes l_{11} \oplus l_{12} \otimes l_{21} & l_{1,2} \otimes l_{11} \oplus l_{12} \otimes l_{22} \\ l_{21} \otimes l_{11} \oplus l_{22} \otimes l_{21} & l_2 \otimes l_{11} \oplus l_{22} \otimes l_{22} \end{pmatrix}$$

$$\tilde{P}^2 = \begin{pmatrix} \max(\log p_{11} + \log p_{11}, \log p_{12} + \log p_{21}) & \max(\log p_{12} + \log p_{11}, \log \\ \max(\log p_{21} + \log p_{11}, \log p_{22} + \log p_{21}) & \max(\log p_{22} + \log p_{11}, \log \end{pmatrix}$$



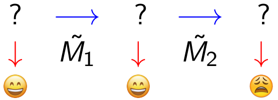


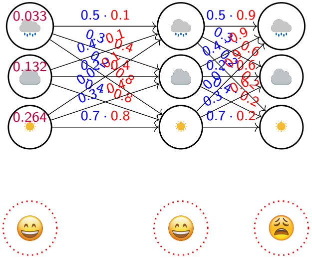
Matrice de Log-probabilité

$$\tilde{M}_1 = \begin{pmatrix} \log 0.05 & \log 0.012 & \log 0.016 \\ \log 0.04 & \log 0.08 & \log 0.032 \\ -\infty & \log 0.012 & \log 0.056 \end{pmatrix}$$

$$\tilde{M}_2 = \begin{pmatrix} \log 0.045 & \log 0.018 & \log 0.04 \\ \log 0.036 & \log 0.012 & \log 0.08 \\ -\infty & \log 0.018 & \log 0.014 \end{pmatrix}$$

Vision (Max, +) de l'algorithme de Viterbi





Calcul de la séquence la plus probable :

$$\tilde{\Pi}_2 = \tilde{\Pi} \otimes \tilde{M}_2 \otimes \tilde{M}_1$$

Application 1

Définition : Speech tagging

Étiquetage morpho-syntaxique (speech tagging) est le processus automatique qui consiste à associer aux mots des informations grammaticales.

Application 1

Définition : Speech tagging

Étiquetage morpho-syntaxique (speech tagging) est le processus automatique qui consiste à associer aux mots des informations grammaticales.

Exemple :

"La vie est belle"

A partir de données empiriques, on a :

vecteur initial

$$\begin{matrix} DT \\ NM \\ VB \\ \vdots \end{matrix} \begin{pmatrix} 0.6 \\ 0.2 \\ 0.1 \\ \vdots \end{pmatrix}$$

Matrices des transitions

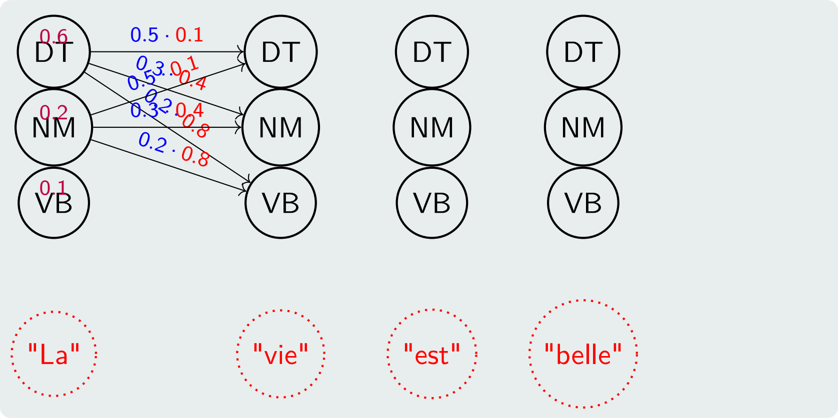
$$\begin{matrix} \\ DT \\ NM \\ VB \\ \dots \end{matrix} \begin{matrix} DT & NM & VB \\ \end{matrix}$$

	DT	NM	VB
DT	0	0.9	0.1
NM	0.4	0.5	0.5
VB	0.5	0.5	0
...		...	

Matrices des émissions

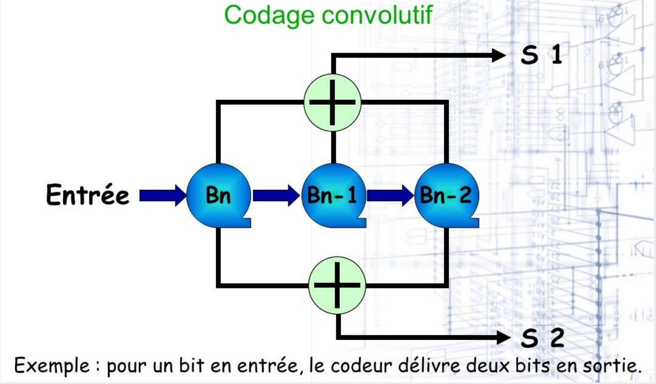
	La	Vie	est	belle
DT	0.2	0	0.1	0
NM	0	0.1	0.2	0.1
VB	0	0.2	0.3	0.3
...		...		

Application 1



Applications

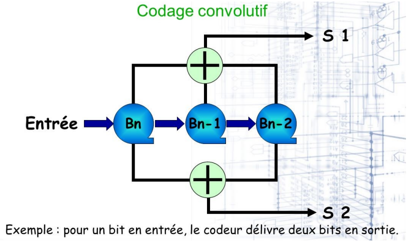
Codeur convolutif



$S_1 = B_n \oplus B_{n-1} \oplus B_{n-2}$

$S_2 = B_n \oplus B_{n-2}$

Application 2



Exemple : Codage de l'entrée 1 0 0 1

pour le 1er bit d'entrée 1 on a la sortie $S_1, S_2 = 1\ 1$

2eme bit 0 on a la sortie $S_1, S_2 = 1\ 0$

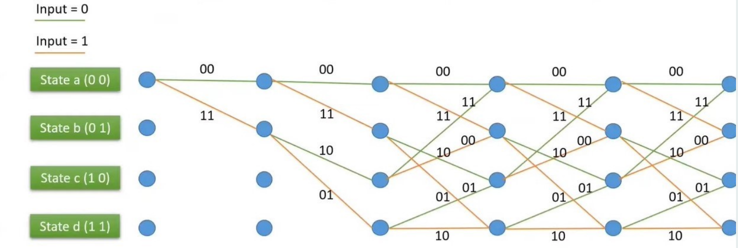
3eme bit 0 on a la sortie $S_1, S_2 = 1\ 0$

4eme bit 1 on a la sortie $S_1, S_2 = 1\ 1$

La sortie est donc : 11 10 11 11.

Application 2

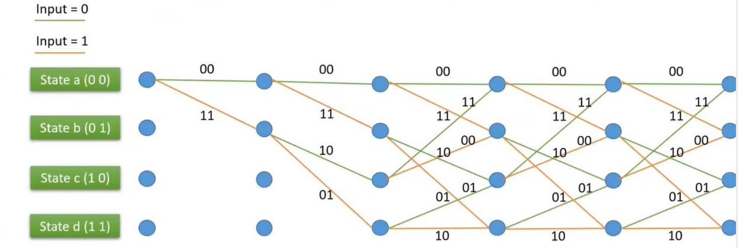
Vision treillis du codage convolutif



Codage de l'entrée 1 0 0 1
donne : 11 10 11 11.

Applications

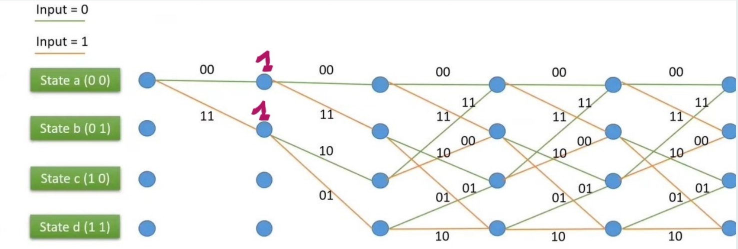
Décodage de la séquence 01 10 11 00 00



01 10 11 00 00

Applications

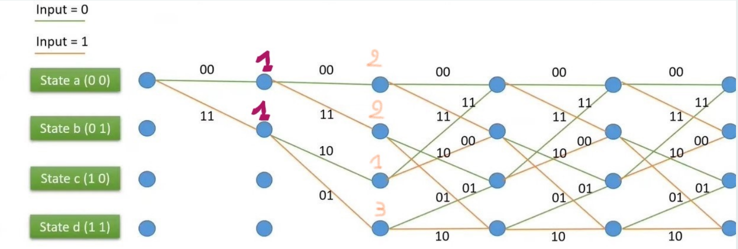
Exemple : Réception de la séquence 01 10 11 00 00



01 10 11 00 00

Applications

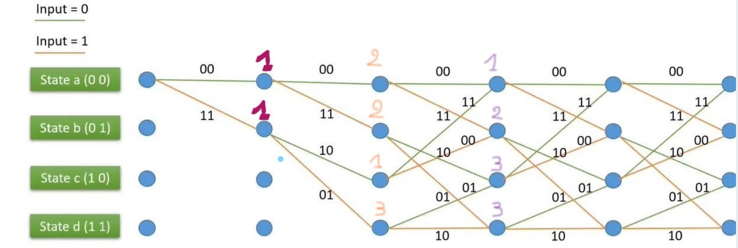
Exemple : Réception de la séquence 01 10 11 00 00



01 10 11 00 00

Applications

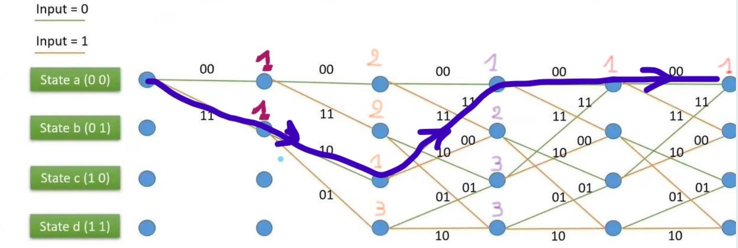
Exemple : Réception de la séquence 01 10 11 00 00



01 10 11 00 00

Application 2

Exemple : Réception de la séquence 01 10 11 00 00



~~01 10 11 00 00~~

11 10 11 00 00

Conclusion

Formulation tropical de l'algorithme de Viterbi pour les chaînes de Markov.

Merci pour votre attention !