# Prédiction dans un HMM

- $\alpha(i,t)$ peut être utilisé pour inférer la distribution de prédiction $\mathbf{P}(H_{t+k}|s_{1:t})$
- On utilise également un programme dynamique
  - on définit $\pi(i,k) = P(H_{t+k} = i|S_{1:t}=s_{1:t})$
  - on note la récursion

$$\begin{aligned}\pi(i,k+1) &= P(H_{t+k+1} = i|S_{1:t}=s_{1:t}) \\ &= \textstyle\sum_j P(H_{t+k+1} = i,\ H_{t+k} = j|S_{1:t}=s_{1:t}) \\ &= \textstyle\sum_j P(H_{t+k+1} = i|H_{t+k} = j)\ P(H_{t+k} = j|S_{1:t}=s_{1:t}) \\ &= \textstyle\sum_j P(H_{t+k+1} = i|H_{t+k} = j)\ \pi(j,k)\end{aligned}$$

  - on a les valeurs initiales

$$\pi(i,0) = P(H_t = i|s_{1:t}) = \alpha(i,t) / \textstyle\sum_j \alpha(j,t) \quad \forall i$$

- On pourrait également faire une prédiction de $S_{t+k}$
  - $P(S_{t+k} = s|S_{1:t}=s_{1:t}) = \sum_j P(S_{t+k}= s|H_{t+k} = j)\ P(H_{t+k} = j|S_{1:t}=s_{1:t})$

$$= \textstyle\sum_j P(S_{t+k}= s|H_{t+k} = j)\ \pi(j,k)$$

# Prédiction dans un HMM

- Exemple: décoder un message binaire avec canal bruité ($T$=4)
  - on souhaite calculer la distribution de prédiction $P(H_6=0 \mid S_1=0, S_2=0, S_3=0, S_4=1)$

**Modèle d'observation**

| | $H_t=0$ | $H_t=1$ |
|---|---|---|
| $P(S_t=0 \mid H_t)$ | 0.9 | 0.2 |
| $P(S_t=1 \mid H_t)$ | 0.1 | 0.8 |

**Modèle de transition**

| | $H_{t-1}=0$ | $H_{t-1}=1$ |
|---|---|---|
| $P(H_t=0 \mid H_{t-1})$ | 0.3 | 0.6 |
| $P(H_t=1 \mid H_{t-1})$ | 0.7 | 0.4 |

**Distribution initiale**

| | $H_1=0$ | $H_1=1$ |
|---|---|---|
| $P(H_1)$ | 0.5 | 0.5 |

**α(i,t)**

| i \ t | ... | 4 |
|---|---|---|
| 0 | ... | 0.00439 |
| 1 | ... | 0.05769 |

**π(i,k)**

| i \ k | 0 | 1 | 2 |
|---|---|---|---|
| 0 | | | |
| 1 | | | |

  - initialisation: $\pi(i,0) = \alpha(i,t) / \sum_j \alpha(j,t)$


Hugo Larochelle et Froduald Kabanza

# Prédiction dans un HMM

- Exemple: décoder un message binaire avec canal bruité ($T$=4)
  - on souhaite calculer la distribution de prédiction $P(H_6=0 \mid S_1=0, S_2=0, S_3=0, S_4=1)$

**Modèle d'observation**

| | $H_t=0$ | $H_t=1$ |
|---|---|---|
| $P(S_t=0 \mid H_t)$ | 0.9 | 0.2 |
| $P(S_t=1 \mid H_t)$ | 0.1 | 0.8 |

**Modèle de transition**

| | $H_{t-1}=0$ | $H_{t-1}=1$ |
|---|---|---|
| $P(H_t=0 \mid H_{t-1})$ | 0.3 | 0.6 |
| $P(H_t=1 \mid H_{t-1})$ | 0.7 | 0.4 |

**Distribution initiale**

| | $H_1=0$ | $H_1=1$ |
|---|---|---|
| $P(H_1)$ | 0.5 | 0.5 |

**α(i,t)**

| i \ t | ... | 4 |
|---|---|---|
| 0 | ... | 0.00439 |
| 1 | ... | 0.05769 |

**π(i,k)**

| i \ k | 0 | 1 | 2 |
|---|---|---|---|
| 0 | | | |
| 1 | | | |

  - initialisation: $\pi(0,0) = \alpha(0,4) \,/\, (\, \alpha(0,4) + \alpha(1,4)\,)$

# Prédiction dans un HMM

- Exemple: décoder un message binaire avec canal bruité ($T$=4)
    - on souhaite calculer la distribution de prédiction $P(H_6=0 \mid S_1=0, S_2=0, S_3=0, S_4=1)$

**Modèle d'observation**

| | $H_t=0$ | $H_t=1$ |
|---|---|---|
| $P(S_t=0 \mid H_t)$ | 0.9 | 0.2 |
| $P(S_t=1 \mid H_t)$ | 0.1 | 0.8 |

**Modèle de transition**

| | $H_{t-1}=0$ | $H_{t-1}=1$ |
|---|---|---|
| $P(H_t=0 \mid H_{t-1})$ | 0.3 | 0.6 |
| $P(H_t=1 \mid H_{t-1})$ | 0.7 | 0.4 |

**Distribution initiale**

| | $H_1=0$ | $H_1=1$ |
|---|---|---|
| $P(H_1)$ | 0.5 | 0.5 |

**α(i,t)**

| i \ t | ... | 4 |
|---|---|---|
| 0 | ... | 0.00439 |
| 1 | ... | 0.05769 |

**π(i,k)**

| i \ k | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 0.07071 | | |
| 1 | | | |

- initialisation: $\pi(0,0) = 0.00439 / (0.00439 + 0.05769) = 0.07071$



Hugo Larochelle et Froduald Kabanza

# Prédiction dans un HMM

- Exemple: décoder un message binaire avec canal bruité ($T$=4)
  - on souhaite calculer la distribution de prédiction $P(H_6=0 \mid S_1=0,S_2=0,S_3=0,S_4=1)$

**Modèle d'observation**

| | $H_t=0$ | $H_t=1$ |
|---|---|---|
| $P(S_t=0 \mid H_t)$ | 0.9 | 0.2 |
| $P(S_t=1 \mid H_t)$ | 0.1 | 0.8 |

**Modèle de transition**

| | $H_{t-1}=0$ | $H_{t-1}=1$ |
|---|---|---|
| $P(H_t=0 \mid H_{t-1})$ | 0.3 | 0.6 |
| $P(H_t=1 \mid H_{t-1})$ | 0.7 | 0.4 |

**Distribution initiale**

| | $H_1=0$ | $H_1=1$ |
|---|---|---|
| $P(H_1)$ | 0.5 | 0.5 |

**α(i,t)**

| i \ t | ... | 4 |
|---|---|---|
| 0 | ... | 0.00439 |
| 1 | ... | 0.05769 |

**π(i,k)**

| i \ k | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 0.07071 | | |
| 1 | 0.92929 | | |

  - initialisation: π(1,0) = 0.05769 / (0.00439 + 0.05769) = 0.92929

# Prédiction dans un HMM

- Exemple: décoder un message binaire avec canal bruité ($T$=4)
  - on souhaite calculer la distribution de prédiction $P(H_6=0 \mid S_1=0, S_2=0, S_3=0, S_4=1)$

**Modèle d'observation**

| | $H_t$=0 | $H_t$=1 |
|---|---|---|
| $P(S_t=0 \mid H_t)$ | 0.9 | 0.2 |
| $P(S_t=1 \mid H_t)$ | 0.1 | 0.8 |

**Modèle de transition**

| | $H_{t-1}$=0 | $H_{t-1}$=1 |
|---|---|---|
| $P(H_t=0 \mid H_{t-1})$ | 0.3 | 0.6 |
| $P(H_t=1 \mid H_{t-1})$ | 0.7 | 0.4 |

**Distribution initiale**

| | $H_1$=0 | $H_1$=1 |
|---|---|---|
| $P(H_1)$ | 0.5 | 0.5 |

**α(i,t)**

| i \ t | ... | 4 |
|---|---|---|
| 0 | ... | 0.00439 |
| 1 | ... | 0.05769 |

**π(i,k)**

| i \ k | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 0.07071 | | |
| 1 | 0.92929 | | |

  - récursion (k=0): $\pi(i,k+1) = \sum_j P(H_{t+k+1} = i \mid H_{t+k} = j)\ \pi(j,k)$

# Prédiction dans un HMM

- Exemple: décoder un message binaire avec canal bruité ($T$=4)
  - on souhaite calculer la distribution de prédiction $P(H_6=0 \mid S_1=0,S_2=0,S_3=0,S_4=1)$

**Modèle d'observation**

| | $H_t$=0 | $H_t$=1 |
|---|---|---|
| $P(S_t=0 \mid H_t)$ | 0.9 | 0.2 |
| $P(S_t=1 \mid H_t)$ | 0.1 | 0.8 |

**Modèle de transition**

| | $H_{t-1}$=0 | $H_{t-1}$=1 |
|---|---|---|
| $P(H_t=0 \mid H_{t-1})$ | 0.3 | 0.6 |
| $P(H_t=1 \mid H_{t-1})$ | 0.7 | 0.4 |

**Distribution initiale**

| | $H_1$=0 | $H_1$=1 |
|---|---|---|
| $P(H_1)$ | 0.5 | 0.5 |

**α(i,t)**

| i \ t | ... | 4 |
|---|---|---|
| 0 | ... | 0.00439 |
| 1 | ... | 0.05769 |

**π(i,k)**

| i \ k | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 0.07071 | | |
| 1 | 0.92929 | | |

  - récursion (k=0): $\pi(0, 1) = P(H_5 = 0 \mid H_4 = 0)\ \pi(0,0) + P(H_5 = 0 \mid H_4 = 1)\ \pi(1,0)$

# Prédiction dans un HMM

- Exemple: décoder un message binaire avec canal bruité ($T$=4)
  - on souhaite calculer la distribution de prédiction $P(H_6=0 \mid S_1=0, S_2=0, S_3=0, S_4=1)$

**Modèle d'observation**

| | $H_t=0$ | $H_t=1$ |
|---|---|---|
| $P(S_t=0 \mid H_t)$ | 0.9 | 0.2 |
| $P(S_t=1 \mid H_t)$ | 0.1 | 0.8 |

**Modèle de transition**

| | $H_{t-1}=0$ | $H_{t-1}=1$ |
|---|---|---|
| $P(H_t=0 \mid H_{t-1})$ | 0.3 | 0.6 |
| $P(H_t=1 \mid H_{t-1})$ | 0.7 | 0.4 |

**Distribution initiale**

| | $H_1=0$ | $H_1=1$ |
|---|---|---|
| $P(H_1)$ | 0.5 | 0.5 |

**α(i,t)**

| i \ t | ... | 4 |
|---|---|---|
| 0 | ... | 0.00439 |
| 1 | ... | 0.05769 |

**π(i,k)**

| i \ k | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 0.07071 | 0.57879 | |
| 1 | 0.92929 | | |

  - récursion (k=0): $\pi(0, 1) = 0.3 \times 0.07071 + 0.6 \times 0.92929 = 0.57879$



Hugo Larochelle et Froduald Kabanza

# Prédiction dans un HMM

- Exemple: décoder un message binaire avec canal bruité ($T$=4)
  - on souhaite calculer la distribution de prédiction $P(H_6=0 \mid S_1=0,S_2=0,S_3=0,S_4=1)$

**Modèle d'observation**

| | $H_t=0$ | $H_t=1$ |
|---|---|---|
| $P(S_t=0 \mid H_t)$ | 0.9 | 0.2 |
| $P(S_t=1 \mid H_t)$ | 0.1 | 0.8 |

**Modèle de transition**

| | $H_{t-1}=0$ | $H_{t-1}=1$ |
|---|---|---|
| $P(H_t=0 \mid H_{t-1})$ | 0.3 | 0.6 |
| $P(H_t=1 \mid H_{t-1})$ | 0.7 | 0.4 |

**Distribution initiale**

| | $H_1=0$ | $H_1=1$ |
|---|---|---|
| $P(H_1)$ | 0.5 | 0.5 |

**α(i,t)**

| i \ t | ... | 4 |
|---|---|---|
| 0 | ... | 0.00439 |
| 1 | ... | 0.05769 |

**π(i,k)**

| i \ k | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 0.07071 | 0.57879 | |
| 1 | 0.92929 | | |

  - récursion (k=0): $\pi(1, 1) = P(H_5 = 1 | H_4 = 0)\ \pi(0,0) + P(H_5 = 1 | H_4 = 1)\ \pi(1,0)$

# Prédiction dans un HMM

- Exemple: décoder un message binaire avec canal bruité ($T$=4)
  - on souhaite calculer la distribution de prédiction $P(H_6=0 \mid S_1=0, S_2=0, S_3=0, S_4=1)$

**Modèle d'observation**

| | $H_t=0$ | $H_t=1$ |
|---|---|---|
| $P(S_t=0 \mid H_t)$ | 0.9 | 0.2 |
| $P(S_t=1 \mid H_t)$ | 0.1 | 0.8 |

**Modèle de transition**

| | $H_{t-1}=0$ | $H_{t-1}=1$ |
|---|---|---|
| $P(H_t=0 \mid H_{t-1})$ | 0.3 | 0.6 |
| $P(H_t=1 \mid H_{t-1})$ | 0.7 | 0.4 |

**Distribution initiale**

| | $H_1=0$ | $H_1=1$ |
|---|---|---|
| $P(H_1)$ | 0.5 | 0.5 |

**α(i,t)**

| i \ t | ... | 4 |
|---|---|---|
| 0 | ... | 0.00439 |
| 1 | ... | 0.05769 |

**π(i,k)**

| i \ k | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 0.07071 | 0.57879 | |
| 1 | 0.92929 | 0.42121 | |

  - récursion (k=0): $\pi(0, 1) = 0.7 \times 0.07071 + 0.4 \times 0.92929 = 0.42121$

# Prédiction dans un HMM

- Exemple: décoder un message binaire avec canal bruité ($T$=4)
  - on souhaite calculer la distribution de prédiction $P(H_6=0 \mid S_1=0, S_2=0, S_3=0, S_4=1)$

**Modèle d'observation**

| | $H_t$=0 | $H_t$=1 |
|---|---|---|
| $P(S_t=0 \mid H_t)$ | 0.9 | 0.2 |
| $P(S_t=1 \mid H_t)$ | 0.1 | 0.8 |

**Modèle de transition**

| | $H_{t-1}$=0 | $H_{t-1}$=1 |
|---|---|---|
| $P(H_t=0 \mid H_{t-1})$ | 0.3 | 0.6 |
| $P(H_t=1 \mid H_{t-1})$ | 0.7 | 0.4 |

**Distribution initiale**

| | $H_1$=0 | $H_1$=1 |
|---|---|---|
| $P(H_1)$ | 0.5 | 0.5 |

**α(i,t)**

| i \ t | ... | 4 |
|---|---|---|
| 0 | ... | 0.00439 |
| 1 | ... | 0.05769 |

**π(i,k)**

| i \ k | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 0.07071 | 0.57879 | 0.42637 |
| 1 | 0.92929 | 0.42121 | 0.57363 |

  - on continue d'appliquer la récursion jusqu'à la fin (k=2)...

# Prédiction dans un HMM

- Exemple: décoder un message binaire avec canal bruité ($T$=4)
  - on souhaite calculer la distribution de prédiction $P(H_6=0 \mid S_1=0,S_2=0,S_3=0,S_4=1)$

**Modèle d'observation**

| | $H_t=0$ | $H_t=1$ |
|---|---|---|
| $P(S_t=0 \mid H_t)$ | 0.9 | 0.2 |
| $P(S_t=1 \mid H_t)$ | 0.1 | 0.8 |

**Modèle de transition**

| | $H_{t-1}=0$ | $H_{t-1}=1$ |
|---|---|---|
| $P(H_t=0 \mid H_{t-1})$ | 0.3 | 0.6 |
| $P(H_t=1 \mid H_{t-1})$ | 0.7 | 0.4 |

**Distribution initiale**

| | $H_1=0$ | $H_1=1$ |
|---|---|---|
| $P(H_1)$ | 0.5 | 0.5 |

**α(i,t)**

| i \ t | ... | 4 |
|---|---|---|
| 0 | ... | 0.00439 |
| 1 | ... | 0.05769 |

**π(i,k)**

| i \ k | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 0.07071 | 0.57879 | 0.42637 |
| 1 | 0.92929 | 0.42121 | 0.57363 |

  - $P(H_6=0 \mid S_1=0,S_2=0,S_3=0,S_4=1) = \pi(0,2) = 0.42637$