

Explication la plus plausible avec un HMM

- On peut également éviter une énumération exponentielle

- ◆ exemple avec $T=3$:

$$\begin{aligned} & \max_{h^*_{1:3}} P(h^*_1) P(s_1|h^*_1) P(h^*_2|h^*_1) P(s_2|h^*_2) P(h^*_3|h^*_2) P(s_3|h^*_3) \\ &= \max_{h^*_3} P(s_3|h^*_3) \max_{h^*_2} P(s_2|h^*_2) P(h^*_3|h^*_2) \max_{h^*_1} P(h^*_2|h^*_1) P(h^*_1) P(s_1|h^*_1) \end{aligned}$$

- Solution: **programmation dynamique**, avec un **max** au lieu de la somme

- ◆ on définit $\alpha^*(i,t) = P(S_{1:t}=s_{1:t}, H_{1:t-1} = h^*_{1:t-1}, H_t = i)$

- ◆ on note la récursion

$$\begin{aligned} \alpha^*(i,t+1) &= \max_j P(S_{1:t+1}=s_{1:t+1}, H_{1:t-1} = h^*_{1:t-1}, H_t = j, H_{t+1} = i) \\ &= \max_j P(S_{t+1}=s_{t+1} \mid H_{t+1} = i) P(H_{t+1} = i \mid H_t = j) P(S_{1:t}=s_{1:t}, H_{1:t-1} = h^*_{1:t-1}, H_t = j) \\ &= P(S_{t+1}=s_{t+1} \mid H_{t+1} = i) \max_j P(H_{t+1} = i \mid H_t = j) \alpha^*(j,t) \end{aligned}$$

- ◆ on a les valeurs initiales: $\alpha^*(i,1) = P(S_1=s_1 \mid H_1 = i) P(H_1 = i) \quad \forall i$

- On a alors que $P(S_{1:T}=s_{1:T}, H_{1:T} = h^*_{1:T}) = \max_j \alpha^*(j,T)$

- On retrouve $h^*_{1:T}$ à partir de tous les argmax_j

Explication la plus plausible avec un HMM

- Exemple: décoder un message binaire avec canal bruité ($T=4$)
 - message observé: $S_1=0, S_2=0, S_3=0, S_4=1$

Modèle d'observation

	$H_t=0$	$H_t=1$
$P(S_t=0 H_t)$	0.9	0.2
$P(S_t=1 H_t)$	0.1	0.8

Modèle de transition

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

Distribution initiale

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

$\alpha^*(i,t)$	$i \backslash t$	1	2	3	4
	0				
	1				

- initialisation: $\alpha^*(i,1) = P(S_1=s_1, H_1 = i) = P(S_1=s_1 | H_1 = i) P(H_1 = i)$

Explication la plus plausible avec un HMM

- Exemple: décoder un message binaire avec canal bruité ($T=4$)
 - message observé: $S_1=0, S_2=0, S_3=0, S_4=1$

Modèle d'observation

	$H_t=0$	$H_t=1$
$P(S_t=0 H_t)$	0.9	0.2
$P(S_t=1 H_t)$	0.1	0.8

Modèle de transition

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

Distribution initiale

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

$\alpha^*(i,t)$	$\begin{matrix} i & t \end{matrix}$		1	2	3	4
	0	1	0.45			
	1					

- initialisation: $\alpha^*(0,1) = P(S_1=0 | H_1=0) P(H_1=0) = 0.9 \times 0.5 = 0.45$

Explication la plus plausible avec un HMM

- Exemple: décoder un message binaire avec canal bruité ($T=4$)
 - message observé: $S_1=0, S_2=0, S_3=0, S_4=1$

Modèle d'observation


	$H_t=0$	$H_t=1$
$P(S_t=0 H_t)$	0.9	0.2
$P(S_t=1 H_t)$	0.1	0.8

Modèle de transition

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

Distribution initiale

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

$\alpha^*(i,t)$			1	2	3	4
	i	t				
0			0.45			
1			0.1			

- initialisation: $\alpha^*(1,1) = P(S_1=0 | H_1=1) P(H_1=1) = 0.2 \times 0.5 = 0.1$

Explication la plus plausible avec un HMM

- Exemple: décoder un message binaire avec canal bruité ($T=4$)
 - ◆ message observé: $S_1=0, S_2=0, S_3=0, S_4=1$

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

$\alpha^*(i,t)$	$i \backslash t$	1	2	3	4
0		0.45			
1		0.1			

- ◆ réursion (t=1): $\alpha^*(i, t+1) = P(S_{t+1} = s_{t+1} \mid H_{t+1} = i) \max_j P(H_{t+1} = i \mid H_t = j) \alpha^*(j, t)$

Explication la plus plausible avec un HMM

- Exemple: décoder un message binaire avec canal bruité ($T=4$)
 - message observé: $S_1=0, S_2=0, S_3=0, S_4=1$

Modèle d'observation

	$H_t=0$	$H_t=1$
$P(S_t=0 H_t)$	0.9	0.2
$P(S_t=1 H_t)$	0.1	0.8

Modèle de transition

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

Distribution initiale

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

$\alpha^*(i,t)$	i \ t	1	2	3	4
0		0.45			
1		0.1			

- réursion: $\alpha^*(0,2) = P(S_2=0 | H_2=0) \max\{P(H_2=0 | H_1=0) \alpha^*(0,1), P(H_2=0 | H_1=1) \alpha^*(1,1)\}$

Explication la plus plausible avec un HMM

- Exemple: décoder un message binaire avec canal bruité ($T=4$)
 - message observé: $S_1=0, S_2=0, S_3=0, S_4=1$

Modèle d'observation

	$H_t=0$	$H_t=1$
$P(S_t=0 H_t)$	0.9	0.2
$P(S_t=1 H_t)$	0.1	0.8

Modèle de transition

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

Distribution initiale

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

$\alpha^*(i,t)$	i \ t	1	2	3	4
0	0	0.45	0.1215		
1	1	0.1			

- réursion: $\alpha^*(0,2) = 0.9 \max\{ \underline{0.3 \times 0.45}, 0.6 \times 0.1 \} = 0.1215$

Explication la plus plausible avec un HMM

- Exemple: décoder un message binaire avec canal bruité ($T=4$)
 - message observé: $S_1=0, S_2=0, S_3=0, S_4=1$

Modèle d'observation

	$H_t=0$	$H_t=1$
$P(S_t=0 H_t)$	0.9	0.2
$P(S_t=1 H_t)$	0.1	0.8

Modèle de transition

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

Distribution initiale

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

$\alpha^*(i,t)$	$i \backslash t$	1	2	3	4
0		0.45	0.1215		
1		0.1			

- récursion: $\alpha^*(1,2) = P(S_2=0 | H_2=1) \max\{P(H_2=1 | H_1=0) \alpha^*(0,1), P(H_2=1 | H_1=1) \alpha^*(1,1)\}$

Explication la plus plausible avec un HMM

- Exemple: décoder un message binaire avec canal bruité ($T=4$)
 - message observé: $S_1=0, S_2=0, S_3=0, S_4=1$

Modèle d'observation

	$H_t=0$	$H_t=1$
$P(S_t=0 H_t)$	0.9	0.2
$P(S_t=1 H_t)$	0.1	0.8

Modèle de transition

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

Distribution initiale

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

$\alpha^*(i,t)$	$i \backslash t$	1	2	3	4
	0	0.45	0.1215		
	1	0.1	0.063		

- réursion: $\alpha^*(1,2) = 0.2 \max\{ \underline{0.7 \times 0.45}, 0.4 \times 0.1 \} = 0.063$

Explication la plus plausible avec un HMM

- Exemple: décoder un message binaire avec canal bruité ($T=4$)
 - message observé: $S_1=0, S_2=0, S_3=0, S_4=1$

Modèle d'observation

	$H_t=0$	$H_t=1$
$P(S_t=0 H_t)$	0.9	0.2
$P(S_t=1 H_t)$	0.1	0.8

Modèle de transition

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

Distribution initiale

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

$\alpha^*(i,t)$	$i \backslash t$	1	2	3	4
0		0.45	0.1215		
1		0.1	0.063		

- récursion ($t=2$): $\alpha^*(i,t+1) = P(S_{t+1}=s_{t+1} | H_{t+1}=i) \max_j P(H_{t+1}=i | H_t=j) \alpha^*(j,t)$

Explication la plus plausible avec un HMM

- Exemple: décoder un message binaire avec canal bruité ($T=4$)
 - message observé: $S_1=0, S_2=0, S_3=0, S_4=1$

Modèle d'observation

	$H_t=0$	$H_t=1$
$P(S_t=0 H_t)$	0.9	0.2
$P(S_t=1 H_t)$	0.1	0.8

Modèle de transition

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

Distribution initiale

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

$\alpha^*(i,t)$	t	1	2	3	4
i					
0		0.45	0.1215	0.03402	
1		0.1	0.063		

- réursion: $\alpha^*(0,3) = 0.9 \max\{0.3 \times 0.1215, \underline{0.6 \times 0.063}\} = 0.03402$

Explication la plus plausible avec un HMM

- Exemple: décoder un message binaire avec canal bruité ($T=4$)
 - message observé: $S_1=0, S_2=0, S_3=0, S_4=1$

Modèle d'observation

	$H_t=0$	$H_t=1$
$P(S_t=0 H_t)$	0.9	0.2
$P(S_t=1 H_t)$	0.1	0.8

Modèle de transition

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

Distribution initiale

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

$\alpha^*(i,t)$	$i \backslash t$	1	2	3	4
0		0.45	0.1215	0.03402	0.001021
1		0.1	0.063	0.01701	0.019051

- on continue d'appliquer la récursion jusqu'à la fin ($t=4$)...

Explication la plus plausible avec un HMM

- Exemple: décoder un message binaire avec canal bruité ($T=4$)
 - message observé: $S_1=0, S_2=0, S_3=0, S_4=1$

Modèle d'observation

	$H_t=0$	$H_t=1$
$P(S_t=0 H_t)$	0.9	0.2
$P(S_t=1 H_t)$	0.1	0.8

Modèle de transition

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

Distribution initiale

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

$\alpha^*(i,t)$	$i \backslash t$				
		1	2	3	4
0		0.45	0.1215	0.03402	0.001021
1		0.1	0.063	0.01701	0.019051

- on trouve le maximum à la dernière colonne...

Explication la plus plausible avec un HMM

- Exemple: décoder un message binaire avec canal bruité ($T=4$)
 - message observé: $S_1=0, S_2=0, S_3=0, S_4=1$

Modèle d'observation

	$H_t=0$	$H_t=1$
$P(S_t=0 H_t)$	0.9	0.2
$P(S_t=1 H_t)$	0.1	0.8

Modèle de transition

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

Distribution initiale

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

$\alpha^*(i,t)$	$i \backslash t$				
		1	2	3	4
0		0.45	0.1215	0.03402	0.001021
1		0.1	0.063	0.01701	0.019051

- ... puis on **retrouve le chemin** qui a mené là

Explication la plus plausible avec un HMM

- Exemple: décoder un message binaire avec canal bruité ($T=4$)
 - message observé: $S_1=0, S_2=0, S_3=0, S_4=1$

Modèle d'observation

	$H_t=0$	$H_t=1$
$P(S_t=0 H_t)$	0.9	0.2
$P(S_t=1 H_t)$	0.1	0.8

Modèle de transition

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

Distribution initiale

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

$\alpha^*(i,t)$		1	2	3	4
i	t	0.45	0.1215	0.03402	0.001021
1	t	0.1	0.063	0.01701	0.019051

$H_1=0$ $H_2=1$ $H_3=0$ $H_4=1$

- ce chemin nous donne la séquence des H_t la plus probable