Probabilité d'une séquence observée

- Comme calculer la probabilité marginale d'une séquence d'observations :

$$P(S_{1:T}) = \sum_{h_{1:T}} P(H_{1:T} = h_{1:T})\, P(S_{1:T} \mid H_{1:T} = h_{1:T})$$

  - la même séquence observée peut être produite par plusieurs séquences cachées différentes
  - en fait, il y a un nombre exponentiel de séquences cachées possibles
  - un calcul naïf est donc très inefficace

Hugo Larochelle et Froduald Kabanza

Programmation dynamique pour HMM

- Une façon plus efficace de calculer la probabilité d'une séquence observée $s_{1:T}$

- Idée: utiliser la **programmation dynamique**
  - on définit $\alpha(i,t) = P(S_{1:t}=s_{1:t}, H_t = i)$
  - on note la récursion
    $\alpha(i,t+1) = P(S_{1:t+1}=s_{1:t+1}, \mathrm{H}_{t+1} = i)$
    $= \sum_j P(S_{1:t+1}=s_{1:t+1}, H_{\mathrm{t}} = j, H_{t+1} = i) = \sum_j P(S_{1:t}=s_{1:t}, H_t = j, H_{t+1} = i, S_{t+1}= \mathrm{s}_{t+1})$
    $= P(S_{t+1}= \mathrm{s}_{t+1} | H_{t+1} = i) \sum_j P(H_{t+1} = i | H_t = j) P(S_{1:t}=s_{1:t}, H_t = j)$
    $= P(S_{t+1}= \mathrm{s}_{t+1} | H_{t+1} = i) \sum_j P(H_{t+1} = i | H_t = j) \alpha(j,t)$
  - on a les valeurs initiales
    $\alpha(i,1) = P(S_1=s_1, H_1 = i) = P(S_1=s_1 | H_1 = i) P(H_1 = i) \quad \forall i$

- Une fois le tableau α calculé, on obtient facilement:
  $P(S_{1:T}=s_{1:T}) = \sum_j P(S_{1:T}=s_{1:T}, H_T = j) = \sum_j \alpha(j,T)$

Programmation dynamique pour 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

α(i,t)

i \ t	1	2	3	4
0				
1				

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

Programmation dynamique pour 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

α(i,t)

i \ t	1	2	3	4
0	0.45			
1				

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

Programmation dynamique pour 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

α(i,t)

i \ t	1	2	3	4
0	0.45			
1	0.1			

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

Programmation dynamique pour 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

α(i,t)

i \ t	1	2	3	4
0	0.45			
1	0.1			

  - récursion (t=1): $\alpha(i,t+1) = P(S_{t+1} = s_{t+1} \mid H_{t+1} = i) \sum_j P(H_{t+1} = i \mid H_t = j)\, \alpha(j,t)$

Programmation dynamique pour 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

α(i,t)

i \ t	1	2	3	4
0	0.45			
1	0.1			

  - récursion: $\alpha(0,2) = P(S_2 = 0 \mid H_2 = 0)\ (\ P(H_2 = 0 \mid H_1 = 0)\ \alpha(0,1) + P(H_2 = 0 \mid H_1 = 1)\ \alpha(1,1))$

Hugo Larochelle et Froduald Kabanza

Programmation dynamique pour 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

α(i,t)

i \ t	1	2	3	4
0	0.45	0.1755		
1	0.1			

  - récursion: $\alpha(0,2)$ = 0.9 (0.3 x 0.45 + 0.6 x 0.1) = 0.1755

Programmation dynamique pour 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

α(i,t)

i \ t	1	2	3	4
0	0.45	0.1755		
1	0.1			

  - récursion: $\alpha(1,2) = P(S_2 = 0 \mid H_2 = 1)\ (\ P(H_2 = 1 \mid H_1 = 0)\ \alpha(0,1) + P(H_2 = 1 \mid H_1 = 1)\ \alpha(1,1))$

Programmation dynamique pour 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

α(i,t)

i \ t	1	2	3	4
0	0.45	0.1755		
1	0.1	0.071		

  - récursion: α(1,2) = 0.2 (0.7 x 0.45 + 0.4 x 0.1) = 0.071

Programmation dynamique pour 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

α(i,t)

i \ t	1	2	3	4
0	0.45	0.1755		
1	0.1	0.071		

  - récursion (t=2): $\alpha(i,t+1) = P(S_{t+1}= s_{t+1} \mid H_{t+1} = i) \sum_j P(H_{t+1} = i \mid H_t = j)\, \alpha(j,t)$

Programmation dynamique pour 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

α(i,t)

i \ t	1	2	3	4
0	0.45	0.1755	0.085725	
1	0.1	0.071		

  - récursion: α(0,3) = 0.9 (0.3 x 0.1755 + 0.6 x 0.071) = 0.085725

Programmation dynamique pour 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

α(i,t)

i \ t	1	2	3	4
0	0.45	0.1755	0.085725	0.004387
1	0.1	0.071	0.03025	0.057686

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

Hugo Larochelle et Froduald Kabanza

Filtrage dans 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

α(i,t)

i \ t	1	2	3	4
0	0.45	0.1755	0.085725	0.004387
1	0.1	0.071	0.03025	0.057686

  - on peut calculer les probabilités de filtrage

$$P(H_4 = 0 \mid S_1=0, S_2=0, S_3=0, S_4=1) = \frac{P(H_4 = 0, S_1=0, S_2=1, S_3=0, S_4=0)}{\sum_i P(H_4 = \mathrm{i}, S_1=0, S_2=1, S_3=0, S_4=0)}$$

$$= \alpha(0,4) \,/\, (\, \alpha(0,4) + \alpha(1,4)\,)$$

$$= 0.004387 \,/\, (0.004387 + 0.057686)$$

$$\approx 0.0707$$

$$P(H_4 = 1 \mid S_1=0, S_2=0, S_3=0, S_4=1) = 0.057686/\, (0.004387 + 0.057686)$$

$$\approx 0.9293$$

Hugo Larochelle et Froduald Kabanza