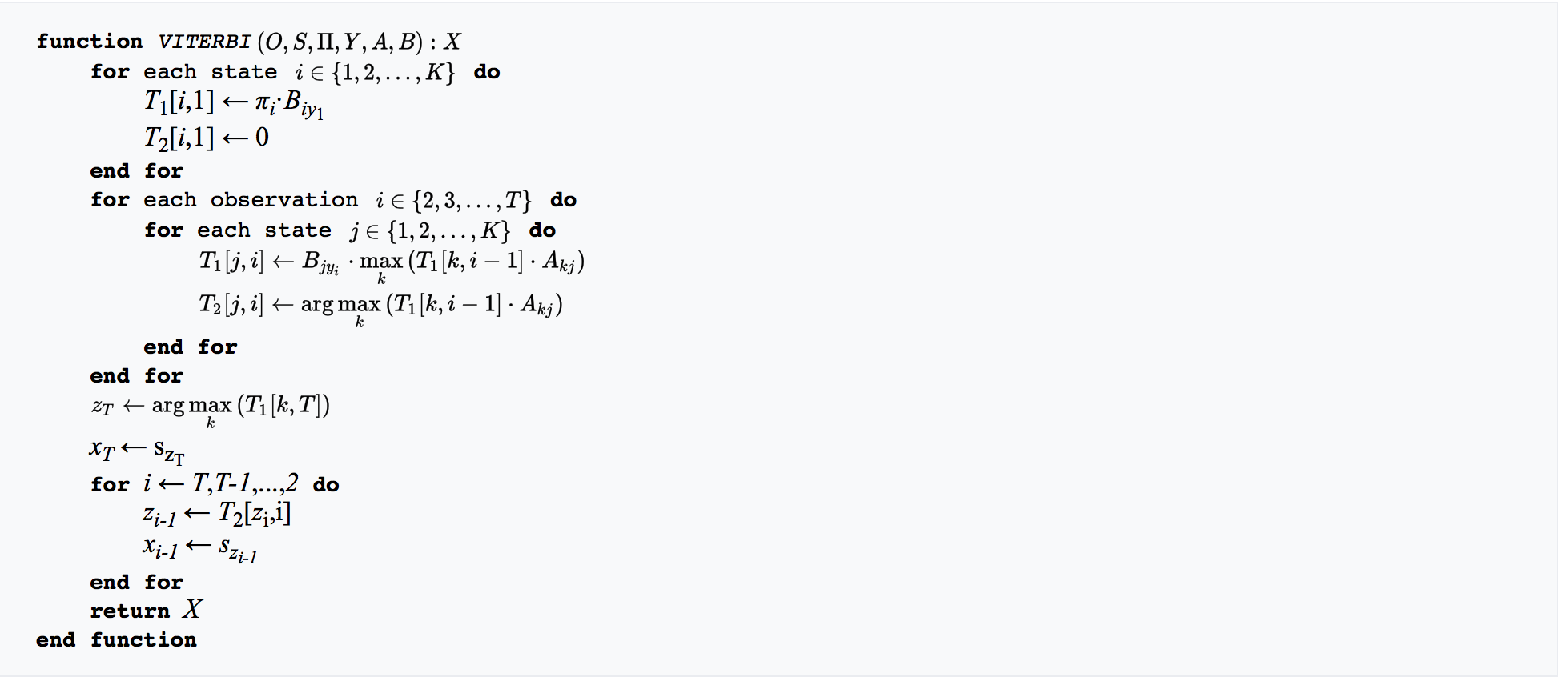
**TD HMM**

**Algorithme de Viterbi (wikipedia EN)**

**Entrées**

* Un ensemble de K états S={s1, s2, …,si, …, sK}
* Un ensemble de taille N des observations O = {o1, o2, … , oj, …, oT}.
* Une matrice de transition A de taille K x K tel que ai,j est la probabilité de transité de l’état i à létat j.
* La matrice de probabilité des observations B de taille K x N. bi,j est la probabilité de l’observation oj dans l’état si.



**Archéologie :**

* Regarder l’article suivant :

<http://documents.irevues.inist.fr/bitstream/handle/2042/10273/AR88.pdf?sequence=1&isAllowed=y>

**Exercice : Viterbi à la main**

* Quelle est la matrice de transition A ?

|  |  |  |  |
| --- | --- | --- | --- |
|  | H | L |  |
| H | 0,5 | 0,5 | = 1 |
| L | 0,6 | 0,4 | = 1 |

* Quelles sont les probabilités initiales π ?

|  |  |
| --- | --- |
| H | 0,5 |
| L | 0,5 |
| Somme = 1 |  |

* Quelles sont les probabilités B ?

|  |  |  |
| --- | --- | --- |
|  | H | L |
| A | 0,2 | 0,3 |
| C | 0,3 | 0,2 |
| G | 0,3 | 0,2 |
| T | 0,2 | 0,3 |
| Somme = 1 | = 1 | = 1 |

* Donnez le vocabulaire des observations qui est connu du HMM ?

Vocabulaire = {A, C, G, T}





* Quelle est la probabilité de la séquence d’observation « GGCAC » (ADN) avec la séquence d’états « LLHHL » ? Indiquez le chemin dans le graphe, calculez la probabilité de ce chemin.

BH(A)=0,2

BH(C)=0,3



0,5

0,5

0,6

0,4

BL(G)=0,2

0,5



BL(C)=0,2

PIL \* BL(G)\*ALL

0,5 \* 0,2 \* 0,4 \* 0,2 \* 0,6 \* 0,3 \* 0,5 \* 0,2 \* 0,5 \* 0,2

* Il existe plusieurs séquences d’états pouvant générer la séquence d’observation « GGCAC ». Appliquer Viterbi pour trouver le chemin de probabilité maximale.



|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| T1 | i=1 O1= « G » | i=2  O1= « G » | i=3, O3= « C » | i=4, O4= « A » | i=5, O5= « C » |
| H | 0,5x0,3=0.15 | H🡪H 0.15x0,5x0,3= **0,0225**  L🡪H  0,3x0,6x0,1= 0,018 | H🡪H  0,0225x0,5x0,3=  **0,0034**  L🡪H  0,015x0,6x0,3=  0,0027 | H🡪H  0,0034x0,5x0,2=  0,00034  L🡪H  0,00225x0,6x0,2=  0,00027 | H🡪H  0,00034x0,5x0,3= 0,000051  L🡪H  0,00051x0,6x0,3= **0,0000918** |
| L | 0,5x0,2 = 0.1 | H🡪L  0,15x0,5x0,2=0,015  L🡪L  0,2x0,4x0,1= 0,008 | H🡪L  0,0225x0,5x0,2=  0,00225  L🡪L  0,015x0,4x0,2=  0,0012 | H🡪L  0,0034x0,5x0,3=  **0,00051**  L🡪L  0,00225x0,4x0,3=  0,00027 | H🡪L  0,00034x0,5x0,2= 0,000034  L🡪L  0,00051x0,4x0,2= 0,0000408 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| T2 | i=1 O1= « G » | i=2  O1= « G » | i=3, O3= « C » | i=4, O4= « A » | i=5, O5= « C » |
| H | - | H | H | H | L |
| L | - | H | H | H | L |

O = {G, G, C, A, C}

Chemin = {H, H, H, L, H}

* Que se passe-t-il si je retire l’arc de L vers H ?

Quand on est dans H et qu’on va dans L on ne peut plus revenir donc ça finira forcément sur L à la fin.