Modèle de Markov caché

- Dans une **modèle de Markov caché** (***hidden Markov model*** ou **HMM**):
  - il y a des **variables cachées** H_t et des **variables d'observation** S_t, toutes les deux discrètes
  - la chaîne de Markov est sur les variables cachées H_t
  - le symbole observé (émis) $S_t=s_t$ dépend uniquement de la variable cachée actuelle H_t

Illustration

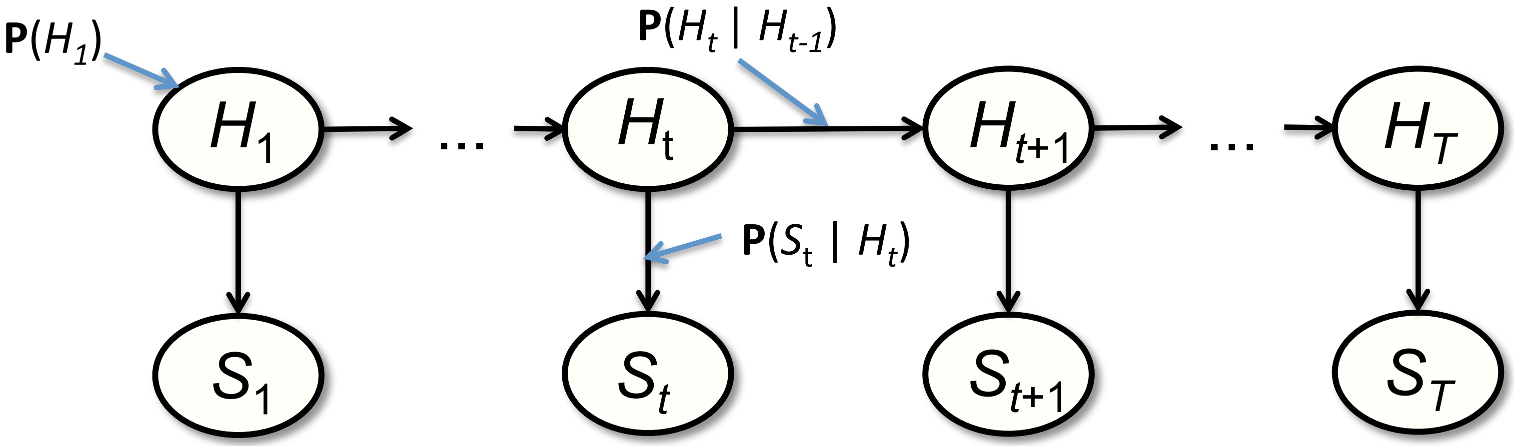

Illustration dans le cas d'une **chaîne finie**

Probabilité de générer une séquence cachée et une séquence visible

distribution initiale
de probabilités

modèle de
transition

$$P(S_{1:T}, H_{1:T}) = P(H_1)\, P(S_1|H_1) \prod_{t=2}^{T} P(H_t \mid H_{t-1})\, P(S_t \mid H_t)$$

séquence de nœuds cachés et de symboles de sortie

probabilité d'observer le symbole S_t à partir de H_t

Simuler d'un HMM

- Il est facile de générer des observations d'un HMM
  - échantillonner une valeur initiale $H_1 = h_1$ de $\mathbf{P}(H_1)$
  - pour t = 2 jusqu'à T, répéter les deux échantillonnages suivants:
    - » utiliser les probabilités de transition de l'état caché courant pour obtenir un échantillon h_t, sachant l'état caché précédent: $\mathbf{P}(H_t \mid H_{t-1} = h_{t-1})$
    - » utiliser les probabilités de sortie de la variable d'observation étant donné l'état caché courant, pour obtenir le symbole d'observation (émission) s_t : $\mathbf{P}(S_t \mid H_t = h_t)$

- On peut aussi générer la séquence des états cachés d'abord et ensuite générer les observations
  - les variables cachées dépendent uniquement des variables cachées précédentes
  - chaque observation (émission) ne dépendra pas des autres

Illustration

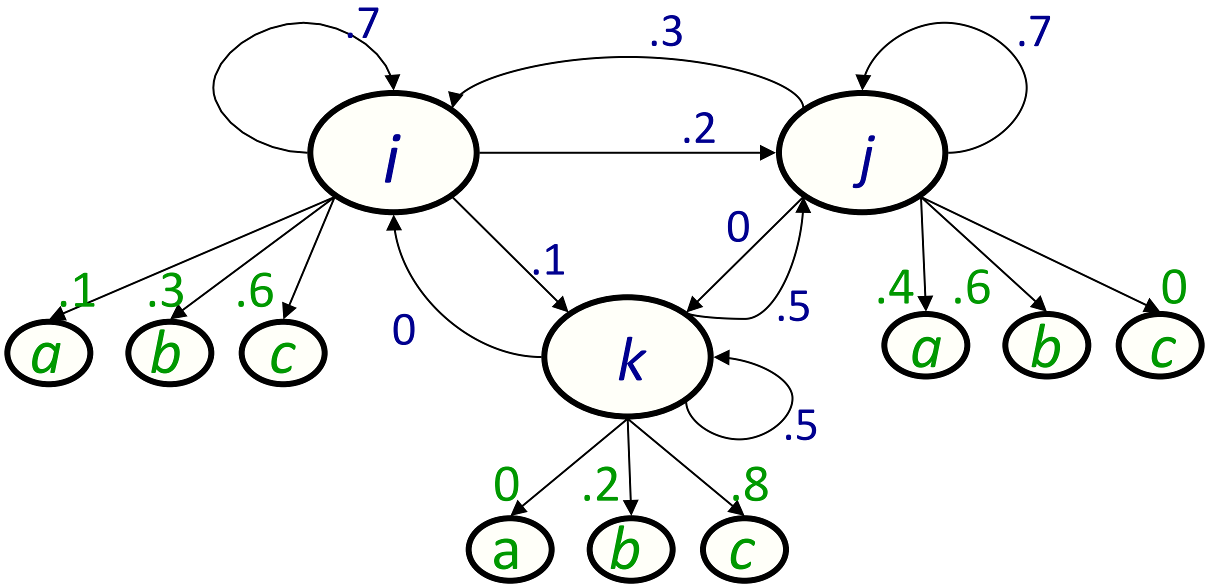

Illustration dans le cas d'une **chaîne infinie**, avec visualisation des valeurs de la variable cachée et la variable d'observation

Chaque **nœud caché** (valeur *i,j,k* possible de *H*) a un vecteur de **probabilités de transitions** et un **vecteur de probabilités d'émission (observations)**