

## Le modèle de Markov en 20 lignes...

### Quoi ? Pourquoi ? Comment ?

Le **modèle de Markov** associe des probabilités d'obtenir un état d'une chaîne de caractères à un processus de progression le long de celle-ci. Dans une séquence, le modèle est la progression. A chaque case (position de la séquence) je réalise **deux tirages**. Le **premier tirage** demande si on doit **passer à la case suivante**. Lui est associée la probabilité de passage à la case suivante. (Deux résultats possibles de ce tirage : je reste à la même case ou je passe à la suivante). Le **second est le tirage d'une base** azotée pour remplir la case (donc il peut y avoir jusqu'à 4 probabilités différentes associées à ce tirage : une pour chaque base).

**Le modèle de Markov est caché lors de l'observation d'une séquence** issue de l'échantillonnage expérimental. J'observe l'état de la séquence, **sans pouvoir voir le processus qui l'a engendré** (combien d'arrêt, combien de fois lorsque l'on est resté sur la même case la même base qu'au précédent tirage de base sur cette position a été retirée...).

Sur un ensemble de séquences de la même famille, je peux réaliser un alignement multiple et un profil. On constitue donc un jeu d'entraînement. **Ce jeu d'entraînement permet de déterminer "une machine" qui aurait pu construire ces séquences.** Autrement dit l'observation de séquences permet de régler les probabilités pour les deux tirages.

Soumettre une nouvelle séquence (la requête) a pour but de vérifier si cette "machine" (le modèle de markov que l'on a déterminé grâce au jeu d'entraînement) **pourrait avoir produit aussi cette séquence. La réponse est un indice de confiance probabiliste.**