

Algorithmes statistiques pour l'alignement multiple

Jotun Hein

Department of Statistics

University College

1 South Parks Road

Oxford OX1 3TG

UNITED KINGDOM

Résumé

Depuis une dizaine d'années on observe un intérêt croissant pour les méthodes statistiques dans le cadre de l'analyse de séquences. Ceci est moins vrai pour le problème de l'alignement de séquences. Dans cet exposé, nous présentons des algorithmes permettant de calculer la probabilité d'un ensemble de séquences reliées par un arbre binaire ayant évolué selon le modèle de Thorne-Kishino-Felsenstein pour un ensemble fixe de paramètres. Les récurrences sont basées sur une chaîne de Markov génératrice de séquences et sur leur alignement à des noeuds de l'arbre. Deux types de récurrences, de complexité similaires mais de probabilité différentes, sont obtenues. Les temps d'exécution des algorithmes sont en $O(L^k)$, où L est la taille des séquences observées et k est le nombre de séquences — feuilles de l'arbre binaire. Si le temps le permet, une approche "Gibbs sampling" et une extension du modèle de base seront discutés.