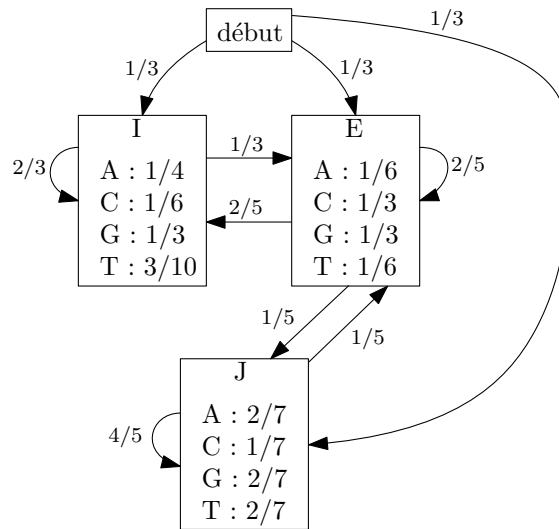


# BIN702 - Série d'exercices #5 : le modélisation de séquences avec modèles de Markov

Manuel Lafond

**Exercice 1:** Le modèle de Markov ci-bas tente, de façon complètement irréaliste, de modéliser une séquence où chaque nucléotide peut être dans un intron (I), un exon (E), ou une région inutile (J, pour “junk”).



- Calculez la probabilité de générer *ACCT* avec la séquence d'états *JEEI*.
- En utilisant l'algorithme de Viterbi, calculez le chemin le plus probable pour la séquence *ACCT*. La façon la plus simple est de remplir une matrice de programmation dynamique.
- Ce modèle ne tient pas compte des dépendances entre les caractères consécutifs. Par exemple, on pourrait modéliser le fait que *CG* est plus rare dans l'état *J* que dans l'état *E*. Proposez un modèle de Markov caché alternatif pour pallier ce manque (ne donnez que les états - inutile de tenter d'inférer les probabilités de transitions et d'émissions).