

μων που θα είναι κατάλληλα για την εργασία αυτή. Οι άνθρωποι συναρπάζονταν πάντα από την καταγωγή (των ειδών), και το μυστήριο της δικής μας καταγωγής έχει κυρίαρχη θέση. Επομένως, δεν περιμένω ότι θα υπάρξει έλλειψη ενδιαφέροντος για τέτοιες έρευνες.

11.7 Προβλήματα

Πρόβλημα 11.1

Το πρόβλημα της Αποκωδικοποίησης μπορεί να διατυπωθεί ως πρόβλημα μεγαλύτερης διαδρομής σε DAG. Αυτό αποτελεί κίνητρο για να βρεθεί μια αποδοτική ως προς το χώρο έκδοση του αλγορίθμου αποκωδικοποίησης. Υπάρχει όντως αλγόριθμος που χρησιμοποιεί γραμμικό χώρο για το πρόβλημα της Αποκωδικοποίησης;

Πρόβλημα 11.2

Ο κρουπιέρης στο καζίνο με τα δίκαια στοιχήματα κρατάει κάθε νόμισμα για τουλάχιστον δέκα ρίψεις, ανεξάρτητα από το αν είναι ιδανικό ή πειραγμένο, έτσι ώστε να μην προκαλέσει υποψίες. Περιγράψτε το αντίστοιχο HMM, καθώς και το πώς μπορεί να τροποποιηθεί ο αλγόριθμος αποκωδικοποίησης γι' αυτή την περίπτωση.

Πρόβλημα 11.3

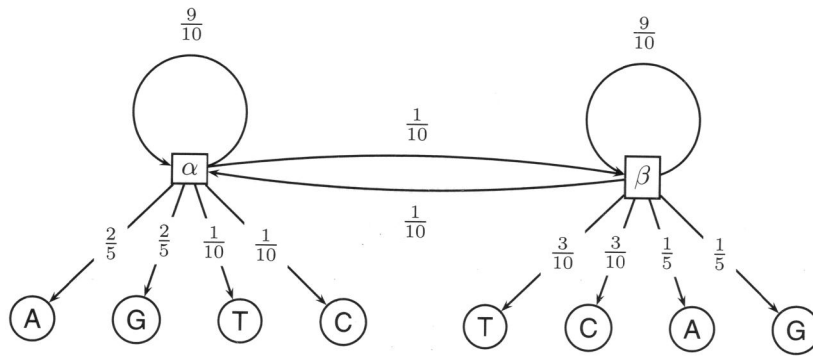
Έστω ότι σας δίνονται οι συχνότητες των δινουκλεοτιδίων μέσα και έξω από τα νησίδια CG. Σχεδιάστε ένα HMM για τον εντοπισμό των νησιδίων CG σε γονιδιωματικές αλληλουχίες.

Πρόβλημα 11.4

Στο Σχήμα 11.7 φαίνεται ένα HMM με δύο καταστάσεις α και β . Όταν το HMM βρίσκεται στην κατάσταση α , έχει μεγαλύτερη πιθανότητα να εκπέμψει πουρίνες (A και G). Όταν βρίσκεται στην κατάσταση β , έχει μεγαλύτερη πιθανότητα να εκπέμψει πυριμιδίνες (C και T). Αποκωδικοποιήστε την πιο πιθανή ακολουθία των καταστάσεων (α/β) για την αλληλουχία GGCT. Χρησιμοποιήστε λογαριθμικές βαθμολογίες αντί για κανονικές βαθμολογίες πιθανοτήτων.

Πρόβλημα 11.5

Έστω ότι ένας ανέντιμος κρουπιέρης έχει δύο νομίσματα, ένα ιδανικό και ένα πειραγμένο. Το πειραγμένο νόμισμα έχει πιθανότητα για κορώνα ίση με $1/4$.



Σχήμα 11.7 Το HMM που περιγράφεται στο Πρόβλημα 11.4.

Υποθέστε ότι ο κρουπιέρης δεν αλλάζει ποτέ νομίσματα. Ποιο νόμισμα έχει μεγαλύτερη πιθανότητα να παραγάγει την αλληλουχία ΗΤΤΗΗΗΗΤΤΤΗΤΗΗΤΤ; Είναι χρήσιμο να γνωρίζετε ότι $\log_2(3) = 1,585$.

Πρόβλημα 11.6

Θεωρήστε ένα διαφορετικό παιχνίδι στο οποίο ο κρουπιέρης δεν ρίχνει νόμισμα αλλά ζάρι με τρεις πλευρές που έχουν ετικέτες 1, 2, και 3. (Μην προσπαθήσετε να σκεφτείτε την εμφάνιση ενός τέτοιου ζαριού.) Ο κρουπιέρης έχει δύο στημένα ζάρια D_1 και D_2 . Για κάθε ζάρι D_i , η πιθανότητα να προκύψει ο αριθμός i είναι ίση με $1/2$, και η πιθανότητα για τα άλλα δύο αποτελέσματα είναι ίση με $1/4$. Σε κάθε γύρο, ο κρουπιέρης πρέπει να αποφασίσει αν (1) θα κρατήσει το ίδιο ζάρι, (2) θα αλλάξει ζάρι, ή (3) θα σταματήσει το παιχνίδι. Επιλέγει το (1) με πιθανότητα $1/2$ και τα (2) και (3) με πιθανότητα $1/4$. Στην αρχή, ο κρουπιέρης επιλέγει ένα από τα δύο ζάρια με την ίδια πιθανότητα.

- Διατυπώστε ένα HMM για την παραπάνω κατάσταση. Προσδιορίστε το αλφάβητο, τις καταστάσεις, τις πιθανότητες μεταβολής κατάστασης, και τις πιθανότητες εκπομπής. Συμπεριλάβετε μια αρχική κατάσταση *start*, και υποθέστε ότι το HMM ξεκινάει στην κατάσταση *start* με πιθανότητα 1. Συμπεριλάβετε επίσης και μια τελική κατάσταση *end*.
- Ας υποθέσουμε ότι παρατηρείτε την εξής ακολουθία από ρίψεις ζαριών: 1 1 2 1 2 2. Βρείτε μια ακολουθία καταστάσεων που εξηγεί καλύτερα την ακολουθία των ρίψεων. Ποια είναι η πιθανότητα της συγκεκριμένης ακολουθίας; Βρείτε την απάντηση συμπληρώνοντας τον πίνακα Viterbi. Συμπεριλάβετε βέλη οπισθοδρόμησης στα κελιά έτσι ώστε να είστε σε θέση να ανιχνεύσετε αντίστροφα την ακολουθία των καταστάσεων. Μερικά από τα παρακάτω δεδομένα ίσως φανούν χρήσιμα: