

10. Übung zur Vorlesung “Algorithmische Massenspektrometrie”

Wintersemester 2020/2021

Sebastian Böcker, Kai Dührkop

Ausgabe: 21. Januar 2021, Abgabe: 27. Januar 2021

1. **Teilsammenformeln** Berechnen Sie für die Molekülformel $C_6H_{12}O_6$ die Anzahl aller Teilsammenformeln. Geben Sie eine allgemeine Formel an, welche die Berechnung der Teilsammenformeln für ein beliebiges Molekül mit Summenformel aus den Elementen C,H und O erlaubt.

(3 Punkte)

2. **Maximaler Farbenfroher Teilbaum** Gegeben sei der Fragmentierungsgraph Figure 1. Der maximale farbenfrohe Teilbaum dieses Graphen ist der Teilbaum, mit maximalem Kantengewicht, in dem jede Farbe maximal einmal vorkommt.
 - (a) Warum muss der Teilbaum farbenfroh sein? Was würde passieren, wenn wir dieses Kriterium fallen lassen und warum wären die Lösungen dieses Maximalen Teilbaum-Problems keine geeigneten Interpretationen des Spektrums?

(1 Punkt)
 - (b) Warum suchen wir nach einem Baum? Welches Problem würde auftreten, wenn wir nach beliebigen zusammenhängenden farbenfrohen Teilgraphen suchen würden?

(1 Punkt)
 - (c) In dem Fragmentierungsgraphen in Figure 1 haben wir einige Kanten nicht eingezeichnet (gehen wir davon aus, dass diese einen sehr schlechten Score haben). Nennen sie ein Beispiel für eine solche weggelassene Kante.

(1 Punkt)
 - (d) In dem Graphen sind die Knoten nicht mit Summenformeln, sondern lediglich mit abstrakten Labels A bis J beschriftet. Wie könnten die Summenformeln für einen solchen Fragmentierungsgraphen aussehen? Geben Sie für die Knoten A bis J sowie die Wurzel R jeweils eine Summenformel an.

(2 Punkte)
 - (e) Wir betrachten die folgenden zwei Heuristiken zum Finden des maximal farbenfrohen Teilbaums:
 - i. **Greedy Heuristik:** Wir bezeichnen eine Kante als *zulässig*, wenn sie in einem Knoten des bisherigen Teilbaums anfängt und in einem Knoten endet, dessen Farbe im bisherigen Teilbaum noch nicht vorkommt. Die Heuristik beginnt mit einem Teilbaum bestehend aus der Wurzel. Aus allen *zulässigen* Kanten fügt nun Heuristik die *zulässige* Kante mit maximalem Gewicht in den bisherigen Teilbaum ein. Sie wiederholt dies, bis es keine *zulässigen* Kanten mehr gibt.

ii. **Kritischer Pfad Heuristik:** Wir bezeichnen einen Pfad als *zulässig*, wenn er, bis auf seinen Startknoten, nur Knoten besucht, deren Farben im bisherigen Teilbaum noch nicht vorkommen. Das Gewicht eines solchen Pfades ist die Summe seiner Kantengewichte. Die Heuristik beginnt mit einem Teilbaum bestehend aus der Wurzel. Aus allen *zulässigen* Pfaden, die in einem Knoten des bisherigen Teilbaums beginnen, fügt die Heuristik den Pfad mit maximalem Gewicht in den bisherigen Teilbaum ein. Sie wiederholt dies, bis es keine solchen *zulässigen* Pfade mehr gibt.

Berechnen Sie die zwei farbenfrohen Teilbäume, die sich aus den beiden Heuristiken ergeben. Welche Heuristik hat für diesen Fragmentierungsgraphen den besseren Baum gefunden?

(5 Punkte)

(f) Die Kritische Pfad Heuristik lässt sich mit Hilfe dynamischer Programmierung berechnen. Hier muss in jedem Schritt der Heuristik eine Tabelle D berechnet werden. Diese hat die folgende Definition: $D[u]$ ist das Gewicht des besten zulässigen Pfades, der im Knoten u beginnt. Geben Sie die Rekurrenz für D an.

(4 Punkte)

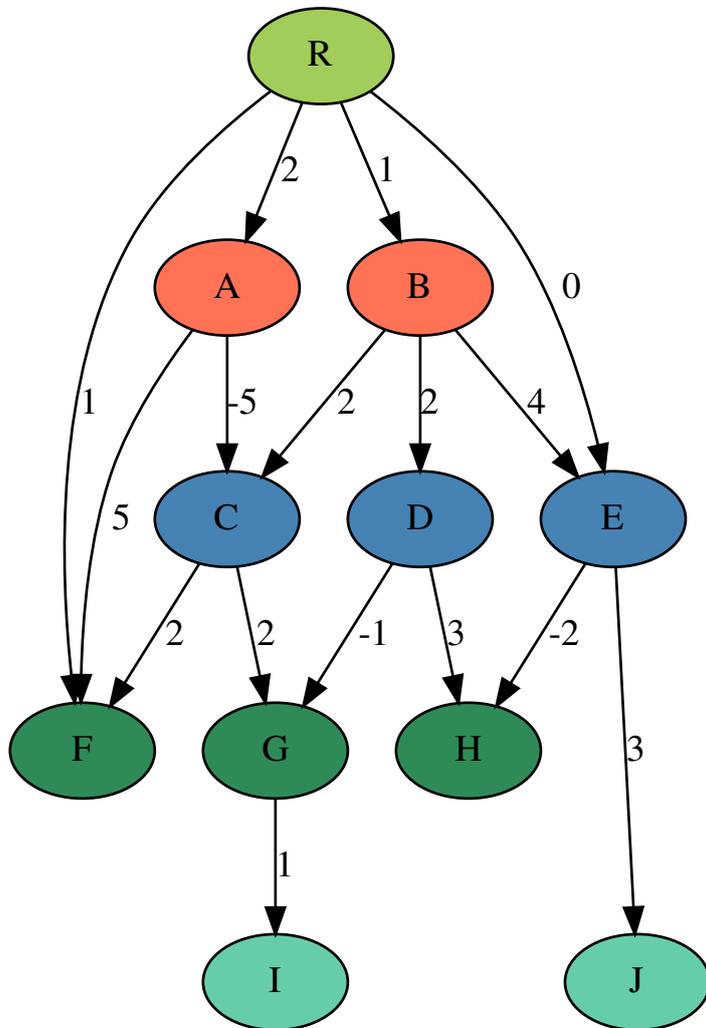


Abbildung 1: Fragmentierungsgraph. Die Knoten sind hier nur zur Vereinfachung mit den Labels A bis J beschriftet.