14. Übung (3.–7. Februar 2020)

Algorithmen und Datenstrukturen

Bitte beachten Sie die Klausurtermine und Raumzuordnungen auf der Lehrveranstaltungswebseite.Am 08. Februar um 14:00–16:00 wird ein Lernraum zur Vorbereitung der Klausur angeboten.

Aufgabe 1 (AGS 10.11)

Bei einem Spiel werfen Sie zwei unabhängige Münzen und erhalten einen Gewinn, wenn nach dem Wurf beide Münzen *auf der gleichen Seite landen*. Sie können bei diesem Spiel nur beobachten, ob Sie gewonnen oder verloren haben. Nehmen Sie an, dass die erste Münze sehr dick ist und daher beim Werfen auch auf dem Rand R landen kann. Die Menge der möglichen Ergebnisse ist daher $X = \{K, Z, R\} \times \{K, Z\}$.

(a) Geben Sie den Analysator A für dieses Szenario an.

(b) Sie spielen das Spiel 24 Mal und gewinnen 6 Mal. Geben Sie den Korpus h mit unvollständigen Daten an.

(c) Gegeben ist die initiale Wahrscheinlichkeitsverteilung $q_0 = q_0^1 \times q_0^2$ über den vollständigen Daten, mit $q_0^1(K) = 2/5$, $q_0^1(R) = 1/5$ und $q_0^2(K) = 1/3$. Dabei ist q_0^1 die Wahrscheinlichkeitsverteilung der ersten und q_0^2 die der zweiten Münze. Führen Sie den E-Schritt des EM-Algorithmus aus, erweitern Sie also den Korpus h zum Korpus h_1 über den vollständigen Daten.

(d) Führen Sie nun den M-Schritt aus. Bestimmen Sie dafür zunächst die Teilkorpora h_1^1 und h_1^2 für die erste bzw. zweite Münze.

(e) Schätzen Sie nun die Wahrscheinlichkeitsverteilungen q_1^1 und q_1^2 der beiden Münzen, indem Sie die relative Häufigkeit der Teilkorpora bestimmen.

Aufgabe 2 (AGS 10.19 $\star$)

Die Personen A und B spielen ein Spiel mit einer Münze mit den Beschriftungen 1 und 2, sowie mit einem dreiseitigen „Würfel“ mit den Beschriftungen 1, 2, und 3. In jeder Runde werden die Münze und der Würfel geworfen. Der Spieler A gewinnt die Runde, falls die gefallene Zahl des Würfels kleiner gleich der auf der Münze ist. Ansonsten gewinnt B die Runde. Die Menge der möglichen Ergebnisse ist somit $X = \{1, 2\} \times \{1, 2, 3\}$.

(a) Geben Sie den Analysator A für dieses Szenario an.

Die Spieler spielen 31 Runden des Spiels und A gewinnt dabei 21 Mal.

(b) Geben Sie den Korpus h mit unvollständigen Daten an.

Wir wollen aus diesem Korpus h mit dem EM-Algorithmus die Wahrscheinlichkeitsverteilungen von Münze und Würfel bestimmen. Die initiale Wahrscheinlichkeitsverteilung $q_0 = q_0^M \times q_0^W$ ist gegeben durch $q_0^M(1) = 1/3$ und $q_0^W(1) = q_0^W(3) = 1/4$. Dabei ist q_0^M die Wahrscheinlichkeitsverteilung der Münze, und q_0^W die des Würfels.

(c) Geben Sie q_0 an.

(d) Führen Sie den E-Schritt aus. Vervollständigen Sie also den Korpus h zum Korpus h_1.

(e) Führen Sie nun den M-Schritt aus. Bestimmen Sie dafür zunächst die Teilkorpora h_1^M und h_1^W für die Münze bzw. den Würfel.

(f) Schätzen Sie nun die Wahrscheinlichkeitsverteilung q_1^M der Münze sowie q_1^W des Würfels, indem Sie die relative Häufigkeit der Teilkorpora bestimmen.

Zusatzaufgabe 1 (AGS 7.2.5 $\star$)

Gegeben seien die Wörter $w = \texttt{espen}$ und $v = \texttt{beispiele}$.

(a) Berechnen Sie die Levenshtein-Distanz $d(w, v)$. Geben Sie dazu die Berechnungsmatrix an. Tragen Sie alle Zelleneinträge zusammen mit den dazugehörigen Pfeilen ein.

(b) Geben Sie die Levenshtein-Distanz $d(\texttt{espe}, \texttt{beispiel})$ an. Beachten Sie, dass `espe` und `beispiel` Präfixe von `espen` bzw. `beispiele` sind.

(c) Geben Sie zwei Alignments zwischen `espen` und `beispiele` an, die zu den minimalen Kosten führen. Dabei sollen die Alignments die jeweils angewendeten Editieroperation enthalten.

(d) Wieviele Backtraces enthält die in Aufgabenteil 2 (a) angegebene Berechnungsmatrix?

Zusatzaufgabe 2 (AGS 8.20 $\star$)

Fügen Sie in die folgenden AVL-Bäume den jeweils angegebenen Schlüssel ein. Stellen Sie nach jedem Einfügen die AVL-Eigenschaft her und dokumentieren Sie hierbei die vom Einfüge- und Balancierungsalgorithmus ausgeführten Operationen. Nutzen Sie dabei die folgenden Abkürzungen:

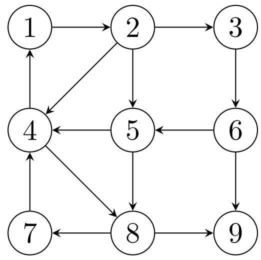
- i(x) – für das Einfügen des Schlüssels x
- L(x) – für die Linksrotation um den Knoten mit dem Schlüsselwert x,
- R(x) – für die Rechtsrotation um den Knoten mit dem Schlüsselwert x.

Geben Sie unmittelbar nach jedem i(x)-Schritt die Balancefaktoren für alle *relevanten* Knoten auf dem Pfad von x zur Wurzel an.

```
   11                  11                 6                       6
  /  \     i(12)      /  \    i(9)       / \           i(12)     / \          i(10)
 8    27   ----->    8    27  ---->     2   8          ----->   2   8         ----->
     /                   /                 / \                   \   \
   13                  13                 7   11                  4   11
```

Zusatzaufgabe 3 (AGS 9.2.14 $\star$)

Der gerichtete Graph G sei durch folgende Darstellung gegeben:



In den folgenden Teilaufgaben müssen Sie lediglich Kanten in den Vordruck eintragen. Zwischenschritte zu den Lösungen brauchen Sie nicht anzugeben.

(a) Wenden Sie auf G wiederholt den DFS-Algorithmus mit dem **Startknoten 2** an und bestimmen Sie auf diese Weise **vier** unterschiedliche depth-first-trees.

(b) Wenden Sie auf G wiederholt den BFS-Algorithmus mit dem **Startknoten 2** an und bestimmen Sie auf diese Weise **drei** unterschiedliche breadth-first-trees.

Zusatzaufgabe 4 (AGS 9.5.32 $\star$)

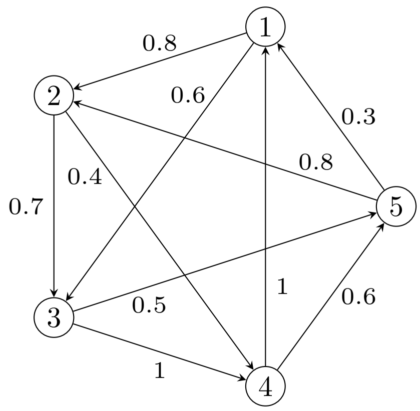
Gegeben ist der Viterbi-Semiring $([0,1], \max, \cdot, 0, 1)$ und der gewichtete Graph G über dem Viterbi-Semiring in der nebenstehenden Abbildung.

(a) Geben Sie die modifizierte Adjazenzmatrix von G vollständig an.

(b) Die Matrix $D_G^{(2)}$ ist gegeben. Geben Sie $D_G^{(3)}$ vollständig an!

(c) Welche Änderungen ergeben sich in der Berechnung von $D_G^{(5)}$ im Vergleich zu $D_G^{(4)}$?

(d) Gegeben sei das Alphabet $\Sigma = \{a, b, c, d\}$, der Semiring der formalen Sprachen $(\mathcal{P}(\Sigma^*), \cup, \cdot, \emptyset, \{\varepsilon\})$ über dem Alphabet Σ und die Matrix $D_{G'}^{(2)}$ eines gewichteten Graphen G' über diesem Semiring. Geben Sie die Werte $D_{G'}^{(3)}(u, v)$ für $u \in \{1, 2, 3\}$ und $v \in \{1, 2\}$ an!



$$D_G^{(2)} = \begin{pmatrix} 1 & 0.8 & 0.6 & 0.32 & 0 \\ 0 & 1 & 0.7 & 0.4 & 0 \\ 0 & 0 & 1 & 1 & 0.5 \\ 1 & 0.8 & 0.6 & 1 & 0.6 \\ 0.3 & 0.8 & 0.56 & 0.32 & 1 \end{pmatrix} \qquad D_{G'}^{(2)} = \begin{pmatrix} \{\varepsilon\} & \emptyset & \{a\} \\ \emptyset & \{\varepsilon\} & \{d\} \\ \emptyset & \{c\} & \{\varepsilon, b, cd\} \end{pmatrix}$$