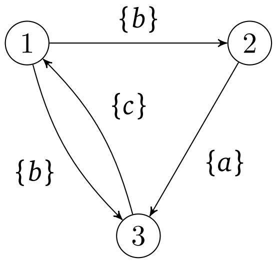
Algorithmen und Datenstrukturen

Aufgabe 1 (AGS 9.4.19)

Der gewichtete Graph $G = (V, E, c)$ über dem Semiring $(\mathscr{P}(\{a, b, c\}^*), \cup, \circ, \emptyset, \{\varepsilon\})$ sei durch folgende graphische Darstellung gegeben:



Es soll für den Graph G das Prozessproblem mit Hilfe des Aho-Algorithmus gelöst werden.

(a) Geben Sie für G die modifizierte Adjazenzmatrix mA_G an.

(b) Berechnen Sie für den Aho-Algorithmus die Matrizen $D_G^{(1)}$ und $D_G^{(2)}$. Sie müssen nur die Matrixelemente aufschreiben, die sich gegenüber mA_G geändert haben.

(c) Geben Sie die letzte Zeile der Ergebnismatrix D_G des Aho-Algorithmus an.

(d) Wie verändert sich der Eintrag $D_G(3,3)$, wenn zu dem Graphen G eine Kante von Knoten 3 zum Knoten 2 mit dem Gewicht $\{b\}$ zugefügt wird?

Aufgabe 2 (AGS 10.7 $\star$)

Bei beschränkten Modellen kann der Maximum-Likelihood-Schätzer im Allgemeinen nicht effizient bestimmt werden. Beinhaltet das Modell $\mathscr{M}$ aber die relative Häufigkeitsverteilung $\mathsf{rfe}(h)$ des gegebenen Korpus h, dann ist $\mathsf{mle}(\mathscr{M}, h) = \mathsf{rfe}(h)$, denn keine andere Wahrscheinlichkeitsverteilung erzeugt eine höhere Likelihood.
Bestimmen Sie für die folgenden Situationen das Wahrscheinlichkeitsmodell, zeigen Sie, dass die relative Häufigkeitsverteilung des betrachteten Korpus in diesem enthalten ist und bestimmen Sie den Maximum-Likelihood-Schätzer.

(a) Werfen eines Würfels, bei dem gegenüberliegende Seiten die gleiche Wahrscheinlichkeit aufweisen. Betrachten Sie den folgenden Korpus:

$$h(1) = 3\,, \quad h(2) = 5\,, \quad h(3) = 1\,, \quad h(4) = 1\,. \quad h(5) = 5\,, \quad h(6) = 3\,.$$

(b) Werfen zweier unabhängiger Münzen. Betrachten Sie den folgenden Korpus:

$$h(K,K) = 2\,, \quad h(K,Z) = 4\,, \quad h(Z,K) = 4\,, \quad h(Z,Z) = 8\,.$$

(c) Ziehen mit Zurücklegen aus einer Urne mit fünf Kugeln. Die Kugeln sind weiß, schwarz oder rot. Betrachten Sie den folgenden Korpus:

$$h(W) = 4\,, \quad h(S) = 2\,, \quad h(R) = 4\,.$$

Aufgabe 3 (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 4 (AGS 10.16)

Bei einem Spiel drehen Sie zwei Glücksräder jeweils ein Mal. Jedes Glücksrad hat drei einfarbige Felder in den Farben rot (r), gelb (g) und blau (b). Die Menge der möglichen Ereignisse ist somit $X = \{r, g, b\} \times \{r, g, b\}$.

Sie gewinnen genau dann, wenn Sie zwei Felder *verschiedener* Farbe erhalten. Der Spielleiter informiert Sie nur darüber, ob Sie gewonnen oder verloren haben; die Farben der erhaltenen Felder können Sie nicht wahrnehmen.

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

(b) Sie spielen 40 Runden und gewinnen dabei 16 mal. Geben Sie den entsprechenden Korpus h mit unvollständigen Daten an.

(c) Wir wollen aus diesem Korpus h mit dem EM-Algorithmus die Wahrscheinlichkeitsverteilungen der beiden Glücksräder bestimmen. Die initiale Wahrscheinlichkeitsverteilung $q_0 = q_0^1 \times q_0^2$ ist gegeben durch $q_0^1(r) = 1/2$, $q_0^1(g) = 1/3$, $q_0^2(r) = 1/4$, und $q_0^2(g) = 1/2$. Dabei ist q_0^i die Wahrscheinlichkeitsverteilung des i-ten Glücksrades. Geben Sie q_0 vollständig an.

(d) Führen Sie den E-Schritt aus, vervollständigen Sie also den Korpus h zum Korpus h_1 und geben Sie diesen an.

(e) Führen Sie nun den M-Schritt aus. Bestimmen Sie dafür zunächst die Teilkorpora h_1^1 und h_1^2 für die zwei Glücksräder.

(f) Schätzen Sie nun die Wahrscheinlichkeitsverteilungen q_1^1 und q_1^2 der beiden Glücksräder, indem Sie die relativen Häufigkeiten der Teilkorpora bestimmen.

Zusatzaufgabe 1 (AGS 4.23 ⋆)

Gegeben sei folgendes C-Programm:

```
1  #include <stdio.h>
2
3  void g (int *x, int y);
4
5  void f (int z) {
6    int a = 1;
7    while (z != 0) {
8      /*label1*/
9      g(&z, a); /*$1*/
10     a = a * 2;
11   }
12   /*label2*/
13 }
14
15 void g (int *x, int y) {
16   /*label3*/
17   if (y >= 2) {
18     *x = 0;
19     /*label4*/
20     f(*x); /*$2*/
21   } else {
22     *x = *x + 1;
23   }
24   /*label5*/
25 }
26
27 int main () {
28   int m;
29   m = 2;
30   /*label6*/
31   f(m); /*$3*/
32   /*label7*/
33   return 0;
34 }
```

(a) Tragen Sie den Gültigkeitsbereich jedes Objektes in eine Tabelle ein. Nutzen Sie dazu die Zeilennummern.

(b) Setzen Sie das folgende Speicherbelegungsprotokoll fort. Dokumentieren Sie die aktuelle Situation beim Passieren der Marken (*label1* bis *label7*). Geben Sie jeweils den Rücksprungmarkenkeller und die *sichtbaren* Variablen mit ihrer Wertebelegung an. Die Inhalte von Speicherzellen nicht sichtbarer Variablen müssen Sie nur bei Änderungen eintragen. Die bereits festgelegten Rücksprungmarken sind *$1* bis *$3*.

Label	RM	1	2	3	4	5	6	7	8	9	10
label6	–	m 2									

Zusatzaufgabe 2 (AGS 6.2.14)

Gegeben sei die Folge $4, 7, 2, 1, 6, 3, 5, 8, 0, 9$. Wenden Sie auf diese Folge den Heapsort-Algorithmus an.

Zusatzaufgabe 3 (AGS 7.1 ⋆)

Geben Sie zu den Pattern

(a) `abaabaaab`

(b) `aaabaaaa`

die jeweils mit Hilfe des KMP-Algorithmus (Knuth-Morris-Pratt) berechnete Verschiebetabelle an.