



Übungsblatt 9

Datenstrukturen und Algorithmen (SS 2018)

Abgabe: Mittwoch, 04.07.2018, 23:55 Uhr — Besprechung: ab Montag, 09.07.2018

Bitte lösen Sie die Übungsaufgabe in **Gruppen von 3 Studenten** und wählen EINEN Studenten aus, welcher die Lösung im ILIAS als **PDF** als **Gruppenabgabe** (unter Angabe aller Gruppenmitglieder) einstellt. Bitte erstellen Sie dazu ein **Titelblatt**, welches die Namen der Studenten, die Matrikelnummern, und die E-Mail-Adressen enthält.

Die Aufgaben mit Implementierung sind mit Impl gekennzeichnet. Das entsprechende Eclipse-Projekt kann im ILIAS heruntergeladen werden. Bitte beachten Sie die Hinweise zu den Implementierungsaufgaben, die im ILIAS verfügbar sind.¹

Dieses Übungsblatt beinhaltet 3 Aufgaben mit einer Gesamtzahl von 30 Punkten.

Aufgabe 1 Textalgorithmen: Knuth-Morris-Pratt [Punkte: 9]

Verwenden Sie folgende Definitionen und gehen Sie von 0-basierten Indizes aus:

$p_1 = \text{„gghf gghf g f“}$; $t_1 = \text{„gghfhgghgghf gghf gghf gghgghf gghf gffgfhgh“}$

- (4 Punkte) Stellen Sie die Präfixtabelle für das Pattern p_1 auf.
- (5 Punkte) Führen Sie den Knuth-Morris-Pratt-Algorithmus durch, um p_1 in t_1 zu finden. Stellen Sie dar, wie Sie vorgegangen sind und fügen Sie Erklärungen hinzu, wo diese zum Verständnis notwendig sind. Orientieren Sie sich dabei an der Darstellungsweise aus der Vorlesung.

Aufgabe 2 Textalgorithmen: Boyer-Moore [Punkte: 9]

- (5 Punkte) Geben Sie für das Suchmuster „piopippi“ die entsprechende shift und last Tabelle an.
- (4 Punkte) Verwenden Sie folgende Definitionen und gehen Sie von 0-basierten Indizes aus:
 $p_2 = \text{„werewerre“}$; $t_2 = \text{„wererrewerrexwerrewxrewerewerrer“}$

Führen Sie den Boyer-Moore-Algorithmus aus, um p_2 in t_2 zu finden. Stellen Sie dar, wie Sie vorgegangen sind, und fügen Sie Erklärungen hinzu, wo diese zum Verständnis notwendig sind. Geben sie bei jedem Sprung an, wie groß die Beiträge aus den Heuristiken (Good Suffix/Match und Bad Character/Occurence) des Algorithmus sind. Greifen Sie auf die folgenden Tabellen für p_2 zurück:

	<i>c</i>	w e r x							
	last[<i>c</i>]	4	8	7	-1				
<i>i</i>	0	1	2	3	4	5	6	7	8
$p_2[i]$	w	e	r	e	w	e	r	r	e
shift[<i>i</i>]	9	9	9	9	9	9	5	3	1

¹https://ilias3.uni-stuttgart.de/goto_Uni_Stuttgart_crs_1442449.html

Aufgabe 3 Reguläre Ausdrücke [*Punkte: 12*]

Geben Sie reguläre Ausdrücke an, welche genau die im Folgenden beschriebenen Sprachen (Mengen von Wörtern) erkennen. Verwenden Sie die Syntax aus dem Skript, Folie *Reguläre Ausdrücke: Notation* (V17). Gehen Sie davon aus, dass jedes Wort eine Zeile für sich hat und dass somit \wedge den Wortanfang und $\$$ das Wortende markieren.

- (a) (*2 Punkte*) Alle Wörter der Länge 5 die mit einem a beginnen und auf a oder b enden.
- (b) (*2 Punkte*) Alle Wörter, die mit *Apfel*, *Kirsch*, *Birn* oder *Obst* beginnen, gefolgt von *baum* und optional mit *plantage* enden.
- (c) (*3 Punkte*) Alle aufgezählten Schreibweisen des Vornamens Rebekka: *Rebekka*, *Rebecca*, *Rebec-
cah*, *Rebecka*, *Rebeka*, *Rebekah*. Vermeiden Sie hier, die Zeichenfolgen (also z.b. *Rebekka*) direkt im regulären Ausdruck zu verwenden.
- (d) (*5 Punkte*) Alle Wörter, die eine Zeichenkette beinhalten, die entweder mit X , Y oder Z beginnt, gefolgt von mindestens einem a und danach optional einem b . Alternativ kann die Zeichenkette mit *Bla* beginnen gefolgt von einer geraden Anzahl (größer 0) von *bla*'s. Unabhängig von dieser Zeichenkette sollen alle Wörter mit *ooh* und danach 2 beliebigen Zeichen enden.