

## Übungsblatt 5

**Aufgabe 1** Sei **WS** die Menge aller Whitespaces. Betrachten Sie das Wort  $w \in (\{a, \dots, z, 0, \dots, 9, <, =, :\} \cup \mathbf{WS})^*$  einer Programmiersprache mit

$w = \mathbf{if} \ x \leq 5 \ \mathbf{then} \ \mathbf{val} := 10$

- Geben Sie eine geeignete Tokenfolge für  $w$  an.
- Welche Tokenfolge erhalten Sie, wenn sie alle Whitespaces aus  $w$  entfernen?
- Entwerfen Sie reguläre Ausdrücke, die die gültigen Tokens der Programmiersprache beschreiben, und geben Sie diesen Prioritäten.
- Geben Sie intuitiv einen deterministischen Automaten an, der das Tokenizing durchführt. Geben Sie für jeden Endzustand an, welche Art Token in ihm erkannt wird.

**Aufgabe 2** Seien  $r_1, r_2 \in \mathcal{E}_{\{a,b\}}$  gegeben durch

$$\begin{aligned} r_1 &= a \\ r_2 &= a^*b \end{aligned}$$

- Definieren Sie den Scanner zu  $r_1|r_2$ .
- Bestimmen Sie die Tokenfolge zu dem Wort  $aaabaabaaa$ .
- Führen Sie den naiven Tokenizing-Algorithmus für  $w = aaaa$  durch.
- Führen Sie anschließend den Maximal-Munch-Algorithmus für  $w$  durch.