6 FORMALE SPRACHEN

6.1 OPERATIONEN AUF FORMALEN SPRACHEN

6.1.1 Produkt oder Konkatenation formaler Sprachen

Produkt von Sprachen

- Def: $L_1 \cdot L_2 = \{w_1 w_2 \mid w_1 \in L_1 \text{ und } w_2 \in L_2\}$
- Beispiel: noch mal: formale Sprache L aller Wörter über $A = \{\mathtt{a}, \mathtt{b}\}$, in denen nirgends das Teilwort `ab` vorkommt.
  kann man jetzt so schreiben: $L = \{\mathtt{b}\}^*\{\mathtt{a}\}^*$
- Beispiel: Menge aller Wörter über A außer dem leeren: $A \cdot A^*$, denn jedes nichtleere Wort hat ein erstes Symbol und dahinter kommt ein beliebiges (evtl auch leeres) Wort
- formale Sprache L_I der legalen Zahlen vom Typ **int**:
  - Versuch: $A = \{\mathtt{0}, \dots, \mathtt{9}\}$
    $L_I = A \cdot A^*$
  - Was fehlt? jedenfalls das Minuszeichen;
    besser: $\{\varepsilon, \mathtt{-}\} \cdot A \cdot A^*$
  - Was ist mit Präfix `0x`? gibts den? ich weiß es nicht
- formale Sprache L_V der legalen Variablennamen in Java:
  - Versuch: $A = \{\mathtt{\_}, \mathtt{a}, \dots, \mathtt{z}, \mathtt{A}, \dots, \mathtt{Z}\}$, $B = A \cup \{\mathtt{0}, \dots, \mathtt{9}\}$
    $L_V = A \cdot B^*$.
  - es fehlen die Umlaute, ...
  - Was ist noch falsch? z. B. Schlüsselwörter (**if**, ...) sind als Variablennamen verboten.
    also eher sowas wie $L_V = (A \cdot B^*) \setminus \{\mathtt{if}, \mathtt{class}, \dots\}$
    Mitteilung: da könnte man jetzt alle endlich vielen Schlüsselwörter aufzählen, aber wenn nur endlich viele Wörter verboten sind, geht es im Prinzip ohne Mengendifferenz
- $L_1 = \{\mathtt{a}^n \mid n \in \mathbb{N}_0\}$ und $L_2 = \{\mathtt{b}^n \mid n \in \mathbb{N}_0\}$
  Achtung: $L_1 L_2 = \{\mathtt{a}^k \mathtt{b}^m \mid k \in \mathbb{N}_0 \text{ und } m \in \mathbb{N}_0\}$ die **Exponenten können verschieden sein!**
  hier steht nichts anderes als $\{\mathtt{a}\}^*\{\mathtt{b}\}^*$

Potenzen von L

- Def: $L^0 = \{\varepsilon\}$ und $L^{i+1} = L^i \cdot L$
- Beispiel: $L = \{\mathtt{a}\}^*\{\mathtt{b}\}^*$
  dann enthält z. B.
  - $L^0 = \{\varepsilon\}$
  - $L^1 = L = \{\varepsilon, \mathtt{aabbbbb}, \mathtt{aaaaab}, \mathtt{aaaa}, \mathtt{bbbbbbbb}, \dots\}$
  - $L^2 = \{\varepsilon, \mathtt{aabbbbaaaaab}, \mathtt{aaaaabab}, \mathtt{aaaaa}, \mathtt{bbbbbbb}, \dots\}$
  - usw.

6.1.2 Konkatenationsabschluss einer formalen Sprache

Kokatenationsabschluss

- Def:
$$L^+ = \bigcup_{i \in \mathbb{N}_+} L^i \qquad \text{und} \qquad L^* = \bigcup_{i \in \mathbb{N}_0} L^i$$
- Beispiel: $L = \{\mathtt{a}\}^*\{\mathtt{b}\}^*$
  - Man mache sich klar: $L^* = \{\mathtt{a}, \mathtt{b}\}^*$, also alles
  - da geht z. B. so: (die Studenten möglichst selber drauf kommen lassen) zerhacke beliebiges aber festes $w \in \{\mathtt{a}, \mathtt{b}\}^*$ an allen Stellen, wo Teilwort $\mathtt{ba}$ vorkommt, zwischen dem $\mathtt{b}$ und dem $\mathtt{a}$. Die entstehenden Teilwörter sind aus L.
- Man beweise: $L^* \cdot L = L^+$
  - Wie beweist man, dass zwei Mengen gleich sind?
  - Zum Beispiel, indem man zeigt, dass $\subseteq$ und $\supseteq$ gelten.
  - Also:
    * $\subseteq$:
      Wenn $w \in L^* \cdot L$, dann $w = w'w''$ mit $w' \in L^*$ und $w'' \in L$.
      Also existiert ein $i \in \mathbb{N}_0$ mit $w' \in L^i$.
      Also $w = w'w'' \in L^i \cdot L = L^{i+1}$.
      Da $i + 1 \in \mathbb{N}_+$, ist $L^{i+1} \subseteq L^+$, also $w \in L^+$.
    * $\supseteq$: Wenn $w \in L^+$, dann existiert ein $i \in \mathbb{N}_+$ mit $w \in L^i$.
      Da $i \in \mathbb{N}_+$ ist $i = j + 1$ für ein $j \in \mathbb{N}_0$,
      also ist für ein $j \in \mathbb{N}_0$: $w \in L^{j+1} = L^j \cdot L$.
      also $w = w'w''$ mit $w' \in L^j$ und $w'' \in L$.
      Wegen $L^j \subseteq L^*$ ist $w = w'w'' \in L^* \cdot L$.