6.11 Infix-Grammatiken

SML erlaubt *Infix-Ausdrücke* wie beispielsweise

```
4 + 6*7 - 9*3
```

Dabei heißen die Symbole (+), (*) und (−) *Infixoperatoren*. Selbstverständlich gibt es feste Regeln, wie solche Infix-Notation zu verstehen ist. Der erste Schritt besteht darin, zu einer *vollständig geklammerten Form* überzugehen. Für den obigen Ausdruck könnte diese wie folgt aussehen:

```
((4 + (6*7)) - (9*3))
```

Diese Klammerung wäre durch die folgenden 2 Annahmen gerechtfertigt:

1. Die Präzedenz von (*) ist größer als die von (+) und (−).

2. (+) und (−) haben die gleiche Präzedenz und sind beide linksassoziativ.

Wir sagen das im Konjunktiv, da Programme die Präzedenzen und Assoziativitäten von Infixoperatoren redeklarieren können. Die Präzedenz eines Operators regelt, wie stark er im Vergleich mit anderen Operatoren bindet. Die Assoziativität eines Operators ist links oder rechts und regelt, nach welcher Richtung bei wiederholter Verwendung des Operators geklammert wird. Wäre (+) rechtassoziativ, dann würde

```
1+2+3+4
```

zu

```
(1+(2+(3+4)))
```

geklammert.

Um Infix-Notation präzise zu erklären, benötigen wir das abstrakte Konzept einer Infix-Grammatik.

Definition 24 *Eine* Infix-Grammatik *ist ein Tupel* $G = (A, O, \mathrm{Pr}, \mathrm{As})$ *wie folgt:*

1. *A und O sind disjunkte Menge. Die Elemente von A heißen* Atome *und die Elemente von O heißen* Operatoren.

2. $\mathrm{Pr} \in O \to \mathbf{Z}$. *Die Zahl* $\mathrm{Pr}(o)$ *heißt* Präzedenz *von o.*

3. $\mathrm{As} \in O \to \{\mathtt{l}, \mathtt{r}\}$. *Das Symbol* $\mathrm{As}(o)$ *heißt* Assoziativität *von o. Ein Operator o heißt* linksassoziativ, *wenn* $\mathrm{As}(o) = \mathtt{l}$ *und* rechtsassoziativ, *wenn* $\mathrm{As}(o) = \mathtt{r}$.

Definition 25 *Sei* $G = (A, O, \mathrm{Pr}, \mathrm{As})$ *eine Infix-Grammatik. Dann ist die Relation* $\vdash \in O \times \{\mathtt{l}, \mathtt{r}\} \times \mathrm{Baum}(A \cup O)$ *wie folgt definiert:*

$$\frac{a \in A}{o, \alpha \vdash (a, \varepsilon)}$$

$$\frac{\mathrm{Pr}(o) = \mathrm{Pr}(o') \qquad \mathrm{As}(o) = \mathrm{As}(o') = \alpha \qquad o', \mathtt{l} \vdash b_1 \qquad o', \mathtt{r} \vdash b_2}{o, \alpha \vdash (o', b_1 b_2)}$$

$$\frac{\mathrm{Pr}(o) < \mathrm{Pr}(o') \qquad o', \mathtt{l} \vdash b_1 \qquad o', \mathtt{r} \vdash b_2}{o, \alpha \vdash (o', b_1 b_2)}$$

Die Menge der Ableitungsbäume von G *ist wie folgt definiert:*

$$\mathrm{ABaum}(G) \;=\; \{b \in \mathrm{Baum}(A \cup O) \mid \exists o \in O \;\; \exists \alpha \in As\colon \;\; o, \alpha \vdash b\}$$

Die Infixprojektion eines Ableitungsbaums *ist wie folgt definiert:*

$$\mathrm{Inp} \in \mathrm{ABaum}(G) \to (A \cup O)^*$$

$$\begin{array}{rcl} \mathrm{Inp}((a,\varepsilon)) & = & (a,\varepsilon) \\ \mathrm{Inp}((o,b_1b_2)) & = & \mathrm{Inp}(b_1)\cdot(o,\varepsilon)\cdot\mathrm{Inp}(b_2) \end{array}$$

Die Ableitungsbäume einer Infix-Grammatik entsprechen vollständig geklammerten Infix-Ausdrücken:

$$\begin{array}{rcl} \mathrm{Ina}((a,\varepsilon)) & = & a \\ \mathrm{Ina}((o,\, b_1b_2)) & = & (\mathrm{Ina}(b_1) \;\; o \;\; \mathrm{Ina}(b_2)) \end{array}$$

Proposition 19 *Sei* $G = (A, O, \mathrm{Pr}, \mathrm{As})$ *eine Infix-Grammatik und* $b = (o, b_1b_2) \in \mathrm{ABaum}(G)$. *Dann gilt:*

1. *Für jeden Operator* o'*, der in* $\mathrm{Inp}(b)$ *vorkommt, gilt* $\mathrm{Pr}(o') \geq \mathrm{Pr}(o)$.
2. *Sei* $\mathrm{As}(o) = \mathtt{l}$. *Dann gilt:*
   (a) *Für jeden Operator* o'*, der in* $\mathrm{Inp}(b_2)$ *vorkommt, gilt* $\mathrm{Pr}(o') > \mathrm{Pr}(o)$.
   (b) *Für jeden Operator* o' *mit* $\mathrm{Pr}(o) = \mathrm{Pr}(o')$*, der in* $\mathrm{Inp}(b_1)$ *vorkommt, gilt* $\mathrm{As}(o') = \mathrm{As}(o)$.
3. *Sei* $\mathrm{As}(o) = \mathtt{r}$. *Dann gilt:*
   (a) *Für jeden Operator* o'*, der in* $\mathrm{Inp}(b_1)$ *vorkommt, gilt* $\mathrm{Pr}(o') > \mathrm{Pr}(o)$.
   (b) *Für jeden Operator* o' *mit* $\mathrm{Pr}(o) = \mathrm{Pr}(o')$*, der in* $\mathrm{Inp}(b_2)$ *vorkommt, gilt* $\mathrm{As}(o') = \mathrm{As}(o)$.

Die Aussagen der Proposition beweist man leicht durch Induktion über $|b|$.

Die nächste Proposition stellt fest, dass eine Infix-Grammatik eine partielle Funktion von ungeklammerten Infix-Ausdrücken auf vollständig geklammerten Infix-Ausdrücke definiert.

Proposition 20 *Sei* $G = (A, O, \mathrm{Pr}, \mathrm{As})$ *eine Infix-Grammatik. Dann existiert zu jedem* $w \in (A \cup O)^*$ *höchstens ein* $b \in \mathrm{ABaum}(G)$ *mit* $w = \mathrm{Inp}(b)$.

Beweis. Wir beweisen die Behauptung durch Induktion über $|w|$. Für $|w| = 1$ ist die Behauptung offensichtlich. Sei $|w| > 1$ und $b, b' \in \mathrm{ABaum}(G)$ mit

$$\mathrm{Inp}(b) = \mathrm{Inp}(b') = w$$

Dann gilt

$$b = (o,\, b_1b_2) \qquad b' = (o',\, b_1'b_2')$$

Mit Proposition 19 folgt $\mathrm{Pr}(o) = \mathrm{Pr}(o')$ und $\mathrm{As}(o) = \mathrm{As}(o')$. Ohne Beschränkung der Allgemeinheit sei $\mathrm{As}(o) = \mathrm{As}(o') = \mathtt{l}$. Da

$$\mathrm{Inp}(b_1) \cdot (o, \varepsilon) \cdot \mathrm{Inp}(b_2) = \mathrm{Inp}(b_1') \cdot (o', \varepsilon) \cdot \mathrm{Inp}(b_2')$$

und alle Operatoren in $\mathrm{Inp}(b_2)$ und $\mathrm{Inp}(b_2')$ wegen Proposition 19 echt größere Präzedenz als o und o' haben, folgt $o = o'$, $\mathrm{Inp}(b_2) = \mathrm{Inp}(b_2')$ und $\mathrm{Inp}(b_1) = \mathrm{Inp}(b_1')$. Also gilt $b_1 = b_1'$ und $b_2 = b_2'$ mit der Induktionsvorraussetzung. Folglich $b = b'$. $\square$

Sei $G = (A, O, \mathrm{Pr}, \mathrm{As})$ eine Infix-Grammatik. Wir sagen, dass ein Wort $w \in (A \cup O)^*$ zu einem Ableitungsbaum $b \in \mathrm{ABaum}(G)$ *analysiert* genau dann, wenn $w = \mathrm{Inp}(b)$. Ein Wort $w \in (A \cup O)^*$ heißt *zulässig* genau dann, wenn es zu einem Ableitungsbaum analysiert.

Sei $G = (A, O, \mathrm{Pr}, \mathrm{As})$ eine Infix-Grammatik wie folgt: $A = \mathbf{N}$, $O = \{<, >\}$, $\mathrm{Pr}(<) = \mathrm{Pr}(>) = 0$, $\mathrm{As}(<) = \mathtt{l}$ und $\mathrm{As}(>) = \mathtt{r}$. Dann gilt:

- `1 < 2 < 3` analysiert zu `((1 < 2) < 3)`.
- `1 > 2 > 3` analysiert zu `(1 > (2 > 3))`.
- `1 < 2 > 3` ist unzulässig.
- `1 > 2 < 3` ist unzulässig.

Wir können jetzt unsere konkrete Syntax für SML (siehe Abschnitt 6.9) um Infix-Applikationen I erweitern:

B	$::=$	e	einfacher Ausdruck
	$\mid$	A	Applikation
I	$::=$	$B\ o\ B$	
	$\mid$	$B\ o\ I$	
a	$::=$	$\cdots$	
	$\mid$	I	Infix-Applikation
e	$::=$	$\cdots$	
	$\mid$	`op` o	

Dabei ist o eine Identifikator, der Infix-Status hat. Wir haben nun eine eindeutige Grammatik, die Infix-Applikationen zuläßt. Die Übersetzung von Infix-Applikationen in abstrakte Applikationen erfolgt gemäß der jeweils gültigen Infix-Grammatik. Damit die Grammatik eindeutig bleibt, werden Identifikatoren mit Infix-Status nicht direkt als Ausdrücke zugelassen. Durch Vorschreiben des reservierten Worts `op` kann man aber sagen, dass ein Identifikator mit Infix-Status als Ausdruck analysiert werden soll. Damit können wir Ausdrücke wie

```
op+(4,6+7)
```

schreiben. Das haben wir bei der Konstruktion von MML ausgenützt.