

Zwischenbericht Compilerbau Gruppe 7 Listen

Christof Weibel

Benjamin Neukom

15. November 2013 Version 1.0

Abstract

Dieses Dokument beschreibt, wie die Sprache IML um Listen erweitert wird und welche Änderungen hierfür am Compiler und an der Virtual Machine gemacht werden müssen. Die Spracherweiterung wird mit den Sprachen Scala und Haskell verglichen und die Designentscheide werden erläutert.

1 Listen als Spracherweiterung

Listen für IML sind wie folgt definiert

1. Listen sind geordnete Ansammlungen von Objekten mit dem gleichen Datentyp T . Listen bestehen aus einem Head vom Typ T und einem Tail einer Liste vom Typ T . Beispiel: $[1, 2, 3]$ ist eine Liste vom Typ int mit den Werten 1, 2 und 3 (Head ist 1 und Tail ist $[2, 3]$). Oder mit Verschachtelung $[[true, false], [true]]$ ist eine Liste vom Typ Liste von bools mit den Werten $[true, false]$ und $[true]$.
2. Stores vom Typ Liste können mit Brackets definiert werden. Beispiel: $var l : [int]$ ist ein Store vom Typ Liste von ints.
3. Der Typ einer Liste ist definiert durch das erste Element (falls das Element eine Liste ist, das erste nicht leere Element) der Liste. Beispiel: $[[1, 2, 3], [3, 4, 5]]$ ist eine Liste vom Typ Liste von $ints$ oder $[], [true]$ ist eine Liste vom Typ Liste von bools.
4. Die leere Liste $[]$ ist vom Typ Any (welcher in IML nicht verwendet werden kann). Der Typ Any ist kompatibel mit allen anderen Typen. Beispiel: für die Deklaration des Stores $var l : [[int]]$ ist $l := []$ (Typ von $[]$ ist $[Any]$ und Any ist kompatibel mit $[int]$) so wie auch $l := [[]]$ (Typ von $[[]]$ ist $[[Any]]$ und Any ist kompatibel mit int) eine gültige Zuweisung.
5. Listen sind immutable, d.h. die Werte einer Liste können nicht verändert werden. Es können nur mit den Operationen $::$ (Cons) und $tail$ neue Listen konstruiert werden.

2 Operationen mit Listen

2.1 Cons

$$Type :: [Type] \rightarrow [Type]$$

Der $::$ Operator erstellt eine neue Liste mit dem Operand auf der linken Seite als Head und dem Operand der rechten Seite als Tail. Beispiel: die Expression $1 :: [2, 3, 4]$ gibt die Liste $[1, 2, 3, 4]$ zurück. Der $::$ Operator ist rechtsassoziativ. Beispiel: die Expression $1 :: 2 :: 3 :: []$ gibt die Liste $[1, 2, 3]$ zurück.

2.2 Head

$$head [Type] \rightarrow Type$$

Der $head$ Operator gibt den Head einer Liste zurück. Beispiel: die Expression $head [2, 3, 4]$ gibt den Wert 2 zurück. Falls versucht wird, den Head einer leeren Liste abzufragen, wirft die VM eine Exception.

2.3 Tail

$$\textit{tail} [Type] \rightarrow [Type]$$

Der *tail* Operator gibt den Tail einer Liste zurück. Beispiel: die Expression *tail* [1, 2, 3, 4] gibt die Liste [2,3,4] zurück. Falls versucht wird, den Tail einer leeren Liste abzufragen, wirft die VM eine Exception.

2.4 Length

$$\textit{length} [Type] \rightarrow Int$$

Der *length* Operator gibt die Länge einer Liste zurück. Beispiel: *length* [1, 2, 3, 4, 5] gibt den Wert 5 zurück.

3 List Comprehension

$$\{ \underbrace{3 * x}_{\text{Out Fun}} \mid \underbrace{x}_{\text{Variable}} \underbrace{\text{from 0 to 100}}_{\text{Range}} \text{ when } \underbrace{x \bmod 2 == 0}_{\text{Predicate}} \}$$

Mittels List Comprehension können auf einfache Weise Listen erzeugt werden. Sie folgt der mathematischen Set-Builder Notation.

Output Function

Funktion welche auf die vom Predicate akzeptieren Elemente angewandt wird, bevor sie in die Liste gespeichert werden.

Variable

Die Zähler Variable welche die angegebene Range durchläuft. Die Variable ist immer vom Typ int, deshalb muss dieser nicht explizit angegeben werden. Die Variable kann nur innerhalb dieser List Comprehension gebraucht werden.

Range

Die Inputmenge. Im Beispiel von 0 bis und mit 100.

Predicate

Filter, welcher auf jedes Element der Inputmenge angewendet wird, um zu entscheiden ob es in der resultierenden Liste enthalten ist.

4 Vergleich mit Haskell und Scala

4.1 Haskell

Die Operatoren *head*, *tail* und *length* verhalten sich identisch zu den Haskell Varianten. Einen Unterschied zu Haskell ist die Präzedenz des Cons-Operators. Bei der IML Erweiterung hat der Operator `::` die tiefste Priorität und bei Haskell liegt die Priorität zwischen den booleschen, logischen und den arithmetischen Operatoren. Dies hat folgende Konsequenzen:

Beispiel für Haskell:

```
1 > 3 : True : [] // Type Error
(1 > (3 : (True : []))) // Mit Klammerung um Operator Praezedenz zu zeigen
```

Listing 1: Ungültige Cons Operation in Haskell

Gleiches Beispiel in IML:

```
1 > 3 :: true : [] // Keinen Type Error
((1 > 3) :: (true :: [])) // Mit Klammerung um Operator Praezedenz zu zeigen
```

Listing 2: Gültige Listen Konkatenation in IML

Wir haben uns für diese Präzedenz entschieden, da es unserer Meinung nach natürlicher ist, dass der `::` Operator die tiefste Priorität hat. Ausserdem haben wir noch kein Beispiel gefunden, wo die Operator Präzedenz wie sie in Haskell implementiert ist, einen Vorteil gegenüber unserer Implementation bietet.

4.2 Scala

In Scala verhält sich die Operator Präzedenz des `::` Operator gleich wie bei Haskell, also unterschiedlich zu IML.

5 Änderungen an der Grammatik

Folgende Änderungen wurden an der Grammatik vorgenommen. Die Änderungen wurden mit einem eigens entwickelten Tool (mehr dazu später) getestet und sind *LL(1)* konform.

5.1 Expression

Vorher:

```
expr      ::= term1 {BOOLOPR term1}
term1     ::= term2 [RELOPR term2]
term2     ::= term3 {ADDOPR term3}
term3     ::= factor {MULTOPR factor}
factor    ::= literal
           | IDENT [INIT | exprList]
           | monadicOpr factor
           | LPAREN expr RPAREN;
exprList ::= LPAREN [expr {COMMA expr}] RPAREN
monadicOpr ::= NOT | ADDOPR
```

Nachher:

```
expr      ::= term0 {CONCATOPR term0}
term0     ::= term1 {BOOLOPR term1}
term1     ::= term2 [RELOPR term2]
term2     ::= term3 {ADDOPR term3}
term3     ::= factor {MULTOPR factor}
factor    ::= literal
           | IDENT [INIT | exprList]
           | monadicOpr factor
           | LPAREN expr RPAREN;
exprList ::= LPAREN [expr {COMMA expr}] RPAREN
monadicOpr ::= NOT | ADDOPR | HEAD | TAIL | SIZE
```

5.2 Type

Vorher:

$\text{atomType} ::= \text{INT} \mid \text{BOOL}$

Nachher:

$\text{type} ::= \text{atomType} \mid \text{LBRACKET } \text{type} \text{ RBRACKET}$
 $\text{atomType} ::= \text{INT} \mid \text{BOOL}$

5.3 Literal

Vorher:

$\text{literal} ::= \text{INTLITERAL} \mid \text{BOOLLITERAL}$

Nachher:

$\text{literal} ::= \text{INTLITERAL} \mid \text{BOOLLITERAL} \mid \text{listLiteral}$
 $\text{listLiteral} ::= \text{LBRACKET } \text{expr} \{ \text{COMMA } \text{expr} \} \text{ RBRACKET}$

6 Grammatik auf LL(1) Fehler überprüfen

Die Grammatik wurde, mit einem eigens entwickelten Tool, auf LL(1) Fehler überprüft. Das Tool akzeptiert eine EBNF Grammatik (wie im Unterricht besprochen) und wandelt diese in eine normale Grammatik um. Danach werden die *NULLABLE*, *FIRST* und *FOLLOW* Sets berechnet und damit die Parse Tabelle generiert.

Grammatik der EBNF Grammatik:

```
grammar ::= production {SEMICOLON production};
production ::= NTIDENT ASSIGN term0;
term0 ::= {term1} {PIPE {term1}};
term1 ::= repTerm | optTerm | symbol;
repTerm ::= LCURL term0 RCURL;
optTerm ::= LBRAK term0 RBRAK;
symbol ::= TIDENT | NTIDENT
```

7 Code Generierung

Da wir für die Listen einen Heap benötigen wurde als Zielplattform die Java Virtual Machine gewählt.

8 Code Beispiele

Summe der Elemente einer int Liste:

```
program listSum()
global
fun sum(in copy l:[int]) returns var r:int
do
    if length l == 0 do
        r init := 0
    else
        r init := head l + sum(tail l)
    endif
endfun;

var l:[int];
var sum:int
do
    l init := [1,2,3,4,5,6,7,8,9,10];
    sum init := sum(l);
    debugout sum
endprogram
```

Listing 3: Beispiel für die Berechnung der Summe der Element einer Liste in IML

Contains:

```
program listContains()
global
fun contains(in copy l:[int], in copy i:int) returns var r:bool
do
    if length l == 0 do
        r init := false
    else
        r init := head l == i || contains(tail l, i)
    endif
endfun;

var l:[int];
var i:int;
var c:bool
do
    l init := [1,2,3,4,5,6,7,8];
    debugout l;
    debugin i init;
    c init := contains(l,i);
    debugout c
endprogram
```

Listing 4: Listen Contains

Liste umkehren:

```
program listReverse()
global

// returns last element
fun last(in copy l:[int]) returns var r:int
do
    if length l == 1 do
        r init := head l
    else
        r init := last(tail l)
    endif
endfun;

// init for haskell (list without last)
fun initial(in copy l:[int]) returns var r:[int]
do
    if length l == 1 do
        r init := []
    else
        r init := head l :: initial(tail l)
    endif
endfun;

// would be easier with ++ operator
// reverses the given list and returns a new one
fun reverse(in copy l:[int]) returns var r:[int]
do
    if length l == 0 do
        r init := []
    else
        r init := last(l) :: reverse(initial(l))
    endif
endfun;

var l:[int];
var r:[int]
do
    l init := [1,2,3,4,5,6,7,8];
    r init := reverse(l)
endprogram
```

Listing 5: Liste reverse 1

```

program listReverse()
global

// reverse 2
fun reverse(in copy l:[int], in copy acc:[int]) returns var r:[int]
do
    if length l == 0 do
        r init := acc
    else
        r init := reverse(tail l, head l :: acc)
    endif
endfun;

var l:[int];
var r:[int]
do
    l init := [1,2,3,4,5,6,7,8];
    r init := reverse(l, [])
endprogram

```

Listing 6: Liste reverse 2

```

program primesList()
global
fun isPrime(in copy const p:int) returns var b:bool
global
local
    var c:int
do
    c init := 2;

    if p > 1 do
        b init := true;
        while c < p do
            if p mod c == 0 do
                b := false
            else
                skip
            endif;
            c := c + 1
        endwhile
    else
        b init := false
    endif
endfun;

fun sum(in copy const l:[int]) returns var r:int
local
var x:int
do
    if length l == 0 do
        r init := 0
    else
        r init := head l + sum(tail l)
    endif
endfun;

var l:[int];
var max:int

```

```

do
    debugin max init;
    l init := { x | x from 0 to max when isPrime(x)};
    debugout l;
    debugout sum(l)
endprogram

```

Listing 7: Primzahlen Liste

```

program divisibility()
global
    var l:[[int]];
    var max:int;
    var counter:int
do
    debugin max init;

    l init := [];
    counter init := max;
    while counter > 0 do
        l := { x | x from 1 to max when x mod counter == 0 } :: l;
        counter := counter - 1
    endwhile;

    debugout l
endprogram

```

Listing 8: Teilbarkeit

9 Quellen

Internet:

Wikipedia	en.wikipedia.org
Haskell Listen	http://andres-loeh.de/haskell/4.pdf
Haskell Language Specification	www.haskell.org/onlinereport/
Scala	www.scala.org

Bücher:

Progranmming in Scala Odersky et. al.