

Übung zur Vorlesung Compilerbau, Sprachreferenz für Modula-2

Einleitung

In den Aufgaben werden wir die Sprache Modula-2 (bzw. eine Untermenge davon, die wir der Einfachheit halber ebenfalls „Modula-2“ nennen) betrachten. Modula-2 ist eine 1978 entstandene Weiterentwicklung der Programmiersprache Pascal und wurde wie diese von Niklaus Wirth (ETH Zürich) entwickelt, der 1984 für die Entwicklung von Programmiersprachen mit dem Turing-Award ausgezeichnet wurde. Hauptkennzeichen von Modula-2 sind die Sprachmerkmale zur Modularisierung von Programmen, die sich auch heute in allen modernen Programmiersprachen wiederfinden.

Schreibweise

Im folgenden wird **f**oo für Terminalzeichen, $\langle \text{bar} \rangle$ für Nichtterminalzeichen und BAZ für Zeichencodes verwendet.

Lexik

Schlüsselworte

Modula-2 hat folgende 23 Schlüsselworte:

ARRAY BEGIN CHAR CONST DO ELSE ELSIF END FOR IF INTEGER MODULE
OF PROCEDURE REAL REPEAT RETURN THEN TO TYPE UNTIL VAR WHILE

Für Schlüsselworte wird der Code durch Voranstellen von KEY_ erzeugt, also z.B. KEY_ARRAY .

Operatoren

Modula-2 hat folgende 17 Operatoren, die hier nach Präzedenz sortiert sind:

Zeichen	Code	Beschreibung	Präzedenz	Assoziativität
: =	ASSIGN	Zuweisung		rechts
 oder OR	OR	Boolesche Diskunktion		links
& oder AND	AND	Boolesche Konjunktion		links
=	EQ	Gleichheit / Typedefinition	5	links
<> oder #	NE	Ungleichheit	5	links
<	LE	Kleiner-als	4	links
>	GE	Größer-als	4	links
<=	LEQ	Kleiner-gleich	4	links
>=	GEQ	Größer-gleich	4	links
+	PLUS	Addition	3	links
-	MINUS	Subtraktion	3	links
*	ASTERISK	Multiplikation	2	links
/	SLASH	Division	2	links
DIV	DIV	ganzzahlige Division	2	links
MOD	MOD	Modulo	2	links
~ oder NOT	NOT	Boolesche Negation	1	rechts

Trennzeichen

Modula-2 kennt folgende 11 Trennzeichen:

Zeichen	Code	Beschreibung
;	SEMICOLON	Semikolon
.	PERIOD	Punkt
:	COLON	Doppelpunkt
,	COMMA	Komma
(LPAREN	öffnende runde Klammer
)	RPAREN	schließende runde Klammer
[LBRACK	öffnende eckige Klammer
]	RBRACK	schließende eckige Klammer
..	RANGE	Intervall

Kommentare und Whitespace

Text zwischen den Zeichen (*** und ***) sind Kommentare. Kommentare können geschachtelt werden. Leerzeichen (ASCII-Code 32), Tabulatoren (ASCII-Code 9) und Zeilenumbrüche (ASCII-Codes 10 und 13) sind Whitespace. Kommentare und Whitespace haben keine Auswirkung auf das Programm.

Konstanten

Modula-2 kennt folgende Konstanten:

Muster	Code	Beschreibung
$\langle letter \rangle \{ \langle letter \rangle \mid \langle digit \rangle \}$	IDENT	Bezeichner
$\langle digit \rangle \{ \langle digit \rangle \}$	INTEGER	Ganzzahl
$\langle digit \rangle \{ \langle digit \rangle \} \cdot \{ \langle digit \rangle \}$	REAL	Gleitkommazahl
$' \langle character \rangle '$	CHAR	Zeichen
$" \{ \langle stringelement \rangle \} "$	STRING	Zeichenkette

Dabei steht $\langle digit \rangle$ für die Ziffern 0-9 und $\langle letter \rangle$ für kleine und große Buchstaben. Das Nichtterminal $\langle character \rangle$ enthält alle Zeichen *außer* einfache Anführungszeichen (') oder einen Zeilenumbruch. Das Nichtterminal $\langle stringelement \rangle$ enthält alle Zeichen *außer* dem Anführungszeichen (") oder dem Dateiende.

Syntax

Unter modula2.xhtml ist die Syntax von Modula-2 in erweiterter Backus-Naur-Form (EBNF) gegeben:

- Das Zeichen **::=** trennt die linke von der rechten Seite der Regel.
- Das Zeichen **|** steht für Alternativen.
- Symbole mit Stern ***** können beliebig oft wiederholt (und auch ausgelassen) werden.
- Symbole mit Fragezeichen **?** sind optional.
- Klammern **()** werden zur Gruppierung genutzt.
- Terminalsymbole (siehe Lexik) werden in Anführungszeichen **' '** geschrieben.

Zusätzlich sind zu den Regeln äquivalente Syntaxdiagramme angegeben.