

---

ALBERT LUDWIGS UNIVERSITÄT FREIBURG

TECHNISCHE FAKULTÄT

# PicoC-Compiler

## Übersetzung einer Untermenge von C in den Befehlssatz der RETI-CPU

BACHELORARBEIT

*Abgabedatum:* 28<sup>th</sup> April 2022

*Author:*  
Jürgen Mattheis

*Gutachter:*  
Prof. Dr. Scholl

*Betreuung:*  
M.Sc. Seufert

---

Eine Bachelorarbeit am Lehrstuhl für  
Betriebssysteme

---

---

---

## **ERKLÄRUNG**

Hiermit erkläre ich, dass ich diese Abschlussarbeit selbständig verfasst habe, keine anderen als die angegebenen Quellen/Hilfsmittel verwendet habe und alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen wurden, als solche kenntlich gemacht habe. Darüber hinaus erkläre ich, dass diese Abschlussarbeit nicht, auch nicht auszugsweise, bereits für eine andere Prüfung angefertigt wurde.

---

---

# Inhaltsverzeichnis

<b>1</b>	<b>Motivation</b>	<b>12</b>
1.1	RETI	12
1.2	PicoC	12
1.3	Aufgabenstellung	12
1.4	Eigenheiten der Sprache C	12
1.5	Richtlinien	13
<b>2</b>	<b>Einführung</b>	<b>14</b>
2.1	Compiler und Interpreter	14
2.1.1	T-Diagramme	17
2.2	Formale Sprachen	19
2.2.1	Mehrdeutige Grammatiken	20
2.2.2	Präzidenz und Assoziativität	21
2.3	Lexikalische Analyse	21
2.4	Syntaktische Analyse	24
2.5	Code Generierung	30
2.6	Fehlermeldungen	30
2.6.1	Kategorien von Fehlermeldungen	30
<b>3</b>	<b>Implementierung</b>	<b>31</b>
3.1	Lexikalische Analyse	32
3.1.1	Konkrete Syntax für die Lexikalische Analyse	32
3.1.2	Basic Lexer	33
3.2	Syntaktische Analyse	33
3.2.1	Konkrete Syntax für die Syntaktische Analyse	33
3.2.2	Umsetzung von Präzidenz	34
3.2.3	Derivation Tree Generierung	35
3.2.3.1	Early Parser	35
3.2.3.2	Codebeispiel	35
3.2.4	Derivation Tree Vereinfachung	36
3.2.4.1	Visitor	36
3.2.4.2	Codebeispiel	36
3.2.5	Abstrakt Syntax Tree Generierung	38
3.2.5.1	PicoC-Knoten	38
3.2.5.2	RETI-Knoten	43
3.2.5.3	Kompositionen von PicoC-Knoten und RETI-Knoten mit besonderer Bedeutung	44
3.2.5.4	Abstrakte Syntax	46
3.2.5.5	Transformer	48
3.2.5.6	Codebeispiel	48
3.3	Code Generierung	49
3.3.1	Übersicht	49
3.3.2	Passes	50
3.3.2.1	PicoC-Shrink Pass	50
3.3.2.1.1	Codebeispiel	50
3.3.2.2	PicoC-Blocks Pass	52
3.3.2.2.1	Abstrakte Syntax	52

3.3.2.2.2	Codebeispiel	52
3.3.2.3	PicoC-Mon Pass	54
3.3.2.3.1	Abstrakte Syntax	54
3.3.2.3.2	Codebeispiel	55
3.3.2.4	RETI-Blocks Pass	57
3.3.2.4.1	Abstrakte Syntax	57
3.3.2.4.2	Codebeispiel	57
3.3.2.5	RETI-Patch Pass	60
3.3.2.5.1	Abstrakte Syntax	60
3.3.2.5.2	Codebeispiel	60
3.3.2.6	RETI Pass	62
3.3.2.6.1	Konkrete und Abstrakte Syntax	62
3.3.2.6.2	Codebeispiel	63
3.3.3	Umsetzung von Pointern	66
3.3.3.1	Referenzierung	66
3.3.3.2	Pointer Dereferenzierung durch Zugriff auf Arrayindex ersetzen	68
3.3.4	Umsetzung von Arrays	69
3.3.4.1	Initialisierung von Arrays	69
3.3.4.2	Zugriff auf Arrayindex	71
3.3.4.3	Zuweisung an Arrayindex	74
3.3.5	Umsetzung von Structs	76
3.3.5.1	Deklaration von Structs	76
3.3.5.2	Initialisierung von Structs	77
3.3.5.3	Zugriff auf Structattribut	81
3.3.5.4	Zuweisung an Structattribut	84
3.3.6	Umsetzung der Derived Datatypes im Zusammenspiel	87
3.3.6.1	Einleitungsteil für Globale Statische Daten und Stackframe	87
3.3.6.2	Mittelteil für die verschiedenen Derived Datatypes	90
3.3.6.3	Schlusssteil für die verschiedenen Derived Datatypes	93
3.3.7	Umsetzung von Funktionen	100
3.3.7.1	Funktionen auflösen zu RETI Code	100
3.3.7.1.1	Sprung zur Main Funktion	103
3.3.7.2	Funktionsdeklaration und -definition	105
3.3.7.3	Funktionsaufruf	106
3.3.7.3.1	Ohne Rückgabewert	106
3.3.7.3.2	Mit Rückgabewert	108
3.3.7.3.3	Umsetzung von Call by Sharing für Arrays	111
3.3.7.3.4	Umsetzung von Call by Value für Structs	114
3.3.8	Umsetzung kleinerer Details	116
3.4	Fehlermeldungen	116
3.4.1	Error Handler	116
3.4.2	Arten von Fehlermeldungen	116
3.4.2.1	Syntaxfehler	116
3.4.2.2	Laufzeitfehler	116
<b>4</b>	<b>Ergebnisse und Ausblick</b>	<b>117</b>
4.1	Compiler	117
4.1.1	Überblick über Funktionen	117
4.1.2	Vergleich mit GCC	117
4.1.3	Showmode	117
4.2	Qualitätssicherung	117
4.3	Erweiterungsideen	117
<b>A</b>	<b>Appendix</b>	<b>121</b>

A.1 Konkrete und Abstrakte Syntax . . . . .	121
A.2 Bedienungsanleitungen . . . . .	121
A.2.1 PicoC-Compiler . . . . .	121
A.2.2 Showmode . . . . .	121
A.2.3 Entwicklertools . . . . .	121

---

---

# Abbildungsverzeichnis

2.1	Horizontale Übersetzungszwischenschritte zusammenfassen . . . . .	19
2.2	Vertikale Interpretierungszwischenschritte zusammenfassen . . . . .	19
2.3	Veranschaulichung der Lexikalischen Analyse . . . . .	24
2.4	Veranschaulichung der Syntaktischen Analyse . . . . .	29
3.1	Cross-Compiler Kompiliervorgang ausgeschrieben . . . . .	49
3.2	Cross-Compiler Kompiliervorgang Kurzform . . . . .	49
3.3	Architektur mit allen Passes ausgeschrieben . . . . .	50
4.1	Cross-Compiler als Bootstrap Compiler . . . . .	118
4.2	Iteratives Bootstrapping . . . . .	120

---

---

# Codeverzeichnis

3.1	PicoC Code für Derivation Tree Generierung . . . . .	35
3.2	Derivation Tree nach Derivation Tree Generierung . . . . .	36
3.3	Derivation Tree nach Derivation Tree Vereinfachung . . . . .	37
3.4	Abstract Syntax Tree aus vereinfachtem Derivation Tree generiert . . . . .	48
3.5	PicoC Code für Codebeispiel . . . . .	50
3.6	Abstract Syntax Tree für Codebeispiel . . . . .	51
3.7	PicoC Shrink Pass für Codebeispiel . . . . .	52
3.8	PicoC-Blocks Pass für Codebeispiel . . . . .	54
3.9	PicoC-Mon Pass für Codebeispiel . . . . .	57
3.10	RETI-Blocks Pass für Codebeispiel . . . . .	59
3.11	RETI-Patch Pass für Codebeispiel . . . . .	62
3.12	RETI Pass für Codebeispiel . . . . .	65
3.13	PicoC Code für Pointer Referenzierung . . . . .	66
3.14	Abstract Syntax Tree für Pointer Referenzierung . . . . .	66
3.15	PicoC Mon Pass für Pointer Referenzierung . . . . .	67
3.16	RETI Blocks Pass für Pointer Referenzierung . . . . .	68
3.17	PicoC Code für Pointer Dereferenzierung . . . . .	68
3.18	Abstract Syntax Tree für Pointer Dereferenzierung . . . . .	68
3.19	PicoC Shrink Pass für Pointer Dereferenzierung . . . . .	69
3.20	PicoC Code für Array Initialisierung . . . . .	69
3.21	Abstract Syntax Tree für Array Initialisierung . . . . .	70
3.22	Symboltabelle für Array Initialisierung . . . . .	70
3.23	PicoC Mon Pass für Array Initialisierung . . . . .	71
3.24	RETI Blocks Pass für Array Initialisierung . . . . .	71
3.25	PicoC Code für Zugriff auf Arrayindex . . . . .	72
3.26	Abstract Syntax Tree für Zugriff auf Arrayindex . . . . .	72
3.27	PicoC Mon Pass für Zugriff auf Arrayindex . . . . .	73
3.28	RETI Blocks Pass für Zugriff auf Arrayindex . . . . .	74
3.29	PicoC Code für Zuweisung an Arrayindex . . . . .	74
3.30	Abstract Syntax Tree für Zuweisung an Arrayindex . . . . .	75
3.31	PicoC Mon Pass für Zuweisung an Arrayindex . . . . .	75
3.32	RETI Blocks Pass für Zuweisung an Arrayindex . . . . .	76
3.33	PicoC Code für Deklaration von Structs . . . . .	76
3.34	Symboltabelle für Deklaration von Structs . . . . .	77
3.35	PicoC Code für Initialisierung von Structs . . . . .	78
3.36	Abstract Syntax Tree für Initialisierung von Structs . . . . .	79
3.37	Symboltabelle für Initialisierung von Structs . . . . .	80
3.38	PicoC Mon Pass für Initialisierung von Structs . . . . .	81
3.39	RETI Blocks Pass für Initialisierung von Structs . . . . .	81
3.40	PicoC Code für Zugriff auf Structattribut . . . . .	81
3.41	Abstract Syntax Tree für Zugriff auf Structattribut . . . . .	82
3.42	PicoC Mon Pass für Zugriff auf Structattribut . . . . .	83
3.43	RETI Blocks Pass für Zugriff auf Structattribut . . . . .	84
3.44	PicoC Code für Zuweisung an Structattribut . . . . .	84
3.45	Abstract Syntax Tree für Zuweisung an Structattribut . . . . .	85
3.46	PicoC Mon Pass für Zuweisung an Structattribut . . . . .	86
3.47	RETI Blocks Pass für Zuweisung an Structattribut . . . . .	86

3.48 PicoC Code für den Einleitungsteil . . . . .	87
3.49 Abstract Syntax Tree für den Einleitungsteil . . . . .	88
3.50 PicoC Mon Pass für den Einleitungsteil . . . . .	89
3.51 RETI Blocks Pass für den Einleitungsteil . . . . .	89
3.52 PicoC Code für den Mittelteil . . . . .	90
3.53 Abstract Syntax Tree für den Mittelteil . . . . .	91
3.54 PicoC Mon Pass für den Mittelteil . . . . .	92
3.55 RETI Blocks Pass für den Mittelteil . . . . .	93
3.56 PicoC Code für den Schlussteil . . . . .	94
3.57 Abstract Syntax Tree für den Schlussteil . . . . .	95
3.58 PicoC Mon Pass für den Schlussteil . . . . .	97
3.59 RETI Blocks Pass für den Schlussteil . . . . .	100
3.60 PicoC Code für 3 Funktionen . . . . .	100
3.61 Abstract Syntax Tree für 3 Funktionen . . . . .	101
3.62 PicoC Blocks Pass für 3 Funktionen . . . . .	101
3.63 PicoC Mon Pass für 3 Funktionen . . . . .	102
3.64 RETI Blocks Pass für 3 Funktionen . . . . .	102
3.65 PicoC Code für Funktionen, wobei die main Funktion nicht die erste Funktion ist . . . . .	103
3.66 PicoC Mon Pass für Funktionen, wobei die main Funktion nicht die erste Funktion ist . . . . .	103
3.67 PicoC Blocks Pass für Funktionen, wobei die main Funktion nicht die erste Funktion ist . . . . .	104
3.68 PicoC Patch Pass für Funktionen, wobei die main Funktion nicht die erste Funktion ist . . . . .	105
3.69 PicoC Code für Funktionen, wobei eine Funktion vorher deklariert werden muss . . . . .	105
3.70 Symboltabelle für Funktionen, wobei eine Funktion vorher deklariert werden muss . . . . .	106
3.71 PicoC Code für Funktionsaufruf ohne Rückgabewert . . . . .	106
3.72 PicoC Mon Pass für Funktionsaufruf ohne Rückgabewert . . . . .	107
3.73 RETI Blocks Pass für Funktionsaufruf ohne Rückgabewert . . . . .	108
3.74 RETI Pass für Funktionsaufruf ohne Rückgabewert . . . . .	108
3.75 PicoC Code für Funktionsaufruf mit Rückgabewert . . . . .	109
3.76 PicoC Mon Pass für Funktionsaufruf mit Rückgabewert . . . . .	109
3.77 RETI Blocks Pass für Funktionsaufruf mit Rückgabewert . . . . .	110
3.78 RETI Pass für Funktionsaufruf mit Rückgabewert . . . . .	111
3.79 PicoC Code für Call by Sharing für Arrays . . . . .	111
3.80 PicoC Mon Pass für Call by Sharing für Arrays . . . . .	112
3.81 Symboltabelle für Call by Sharing für Arrays . . . . .	113
3.82 RETI Block Pass für Call by Sharing für Arrays . . . . .	114
3.83 PicoC Code für Call by Value für Structs . . . . .	114
3.84 PicoC Mon Pass für Call by Value für Structs . . . . .	115
3.85 RETI Block Pass für Call by Value für Structs . . . . .	116



---

---

# Tabellenverzeichnis

3.1	Präzidenzregeln von PicoC . . . . .	34
3.2	PicoC-Knoten Teil 1 . . . . .	38
3.3	PicoC-Knoten Teil 2 . . . . .	39
3.4	PicoC-Knoten Teil 3 . . . . .	40
3.5	PicoC-Knoten Teil 4 . . . . .	41
3.6	RETI-Knoten . . . . .	43
3.7	Kompositionen von PicoC-Knoten und RETI-Knoten mit besonderer Bedeutung . . . . .	45

---

---

# Definitionsverzeichnis

1.1	Caller-save Register	12
1.2	Callee-save Register	12
1.3	Deklaration	12
1.4	Definition	12
1.5	Initialisierung	12
1.6	Call by value	13
1.7	Call by reference	13
2.1	Interpreter	14
2.2	Compiler	14
2.3	Maschinensprache	15
2.4	Assemblersprache (bzw. engl. Assembly Language)	15
2.5	Assembler	16
2.6	Objectcode	16
2.7	Linker	16
2.8	Immediate	16
2.9	Transpiler (bzw. Source-to-source Compiler)	17
2.10	Cross-Compiler	17
2.11	T-Diagram Programm	17
2.12	T-Diagram Übersetzer (bzw. eng. Translator)	18
2.13	T-Diagram Interpreter	18
2.14	T-Diagram Maschine	18
2.15	Sprache	19
2.16	Chomsky Hierarchie	19
2.17	Grammatik	20
2.18	Reguläre Sprachen	20
2.19	Kontextfreie Sprachen	20
2.20	Ableitung	20
2.21	Links- und Rechtsableitung	20
2.22	Linksrekursive Grammatiken	20
2.23	Ableitungsbaum	20
2.24	Mehrdeutige Grammatik	21
2.25	Assoziativität	21
2.26	Präzidenz	21
2.27	Wortproblem	21
2.28	LL(k)-Grammatik	21
2.29	Pattern	22
2.30	Lexeme	22
2.31	Lexer (bzw. Scanner oder auch Tokenizer)	22
2.32	Bezeichner (bzw. Identifier)	23
2.33	Literal	24
2.34	Konkrete Syntax	25
2.35	Derivation Tree (bzw. Parse Tree)	25
2.36	Parser	25
2.37	Recognizer (bzw. Erkennen)	26
2.38	Transformer	27
2.39	Visitor	27
2.40	Abstrakte Syntax	28

2.41	Abstract Syntax Tree . . . . .	28
2.42	Pass . . . . .	30
2.43	Monadische Normalform . . . . .	30
2.44	Fehlermeldung . . . . .	30
3.1	Token-Knoten . . . . .	42
3.2	Container-Knoten . . . . .	42
3.3	Symboltabelle . . . . .	54
4.1	Self-compiling Compiler . . . . .	117
4.2	Minimaler Compiler . . . . .	118
4.3	Bootstrap Compiler . . . . .	119
4.4	Bootstrapping . . . . .	119

---

---

# Grammatikverzeichnis

3.1.1 Konkrete Syntax für die Lexikalische Analyse in EBNF . . . . .	32
3.2.1 Konkrete Syntax Syntaktische Analyse in EBNF, Teil 1 . . . . .	33
3.2.2 Konkrete Syntax für die Syntaktische Analyse in EBNF, Teil 2 . . . . .	34
3.2.3 Abstrakte Syntax für $L_{PioC}$ . . . . .	47
3.3.1 Abstrakte Syntax für $L_{PicoC\_Blocks}$ . . . . .	52
3.3.2 Abstrakte Syntax für $L_{PicoC\_Mon}$ . . . . .	54
3.3.3 Abstrakte Syntax für $L_{RETI\_Blocks}$ . . . . .	57
3.3.4 Abstrakte Syntax für $L_{RETI\_Patch}$ . . . . .	60
3.3.5 Konkrete Syntax für $L_{RETI\_Lex}$ . . . . .	62
3.3.6 Konkrete Syntax für $L_{RETI\_Parse}$ . . . . .	63
3.3.7 Abstrakte Syntax für $L_{RETI}$ . . . . .	63

---

---

# 1 Motivation

## 1.1 RETI

... basiert auf ... der Vorlesung C. Scholl, „Betriebssysteme“.

### Definition 1.1: Caller-save Register

*a*

<sup>a</sup>G. Siek, *Course Webpage for Compilers (P423, P523, E313, and E513)*.

### Definition 1.2: Callee-save Register

*a*

<sup>a</sup>G. Siek, *Course Webpage for Compilers (P423, P523, E313, and E513)*.

## 1.2 PicoC

## 1.3 Aufgabenstellung

## 1.4 Eigenheiten der Sprache C

### Definition 1.3: Deklaration

*a*

<sup>a</sup>P. Scholl, „Einführung in Embedded Systems“.

### Definition 1.4: Definition

*a*

<sup>a</sup>P. Scholl, „Einführung in Embedded Systems“.

### Definition 1.5: Initialisierung

*a*

<sup>a</sup>Thiemann, „Einführung in die Programmierung“.

**Definition 1.6: Call by value** $a$ 

---

<sup>a</sup>Bast, „Programmieren in C“.**Definition 1.7: Call by reference** $a$ 

---

<sup>a</sup>Bast, „Programmieren in C“.

## 1.5 Richtlinien

# 2 Einführung

## 2.1 Compiler und Interpreter

Der wohl wichtigsten zu klärenden Begriffe, sind die eines **Compilers** (Definition 2.2) und eines **Interpreters** (Definition 2.1), da das Schreiben eines Compilers von der **PicoC-Sprache**  $L_{PicoC}$  in die **RETI-Sprache**  $L_{RETI}$  das Thema dieser Bachelorarbeit ist und die Definition eines **Interpreters** genutzt wird, um zu definieren was ein **Compiler** ist. Des Weiteren wurde zur **Qualitätsicherung** ein **RETI-Interpreter** implementiert, um mithilfe des **GCC**<sup>1</sup> und von **Tests** die **Beziehungen** in 2.2.1 zu belegen (siehe Subkapitel 4.2).

### Definition 2.1: Interpreter

*Interpretiert die **Instructions** bzw. **Statements** eines Programmes  $P$  direkt.*

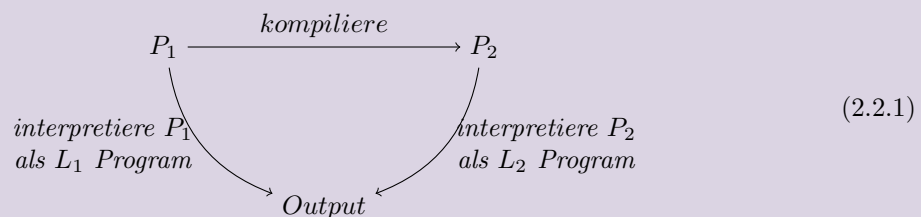
*Auf die Implementierung bezogen arbeitet ein Interpreter auf den compilerinternen **Sub-Bäumen** des **Abstract Syntax Tree** (Definition 2.41) und führt je nach Komposition der **Nodes** des Abstract Syntax Tree, auf die er während des Darüber-Iterierens stösst unterschiedliche Anweisungen aus.<sup>a</sup>*

<sup>a</sup>G. Siek, *Course Webpage for Compilers (P423, P523, E313, and E513)*.

### Definition 2.2: Compiler

***Kompiliert** ein Program  $P_1$ , welches in einer Sprache  $L_1$  geschrieben ist, in ein Program  $P_2$ , welches in einer Sprache  $L_2$  geschrieben ist.*

*Wobei **Kompilieren** meint, dass das Program  $P_1$  in das Program  $P_2$  so übersetzt wird, dass bei beiden Programmen, wenn sie von **Interpretern** ihrer jeweiligen Sprachen  $L_1$  und  $L_2$  **interpretiert** werden, der gleiche **Output** rauskommt. Also beide Programme  $P_1$  und  $P_2$  die gleiche **Semantik** haben und sich nur **syntaktisch** durch die Sprachen  $L_1$  und  $L_2$ , in denen sie geschrieben stehen unterscheiden.<sup>a</sup>*



<sup>a</sup>G. Siek, *Course Webpage for Compilers (P423, P523, E313, and E513)*.

<sup>1</sup>Sammlung von Compilern für Linux bzw. GNU-Linux, steht für **GNU Compiler Collection**

Im Folgenden wird ein voll ausgeschriebenener **Compiler** als  $C_{i.w.k.min}^{o-j}$  geschrieben, wobei  $C_w$  die **Sprache** bezeichnet, die der Compiler als **Input** nimmt und zu einer nicht näher spezifizierten Maschinensprache  $L_{B_i}$  einer Maschine  $M_i$  kompiliert. Fall die Notwendigkeit besteht die **Maschine**  $M_i$  anzugeben, zu dessen **Maschinensprache**  $L_{B_i}$  der Compiler kompiliert, wird das als  $C_i$  geschrieben. Falls die Notwendigkeit besteht die **Sprache**  $L_o$  anzugeben, in der der Compiler selbst geschrieben ist, wird das als  $C^o$  geschrieben. Falls die Notwendigkeit besteht die Version der Sprache, in die der Compiler kompiliert ( $L_{w.k}$ ) oder in der er selbst geschrieben ist ( $L_{o.j}$ ) anzugeben, wird das als  $C_{w.k}^{o-j}$  geschrieben. Falls es sich um einen **minimalen Compiler** handelt (Definition 4.2) kann man das als  $C_{min}$  schreiben.

Üblicherweise kompiliert ein **Compiler** ein **Program**, dass in einer **Programmiersprache** geschrieben ist zu **Maschinenenncode**, der in **Maschinensprache** (Definition 2.3) geschrieben ist, aber es gibt z.B. auch **Transpiler** (Definition 2.9) oder **Cross-Compiler** (Definition 2.10). Des Weiteren sind **Maschinensprache** und **Assemblersprache** (Definition 2.4) voneinander zu unterscheiden.

### Definition 2.3: Maschinensprache

*Programmiersprache, deren mögliche Programme die **hardwarenaheste Repräsentation** eines möglicherweise zuvor hierzu kompilierten bzw. assemblierten Programmes darstellen. Jeder Maschinenbefehl entspricht einer bestimmten **Aufgabe**, die die CPU im **vereinfachten Fall** in einem **Zyklus** der **Fetch- und Execute-Phase**, genauer gesagt in der **Execute-Phase** übernehmen kann oder allgemein in einer **geringen konstanten** Anzahl von Fetch- und Execute Phasen im **komplexeren Fall**. Die Maschinenbefehle sind meist so designed, dass sie sich innerhalb bestimmter **Wortbreiten**, die 2er Potenzen sind codieren lassen. Im einfachsten Fall innerhalb einer **Speicherzelle** des **Hauptspeichers**.<sup>a,b</sup>*

<sup>a</sup>Viele Prozessorarchitekturen erlauben es allerdings auch z.B. **zwei** Maschinenbefehle in **eine** Speicherzelle des Hauptspeichers zu komprimieren, wenn diese zwei Maschinenbefehle keine Operanden mit zu großen **Immediates** (Definition 2.8) haben.

<sup>b</sup>C. Scholl, „Betriebssysteme“.

### Definition 2.4: Assemblersprache (bzw. engl. Assembly Language)

*Eine sehr **hardwarenahe** Programmiersprache, deren **Instructions** eine starke Entsprechung zu bestimmten Maschinenbefehlen bzw. Folgen von Maschinenbefehlen<sup>a</sup> haben. Viele **Instructions** haben eine ähnliche übliche Struktur **Operation** <Operanden>, mit einer **Operation**, die einem **Opcode** eines Maschinenbefehls bezeichnet und keinen oder mehreren **Operanden**, wie die späteren Maschinenbefehle, denen sie entsprechen. Allerdings gibt es oftmals noch viel „syntaktischen Zucker“ innerhalb<sup>b</sup> der Instructions und drumherum<sup>c,d</sup>.*

<sup>a</sup>Instructions der Assemblersprache, die mehreren Maschinenbefehlen entsprechen werden auch als **Pseudo-Instructions** bezeichnet und entsprechen dem, was man im allgemeinen als Macro bezeichnet.

<sup>b</sup>Z.B. erlaubt die Assemblersprache des **GCC** für die **X86\_64-Architektur** für manche Operanden die Syntax **n(%r)**, die einen **Speicherzugriff** mit **Offset**  $n$  zur Adresse, die im **Register** **%r** steht durchführt, wobei z.B. die Klammern **()** usw. nur „syntaktischer Zucker“ sind und natürlich nicht mitcodiert werden.

<sup>c</sup>Z.B. sind im X86\_64 Assembler die Instructions in **Blöcken** untergebracht, die ein **Label** haben und zu denen mittels **jmp <label>** gesprungen werden kann. Ein solches Konstrukt, was vor allem auch noch relativ beliebig wählbare Bezeichner verwendet hat keine direkte Entsprechung in einem handelsüblichen Prozessor und Hauptspeicher.

<sup>d</sup>P. Scholl, „Einführung in Embedded Systems“.

Ein **Assembler** (Definition 2.5) ist in üblichen Compilern in einer bestimmten Form meist schon integriert sein, da Compiler üblicherweise direkt **Maschinenenncode** bzw. **Objectcode** (Definition 2.6) erzeugen. Ein **Compiler** soll möglichst viel von seiner internen Funktionsweise und der damit verbundenen Theorie für den Benutzer abstrahieren und dem Benutzer daher standardmäßig einfach nur den Output liefern, den er in den allermeisten Fällen haben will, nämlich den **Maschinenenncode** bzw. **Objectcode**, der direkt ausführbar ist bzw. wenn er später mit dem **Linker** (Definition 2.7) zu Maschiendencode zusammengesetzt wird ausführbar



ist.

### Definition 2.5: Assembler

Übersetzt im allgemeinen **Assemblercode**, der in **Assemblersprache** geschrieben ist zu **Maschinencode** bzw. **Objectcode** in **binärer Repräsentation**, der in **Maschiensprache** geschrieben ist.<sup>a</sup>

<sup>a</sup>P. Scholl, „Einführung in Embedded Systems“.

### Definition 2.6: Objectcode

Bei komplexeren Compilern, die es erlauben den Programmcode in **mehrere Dateien** aufzuteilen wird häufig **Objectcode** erzeugt, der neben der Folge von Maschinenbefehlen in **binärer Repräsentation** auch noch Informationen für den **Linker** enthält, die im späteren **Maschiencode** nicht mehr enthalten sind, sobald der **Linker** die Objektdateien zum Maschinencode zusammengesetzt hat.<sup>a</sup>

<sup>a</sup>P. Scholl, „Einführung in Embedded Systems“.

### Definition 2.7: Linker

Programm, das **Objektcode** aus mehreren Objektdateien zu ausführbarem **Maschinencode** in eine ausführbare Datei oder Bibliotheksdatei **linkt**, sodass unter anderem kein vermeidbarer **doppelter Code** darin vorkommt.<sup>a</sup>

<sup>a</sup>P. Scholl, „Einführung in Embedded Systems“.

Der **Maschinencode**, denn ein üblicher Compiler einer Programmiersprache generiert, enthält seine Folge von Maschinenbefehlen üblicherweise in **binärer Repräsentation**, da diese in erster Linie für die Maschine, die binär arbeitet verständlich sein sollen und nicht für den Programmierer.

Der **PicoC-Compiler**, der den Zweck erfüllt für Studenten ein **Anschauungs- und Lernwerkzeug** zu sein, generiert allerdings Maschinencode, der die Maschinenbefehle bzw. RETI-Befehle in **menschenlesbarer Form** mit ausgeschriebenem RETI-Operationen, RETI-Registern und Immediates (Definition 2.8) enthält. Für den **RETI-Interpreter** ist es ebenfalls nicht notwendig, dass der Maschinencode, denn der PicoC-Compiler generiert in binärer Darstellung ist, denn es ist für den RETI-Interpreter ebenfalls leichter diese einfach direkt in menschenlesbarer Form zu interpretieren, da der RETI-Interpreter nur die sichtbare Funktionsweise einer RETI-CPU **simulieren** soll und nicht deren mögliche interne Umsetzung<sup>2</sup>.

### Definition 2.8: Immediate

**Konstanter Wert**, der als **Teil eines Maschinenbefehls** gespeichert ist und dessen **Wertebereich** dementsprechend auch durch die die Anzahl an Bits, die ihm innerhalb dieses **Maschinenbefehls** zur Verfügung gestellt sind, **beschränkter** ist als bei sonstigen Werten innerhalb des Hauptspeichers, denen eine ganze Speicherzelle des Hauptspeichers zur Verfügung steht.<sup>a</sup>

<sup>a</sup>Ljohhuh, *What is an immediate value?*

<sup>2</sup>Eine **RETI-CPU** zu bauen, die menschenlesbaren Maschinencode in z.B. **UTF-8 Codierung** ausführen kann, wäre dagegen unnötig kompliziert und aufwändig, da Hardware **binär** arbeitet und man dieser daher lieber direkt die binär codierten Maschinenbefehle übergibt, anstatt z.B. eine unnötig **platzverbrauchenden** UTF-8 Codierung zu verwenden, die nur in sehr vielen Schritten einen Befehl verarbeiten kann, da die Register und Speicherzellen des Hauptspeichers üblicherweise nur **32- bzw. 64-Bit Breite** haben.

**Definition 2.9: Transpiler (bzw. Source-to-source Compiler)**

Kompiliert zwischen Sprachen, die ungefähr auf dem *gleichen* Level an *Abstraktion* arbeiten<sup>a,b</sup>

<sup>a</sup>Die Programmiersprache **TypeScript** will als **Obermenge** von **JavaScript** die Sprachhe Javascript **erweitern** und gleichzeitig die **syntaktischen Mittel** von JavaScript unterstützen. Daher bietet es sich Typescript zu Javascript zu **transpilieren**.

<sup>b</sup>Thiemann, „Compilerbau“.

**Definition 2.10: Cross-Compiler**

Kompiliert auf einer **Maschine**  $M_1$  ein Program, dass in einer **Sprache**  $L_w$  geschrieben ist für eine **andere Maschine**  $M_2$ , wobei beide Maschinen  $M_1$  und  $M_2$  unterschiedliche **Maschinen Sprachen**  $B_1$  und  $B_2$  haben.<sup>a,b</sup>

<sup>a</sup>Beim **PicoC-Compiler** handelt es sich um einen **Cross-Compiler**  $C_{PicoC}^{Python}$ .

<sup>b</sup>Earley und Sturgis, „A formalism for translator interactions“.

Ein **Cross-Compiler** ist entweder notwendig, wenn eine Zielmaschine  $M_2$  nicht ausreichend **Rechenleistung** hat, um ein Programm in der Wunschsprache  $L_w$  selbst **zeitnah** zu kompilieren oder wenn noch kein Compiler  $C_w$  für die **Wunschsprache**  $L_w$  und andere Programmiersprachen  $L_o$ , in denen man Programmieren wollen würde existiert, der unter der **Maschinen Sprache**  $B_2$  einer Zielmaschine  $M_2$  läuft.<sup>3</sup>

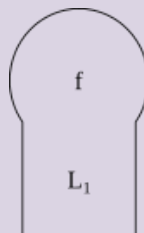
**2.1.1 T-Diagramme**

Um die Architektur von **Compilern** und **Interpretern** übersichtlich darzustellen eignen sich **T-Diagramme**, deren Spezifikation aus dem Paper Earley und Sturgis, „A formalism for translator interactions“ entnommen ist besonders gut, da diese optimal darauf **zugeschnitten** sind die Eigenheiten von Compilern in ihrer Art der Darstellung unterzubringen.

Die **Notation** setzt sich dabei aus den **Blöcken** für ein Programm (Definition 2.11), einen Übersetzer (Definition 2.12), einen Interpreter (Definition 2.13) und eine Maschine (Definition 2.14) zusammen.

**Definition 2.11: T-Diagram Programm**

Repräsentiert ein **Programm**, dass in der **Sprache**  $L_1$  geschrieben ist und die **Funktion**  $f$  berechnet.<sup>a</sup>



<sup>a</sup>Earley und Sturgis, „A formalism for translator interactions“.

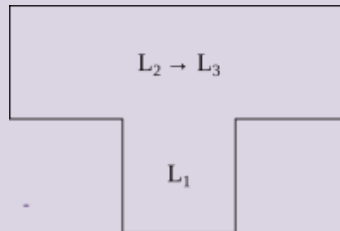
Es ist bei **T-Diagrammen** nicht notwendig beim entsprechenden **Platzhalter**, in den man die genutzte **Sprache** schreibt, den **Namen der Sprache** an ein  $L$  dranzuhängen, weil hier immer eine **Sprache** steht. Es würde in Definition 2.11 also reichen einfach eine 1 hinzuschreiben.

<sup>3</sup>Die an vielen Universitäten und Schulen eingesetzten programmierbaren Roboter von **Lego Mindstorms** nutzen z.B. einen **Cross-Compiler**, um für den programmierbaren Microcontroller eine **C-ähnliche Sprache** in die Maschinent Sprache des Microcontrollers zu kompilieren, da der Microcontroller selbst nicht genug Rechenleistung besitzt, um ein Programm selbst **zeitnah** zu kompilieren.

**Definition 2.12: T-Diagramm Übersetzer (bzw. eng. Translator)**

Repräsentiert einen **Übersetzer**, der in der **Sprache**  $L_1$  geschrieben ist und **Programme** von der **Sprache**  $L_2$  in die **Sprache**  $L_3$  kompiliert.

Für den **Übersetzer** gelten genauso, wie für einen **Compiler**<sup>a</sup> die **Beziehungen** in 2.2.1.<sup>b</sup>

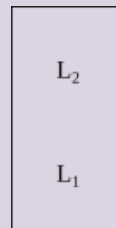


<sup>a</sup>Zwischen den Begriffen **Übersetzung** und **Kompilierung** gibt es einen kleinen Unterschied, **Übersetzung** ist **kleinschrittiger** als **Kompilierung** und ist auch zwischen **Passes** möglich, **Kompilierung** beinhaltet dagegen bereits alle **Passes** in einem Schritt. **Kompilieren** ist also auch **Übersetzen**, aber **Übersetzen** ist nicht immer auch **Kompilieren**.

<sup>b</sup>Earley und Sturgis, „A formalism for translator interactions“.

**Definition 2.13: T-Diagramm Interpreter**

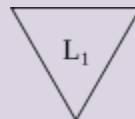
Repräsentiert einen **Interpreter**, der in der **Sprache**  $L_1$  geschrieben ist und **Programme** in der **Sprache**  $L_2$  interpretiert.<sup>a</sup>



<sup>a</sup>Earley und Sturgis, „A formalism for translator interactions“.

**Definition 2.14: T-Diagramm Maschine**

Repräsentiert eine **Maschine**, welche ein **Programm** in **Maschinensprache**  $L_1$  ausführt.<sup>a,b</sup>



<sup>a</sup>Wenn die Maschine **Programme** in einer höheren Sprache als **Maschinensprache** ausführt, ist es auch erlaubt diese Notation zu verwenden, dann handelt es sich um eine **Abstrakte Maschine**, wie z.B. die **Python Virtual Machine** (PVM) oder **Java Virtual Machine** (JVM).

<sup>b</sup>Earley und Sturgis, „A formalism for translator interactions“.

Aus den verschiedenen **Blöcken** lassen sich **Kompositionen** bilden, indem man sie **adjazent** zueinander platziert. Allgemein lässt sich grob sagen, dass **vertikale Adjazents** für **Interpretation** und **horizontale Adjazents** für **Übersetzung** steht.

Sowohl **horizontale** als auch **vertikale Adjazents** lassen sich, wie man in den Abbildungen 2.1 und 2.2 erkennen kann zusammenfassen.

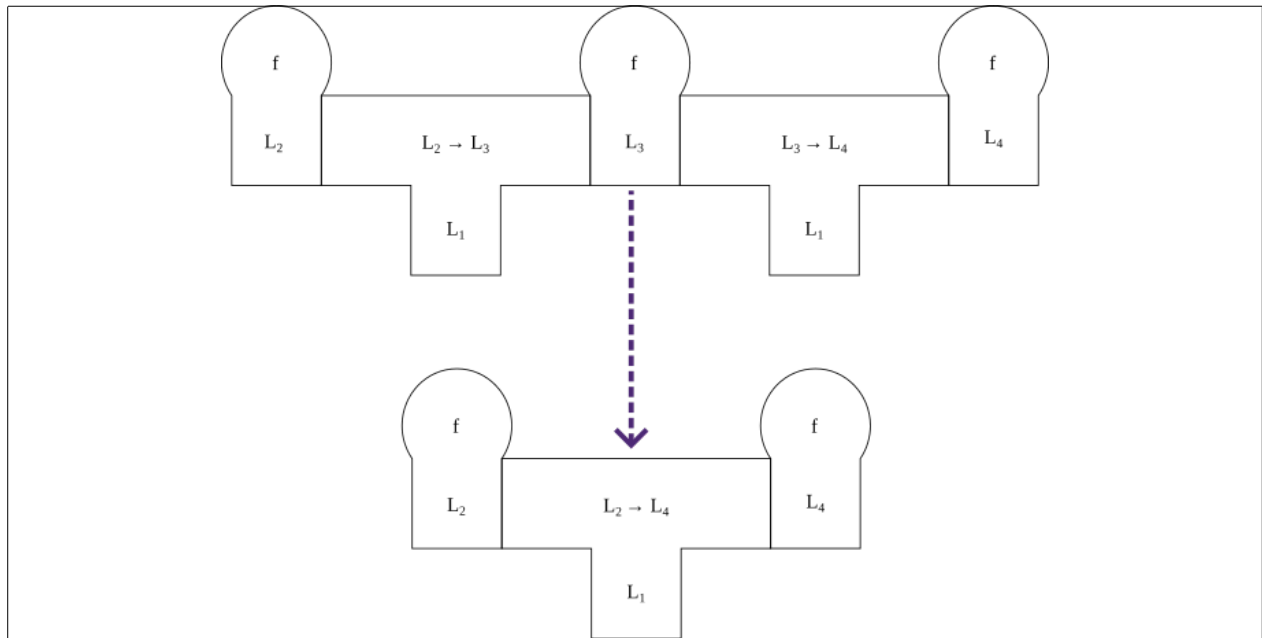


Abbildung 2.1: Horizontale Übersetzungszwischenschritte zusammenfassen

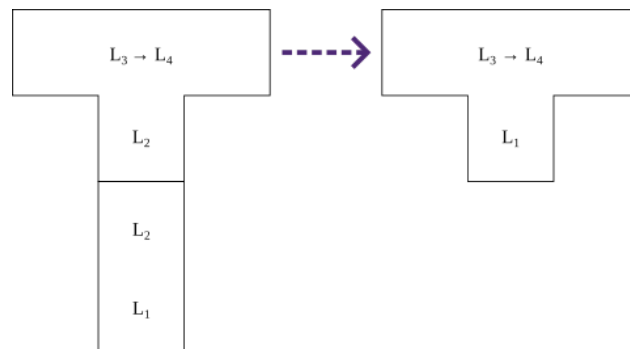


Abbildung 2.2: Vertikale Interpretierungszwischenschritte zusammenfassen

## 2.2 Formale Sprachen

### Definition 2.15: Sprache

*a*

<sup>a</sup>Nebel, „Theoretische Informatik“.

### Definition 2.16: Chomsky Hierarchie

*a*

<sup>a</sup>Nebel, „Theoretische Informatik“.

**Definition 2.17: Grammatik***a*<sup>a</sup>Nebel, „Theoretische Informatik“.**Definition 2.18: Reguläre Sprachen***a*<sup>a</sup>Nebel, „Theoretische Informatik“.**Definition 2.19: Kontextfreie Sprachen***a*<sup>a</sup>Nebel, „Theoretische Informatik“.**Definition 2.20: Ableitung***a*<sup>a</sup>Nebel, „Theoretische Informatik“.**Definition 2.21: Links- und Rechtsableitung***a*<sup>a</sup>Nebel, „Theoretische Informatik“.**Definition 2.22: Linksrekursive Grammatiken**

Eine **Grammatik** ist **linksrekursiv**, wenn sie ein **Nicht-Terminalsymbol** enthält, dass **linksrekursiv** ist.

Ein **Nicht-Terminalsymbol** ist **linksrekursiv**, wenn das **linkeste Symbol** in einer seiner **Produktionen** es selbst ist oder zu sich selbst gemacht werden kann durch eine Folge von Ableitungen:

$$A \Rightarrow^* Aa,$$

wobei *a* eine beliebige Folge von **Terminalsymbolen** und **Nicht-Terminalsymbolen** ist.<sup>a</sup>

<sup>a</sup>Parsing Expressions · Crafting Interpreters.**2.2.1 Mehrdeutige Grammatiken****Definition 2.23: Ableitungsbaum***a*<sup>a</sup>Nebel, „Theoretische Informatik“.

**Definition 2.24: Mehrdeutige Grammatik***a*<sup>a</sup>Nebel, „Theoretische Informatik“.**2.2.2 Präzidenz und Assoziativität****Definition 2.25: Assoziativität***a*<sup>a</sup>*Parsing Expressions · Crafting Interpreters.***Definition 2.26: Präzidenz***a*<sup>a</sup>*Parsing Expressions · Crafting Interpreters.***Definition 2.27: Wortproblem***a*<sup>a</sup>Nebel, „Theoretische Informatik“.**Definition 2.28: LL(k)-Grammatik**

Eine Grammatik ist **LL(k)** für  $k \in \mathbb{N}$ , falls jeder Ableitungsschritt eindeutig durch die nächsten  $k$  **Symbole** des **Eingabeworts** bzw. in Bezug zu Compilerbau **Token** des **Inputstrings** zu bestimmen ist<sup>a</sup>. Dabei steht **LL** für *left-to-right* und *leftmost-derivation*, da das **Eingabewort** von **links nach rechts** geparsed und immer **Linksableitungen** genommen werden müssen<sup>b</sup>, damit die obige Bedingung mit den **nächsten**  $k$  Symbolen gilt.<sup>c</sup>

<sup>a</sup>Das wird auch als **Lookahead** von  $k$  bezeichnet.<sup>b</sup>Wobei sich das mit den **Linksableitungen** automatisch ergibt, wenn man das Eingabewort von **links-nach-rechts** parsed und jeder der nächsten  $k$  **Ableitungsschritte** eindeutig sein soll.<sup>c</sup>Nebel, „Theoretische Informatik“.**2.3 Lexikalische Analyse**

Die **Lexikalische Analyse** bildet üblicherweise die erste Ebene innerhalb der **Pipe Architektur** bei der Implementierung von Compilern. Die Aufgabe der lexikalischen Analyse ist vereinfacht gesagt, in einem Inputstring, z.B. dem Inhalt einer Datei, welche in **UTF-8** codiert ist, Folgen endlicher Symbole (auch **Wörter** genannt) zu finden, die bestimmte **Pattern** (Definition 2.29) matchen, die durch eine **reguläre Grammatik** spezifiziert sind.

Diese Folgen endlicher Symbole werden auch **Lexeme** (Definition 2.30) genannt.

**Definition 2.29: Pattern**

*Beschreibung* aller möglichen **Lexeme**, die eine Menge  $\mathbb{P}_T$  bilden und einem bestimmten **Token**  $T$  zugeordnet werden. Die Menge  $\mathbb{P}_T$  ist eine möglicherweise unendliche Menge von **Wörtern**, die sich mit den Produktionen einer **regulären Grammatik**  $G_{Lex}$  einer **regulären Sprache**  $L_{Lex}$  beschreiben lassen<sup>a</sup>, die für die Beschreibung eines **Tokens**  $T$  zuständig sind.<sup>b</sup>

<sup>a</sup>Als Beschreibungswerkzeug können aber auch z.B. reguläre Ausdrücke hergenommen werden.

<sup>b</sup>Thiemann, „Compilerbau“.

**Definition 2.30: Lexeme**

Ein **Lexeme** ist ein **Wort** aus dem Inputstring, welches das **Pattern** für eines der **Token**  $T$  einer **Sprache**  $L_{Lex}$  matched.<sup>a</sup>

<sup>a</sup>Thiemann, „Compilerbau“.

Diese **Lexeme** werden vom **Lexer** (Definition 2.31) im **Inputstring** identifiziert und **Tokens**  $T$  zugeordnet. Das jeweils nächste **Lexeme** fängt dabei genau nach dem letzten Symbol des **Lexemes** an, das zuletzt vom **Lexer** erkannt wurde. Die **Tokens** (Definition 2.31) sind es, die letztendlich an die **Syntaktische Analyse** weitergegeben werden.

**Definition 2.31: Lexer (bzw. Scanner oder auch Tokenizer)**

Ein **Lexer** ist eine **partielle Funktion**  $lex : \Sigma^* \rightarrow (N \times W)^*$ , welche ein **Wort** bzw. **Lexeme** aus  $\Sigma^*$  auf ein **Token**  $T$  mit einem **Tokennamen**  $N$  und einem **Tokenwert**  $W$  abbildet, falls dieses **Wort** sich unter der **regulären Grammatik**  $G_{Lex}$ , der **regulären Sprache**  $L_{Lex}$  ableiten lässt bzw. einem der **Pattern** der Sprache  $L_{Lex}$  entspricht.<sup>a</sup>

<sup>a</sup>Thiemann, „Compilerbau“.

Ein **Lexer** ist im Allgemeinen eine **partielle Funktion**, da es Zeichenfolgen geben kann, die kein **Pattern** eines **Tokens** der Sprache  $L_{Lex}$  matchen. In Bezug auf eine Implementierung, wird, wenn der Lexer Teil der Implementierung eines Compilers ist, in diesem Fall eine **Fehlermeldung** ausgegeben.

Um Verwirrung verzubäuen ist es wichtig folgende Unterscheidung hervorzuheben:

Wenn von **Symbolen** die Rede ist, so werden in der **Lexikalischen Analyse**, der **Syntaktische Analyse** und der **Code Generierung**, auf diesen verschiedenen Ebenen unterschiedliche Konzepte als Symbole bezeichnet.

In der Lexikalischen Analyse sind einzelne **Zeichen eines Zeichensatzes** die Symbole.

In der Syntaktischen Analyse sind die **Tokennamen** die Symbole.

In der Code Generierung sind die **Bezeichner** (Definition 2.32) von **Variablen, Konstanten und Funktionen** die Symbole<sup>a</sup>.

<sup>a</sup>Das ist der Grund, warum die **Tabelle**, in der Informationen zu **Bezeichnern** gespeichert werden, in Kapitel 3 **Symoltabelle** genannt wird.

**Definition 2.32: Bezeichner (bzw. Identifier)**

*Tokenwert, der eine Konstante, Variable, Funktion usw. **eindeutig** benennt.<sup>a,b</sup>*

<sup>a</sup>Außer wenn z.B. bei Funktionen die Programmiersprache das **Überladen** erlaubt usw. In diesem Fall wird die **Signatur** der Funktion als weiteres Unterscheidungsmerkmal hinzugenommen, damit es eindeutig ist.

<sup>b</sup>Thiemann, „Einführung in die Programmierung“.

Eine weitere Aufgabe der **Lexikalischen Analyse** ist es jegliche für die Weiterverarbeitung unwichtigen Symbole, wie Leerzeichen `␣`, Newline `\n`<sup>4</sup> und Tabs `\t` aus dem Inputstring herauszufiltern. Das geschieht mittels des **Lexers**, der allen für die **Syntaktische Analyse** unwichtige Zeichen das leere Wort  $\epsilon$  zuordnet. Das ist auch im Sinne der Definition, denn  $\epsilon \in (N \times W)^*$  ist immer der Fall beim **Kleene Stern Operator**  $*$ . Nur das, was für die **Syntaktische Analyse** wichtig ist, soll weiterverarbeitet werden, alles andere wird herausgefiltert.

Der Grund warum nicht einfach nur die **Lexeme** an die **Syntaktische Analyse** weitergegeben werden und der Grund für die Aufteilung des **Tokens** in **Tokenname** und **Tokenwert** ist, weil z.B. die Bezeichner von Variablen, Konstanten und Funktionen beliebige Zeichenfolgen sein können, wie `my_fun`, `my_var` oder `my_const` und es auch viele verschiedenen Zahlen gibt, wie 42, 314 oder 12. Die **Überbegriffe** bzw. **Tokennamen** für beliebige Bezeichner von Variablen, Konstanten und Funktionen und beliebige Zahlen sind aber trotz allem z.B. `NAME` und `NUM`<sup>5</sup>, bzw. wenn man sich nicht Kurzformen sucht `IDENTIFIER` und `NUMBER`. Für **Lexeme**, wie `if` oder `}` sind die **Tokennamen** bzw. Überbegriffe genau die Bezeichnungen, die man diesen Zeichenfolgen geben würde, nämlich `IF` und `RBRACE`.

Ein **Lexeme** ist damit aber nicht immer das gleiche, wie der **Tokenwert**, denn z.B. im Falle von PicoC kann der Wert 99 durch zwei verschiedene **Literale** (Definition 2.33) dargestellt werden, einmal als ASCII-Zeichen `'c'`, dass den entsprechenden Wert in der ASCII-Tabelle hat und des Weiteren auch in Dezimalschreibweise als 99<sup>6</sup>. Der **Tokenwert** ist jedoch der letztendlich verwendete Wert an sich, unabhängig von der Darstellungsform.

Die **Grammatik**  $G_{Lex}$ , die zur Beschreibung der Token  $T$  der Sprache  $L_{Lex}$  verwendet wird ist üblicherweise **regulär**, da ein typischer **Lexer** immer nur **ein Symbol** vorausschaut<sup>7</sup>, sich nichts merken muss und unabhängig davon, was für Symbole davor aufgetaucht sind läuft. Die Grammatik 3.1.1 liefert den Beweis, dass die Sprache  $L_{PicoC\_Lex}$  des **PicoC-Compilers** auf jeden Fall **regulär** ist, da sie fast die Definition 2.18 erfüllt. Einzig die Produktion `CHAR ::= "'ASCII_CHAR'"` sieht problematisch aus, kann allerdings auch als `{CHAR ::= "'CHAR2', CHAR2 ::= ASCII_CHAR'"}` **regulär** ausgedrückt werden<sup>8</sup>. Somit existiert eine **reguläre Grammatik**, welche die **Sprache**  $L_{PicoC\_Lex}$  beschreibt und damit ist die **Sprache**  $L_{PicoC\_Lex}$  **regulär**.

<sup>4</sup>In Unix Systemen wird für Newline das ASCII Symbol **line feed**, in Windows hingegen die ASCII Symbole **carriage return** und **line feed** nacheinander verwendet. Das wird aber meist durch die verwendete Programmiersprache, die man zur Implementierung des Lexers nutzt wegabstrahiert.

<sup>5</sup>Diese **Tokennamen** wurden im **PicoC-Compiler** verwendet, da man beim Programmieren möglichst **kurze** und **leicht verständliche** Bezeichner für seine Nodes haben will, damit unter anderem **mehr Code** in eine Zeile passt.

<sup>6</sup>Die Programmiersprache **Python** erlaubt es z.B. dieser Wert auch mit den Literalen `0b1100011` und `0x63` darzustellen.

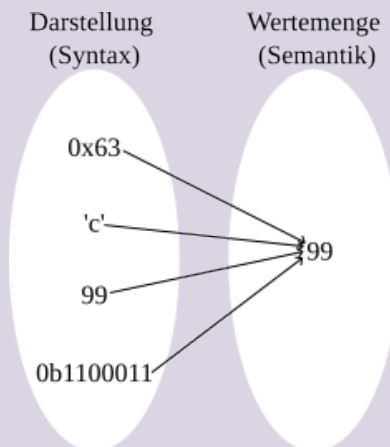
<sup>7</sup>Man nennt das auch einem **Lookahead** von 1

<sup>8</sup>Eine derartige Regel würde nur Probleme bereiten, wenn sich aus `ASCII_CHAR` **beliebig breite** Wörter ableiten lassen.



**Definition 2.33: Literal**

Eine von möglicherweise vielen weiteren *Darstellungsformen* (als *Zeichenkette*) für ein und denselben *Wert* eines *Datentyps*.<sup>a</sup>



<sup>a</sup>Thiemann, „Einführung in die Programmierung“.

Um eine Gesamtübersicht über die **Lexikalische Analyse** zu geben, ist in Abbildung 2.3 die Lexikalische Analyse an einem Beispiel veranschaulicht.

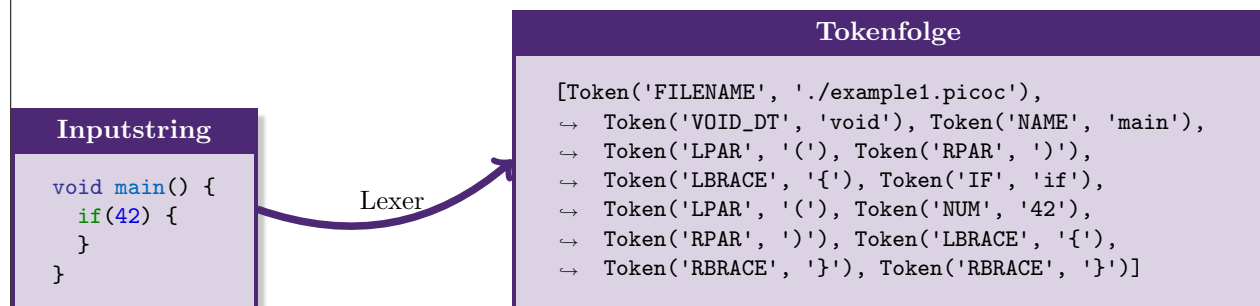


Abbildung 2.3: Veranschaulichung der Lexikalischen Analyse

## 2.4 Syntaktische Analyse

In der **Syntaktischen Analyse** ist für einige Sprachen eine **Kontextfreie Grammatik**  $G_{Parse}$  notwendig, um diese Sprachen zu beschreiben, da viele Programmiersprachen z.B. für **Funktionsaufrufe** `fun(arg)` und **Codeblöcke** `if(1){}` syntaktische Mittel verwenden, die es notwendig machen sich zu merken, wieviele öffnende runde Klammern '(' bzw. öffnende geschweifte Klammern '{' es momentan gibt, die noch nicht durch eine entsprechende schließende runde Klammer ')' bzw. schließende geschweifte Klammer '}' geschlossen wurden.

Die **Syntax**, in welcher der **Inputstring** aufgeschrieben ist, wird auch als **Konkrete Syntax** (Definition 2.34) bezeichnet. In einem Zwischenschritt, dem **Parsen** wird aus diesem Inputstring mithilfe eines **Parsers** (Definition 2.36), ein **Derivation Tree** (Definition 2.35) generiert, der als Zwischenstufe hin zum einem **Abstract Syntax Tree** (Definition 2.41) dient. Beim Compilerbau ist es förderlich kleinschrittig vorzugehen, deshalb erst die Generierung des **Derivation Tree** und dann erst des **Abstract Syntax Tree**.

**Definition 2.34: Konkrete Syntax**

*Syntax* einer *Sprache*, die durch die *Grammatiken*  $G_{Lex}$  und  $G_{Parse}$  zusammengenommen beschrieben wird.

Ein *Programm* in seiner *Textrepräsentation*, wie es in einer Textdatei nach den Produktionen der *Grammatiken*  $G_{Lex}$  und  $G_{Parse}$  abgeleitet steht, bevor man es kompiliert, ist in *Konkreter Syntax* aufgeschrieben.<sup>a</sup>

<sup>a</sup>G. Siek, *Course Webpage for Compilers (P423, P523, E313, and E513)*.

**Definition 2.35: Derivation Tree (bzw. Parse Tree)**

*Compilerinterne Darstellung* eines in *Konkreter Syntax* geschriebenen Inputstrings als *Baumdatenstruktur*, in der *Nichtterminalsymbole* die *Inneren Knoten* der Baumdatenstruktur und *Terminalsymbole* die *Blätter* der Baumdatenstruktur bilden. Jedes zum Ableiten des Inputstrings verwendete *Nicht-Terminalsymbol* einer *Produktion* der *Grammatik*  $G_{Parse}$ , die ein Teil der *Konkrete Syntax* ist, bildet einen eigenen *Inneren Knoten*.

Der *Derivation Tree* wird optimalerweise immer so konstruiert bzw. die *Konkrete Syntax* immer so definiert, dass sich möglichst einfach ein *Abstract Syntax Tree* daraus konstruieren lässt.<sup>a</sup>

<sup>a</sup>JSON parser - Tutorial — Lark documentation.

**Definition 2.36: Parser**

Ein *Parser* ist ein Programm, dass aus einem Inputstring, der in *Konkreter Syntax* geschrieben ist, eine compilerinterne Darstellung, den *Derivation Tree* generiert, was auch als *Parsen* bezeichnet wird.<sup>a, b</sup>

<sup>a</sup>Es gibt allerdings auch alternative Definitionen, denen nach ein Parser in Bezug auf Compilerbau ein Programm ist, dass einen Inputstring von *Konkreter Syntax* in *Abstrakte Syntax* übersetzt. Im Folgenden wird allerdings die Definition 2.36 verwendet.

<sup>b</sup>JSON parser - Tutorial — Lark documentation.

An dieser Stelle könnte möglicherweise eine Verwirrung entstehen, welche Rolle dann überhaupt ein *Lexer* hier spielt.

In Bezug auf Compilerbau ist ein *Lexer* ein Teil eines *Parsers*. Der *Lexer* ist ausschließlich für die *Lexikalische Analyse* verantwortlich und entspricht z.B., wenn man bei einem Wanderausflug verschiedenen Insekten entdeckt, dem Nachschlagen in einem Insektenlexikon und dem Aufschreiben, welchen Insekten man in welcher *Reihenfolge* begegnet ist. Zudem kann man bestimmte *Sehenswürdigkeiten* an denen man während des Ausflugs vorbeikommt ebenfalls festhalten, da es eine Rolle spielen kann in welchem örtlichen *Kontext* man den Insekten begegnet ist.<sup>a</sup>

Der *Parser* vereinigt sowohl die *Lexikalische Analyse*, als auch einen Teil der *Syntaktischen Analyse* in sich und entspricht, um auf das Beispiel zurückzukommen, dem Darstellen von *Beziehungen* zwischen den Insektenbegegnungen in einer für die *Weiterverarbeitung tauglichen Form*.<sup>b</sup>

In der Weiterverarbeitung kann der *Interpreter* das interpretieren und daraus bestimmte Schlüsse ziehen und ein *Compiler* könnte es vielleicht in eine für Menschen leichter entschlüsselbare Sprache kompilieren.

<sup>a</sup>Das würde z.B. der Rolle eines *Semikolon* ; in der Sprache  $L_{PicoC}$  entsprechen.

<sup>b</sup>Z.B. gibt es bestimmte *Wechselbeziehungen* zwischen Insekten, Insekten beeinflussen sich gegenseitig.

Die vom **Lexer** im Inputstring identifizierten **Token** werden in der **Syntaktischen Analyse** vom **Parser** als **Wegweiser** verwendet, da je nachdem, in welcher Reihenfolge die **Token** auftauchen, dies einer anderen Ableitung in der **Grammatik**  $G_{Parse}$  entspricht. Dabei wird in der Grammatik  $L_{Parse}$  nach dem **Tokennamen** unterschieden und nicht nach dem Tokenwert, da es nur von Interesse ist, ob an einer bestimmten Stelle z.B. eine **Zahl** steht und nicht, welchen konkreten Wert diese **Zahl** hat. Der **Tokenwert** ist erst später in der **Code Generierung** in 2.5 wieder relevant.

Ein **Parser** ist genauergesagt ein erweiterter **Recognizer** (Definition 2.37), denn ein Parser löst das **Wortproblem** (Definition 2.27) für die **Sprache**, die durch die **Konkrete Syntax** beschrieben wird und konstruiert parallel dazu oder im Nachgang aus den Informationen, die während der Ausführung des Recognition Algorithmus gesichert wurden den **Derivation Tree**.

#### Definition 2.37: Recognizer (bzw. Erkennen)

*Entspricht dem Maschinenmodell eines **Automaten**. Im Bezug auf Compilerbau entspricht der **Recognizer** einem **Kellerautomaten**, in dem **Wörter** bestimmter **Kontextfreier Sprachen** erkannt werden. Der **Recognizer** erkennt, ob ein Inputstring bzw. **Wort** sich mit den Produktionen der **Konkrete Syntax** ableiten lässt, also ob er bzw. es Teil der Sprache ist, die von der **Konkreten Syntax** beschrieben wird oder nicht<sup>ab</sup>*

<sup>a</sup>Das vom **Recognizer** gelöste Problem ist auch als **Wortproblem** bekannt.

<sup>b</sup>Thiemann, „Compilerbau“.

Für das **Parsen** gibt es grundsätzlich **zwei** verschiedene Ansätze:

- **Top-Down Parsing:** Der **Derivation Tree** wird von **oben-nach-unten** generiert, also von der **Wurzel** zu den **Blättern**. Dementsprechend fängt die Generierung des **Derivation Tree** mit dem **Startsymbol** der **Grammatik** an und wendet in jedem Schritt eine **Linksableitung** auf die **Nicht-Terminalsymbole** an, bis man **Terminalsymbole** hat, die sich zum gewünschten **Inputstring** abgeleitet haben oder sich herausstellt, dass dieser nicht abgeleitet werden kann.<sup>a</sup>

Der Grund, warum die **Linksableitung** verwendet wird und nicht z.B. die **Rechtsableitung**, ist, weil der **Eingabewert** bzw. der **Inputstring** von **links nach rechts** eingelesen wird, was gut damit zusammenpasst, dass die **Linksableitung** die **Blätter** von **links-nach-rechts** generiert.

Welche der **Produktionen** für ein **Nicht-Terminalsymbol** angewandt wird, wenn es mehrere **Alternativen** gibt, wird entweder durch **Backtracking** oder durch **Vorausschauen** gelöst.

Eine sehr einfach zu implementierende Technik für **Top-Down Parser** ist hierbei der **Rekursive Abstieg**. Dabei wird jedem **Nicht-Terminalsymbol** eine **Prozedur** zugeordnet, welche die **Produktionen** dieses Nicht-Terminalsymbols umsetzt. **Prozeduren** rufen sich dabei wechselseitig gegenseitig entsprechend der Produktionsregeln auf, falls eine Produktionsregel ein entsprechendes **Nicht-Terminal** enthält.

Mit dieser Methode ist das Parsen **Linksrekursiver Grammatiken** (Definition 2.22) allerdings nicht möglich, ohne die Grammatik vorher umgeformt zu haben und jegliche **Linksrekursion** aus der **Grammatik** entfernt zu haben, da diese zu **Unendlicher Rekursion** führt.

**Rekursiver Abstieg** kann mit **Backtracking** verbunden werden, um auch Grammatiken parsen zu können, die nicht **LL(k)** (Definition 2.28) sind. Dabei werden meist nach dem **Depth-First-Search Prinzip** alle **Produktionen** für ein **Nicht-Terminalsymbol** solange durchgegangen bis der gewünschte Inputstring abgeleitet ist oder alle **Alternativen** für einen Schritt abgesucht sind, bis man wieder beim ersten Schritt angekommen ist und da auch alle **Alternativen** abgesucht sind, was dann bedeutet, dass der **Inputstring** sich **nicht** mit der verwendeten Grammatik

ableiten lässt.<sup>b</sup>

Wenn man eine **LL(k)** Grammatik hat, kann man auf **Backtracking verzichten** und es reicht einfach nur immer  $k$  **Token** im Inputstring **vorauszuschauen**. **Mehrdeutige Grammatiken** sind dadurch ausgeschlossen, weil **LL(k)** keine **Mehrdeutigkeit** zulässt.<sup>c</sup>

- **Bottom-Up Parsing:** Es wird mit dem **Eingabewort** bzw. **Inputstring** gestartet und versucht **Rechtsableitungen** entsprechend der **Produktionen** der **Konkreten Syntax** rückwärts anzuwenden, bis man beim **Startsymbol** landet.<sup>d</sup>
- **Chart Parser:** Es wird **Dynamische Programmierung** verwendet und partielle Zwischenergebnisse werden in einer **Tabelle** (bzw. einem **Chart**) gespeichert und können wiederverwendet werden. Das macht das Parsen **Kontextfreier Grammatiken** effizienter, sodass es nur noch **polynomielle** Zeit braucht, da **Backtracking** nicht mehr notwendig ist.<sup>e</sup>

<sup>a</sup>What is Top-Down Parsing?

<sup>b</sup>Diese Form von Parsing wurde im **PicoC-Compiler** implementiert, als dieser noch auf dem Stand des **Bachelorprojektes** war, bevor er durch den nicht selbst implementierten **Earley Parser** von **Lark** (siehe *Lark - a parsing toolkit for Python*) ersetzt wurde.

<sup>c</sup>Diese Art von Parser ist im **RETI-Interpreter** implementiert, da die **RETI-Sprache** eine besonders simple **LL(1) Grammatik** besitzt. Diese Art von **Parser** wird auch als **Predictive Parser** oder **LL(k) Recursive Descent Parser** bezeichnet, wobei **Recursive Descent** das englische Wort für **Rekursiven Abstieg** ist.

<sup>d</sup>What is Bottom-up Parsing?

<sup>e</sup>Der **Earley Parser**, den **Lark** und damit der **PicoC-Compiler** verwendet fällt unter diese Kategorie.

Der **Abstract Syntax Tree** wird mithilfe von **Transformern** (Definition 2.38) und **Visitors** (Definition 2.39) generiert und ist das Endprodukt der **Syntaktischen Analyse**. Wenn man die gesamte **Syntaktische Analyse** betrachtet, so übersetzt diese einen Inputstring von der **Konkreten Syntax** in die **Abstrakte Syntax** (Definition 2.40).

#### Definition 2.38: Transformer

Ein Programm, dass von **unten-nach-oben**, nach dem **Breadth First Search** Prinzip alle Knoten des **Derivation Tree** besucht und beim Antreffen eines bestimmten Knoten des **Derivation Tree** einen entsprechenden Knoten des **Abstract Syntax Tree** erzeugt und diesen anstelle des Knotens des **Derivation Tree** setzt und so Stück für Stück den **Abstract Syntax Tree** konstruiert.<sup>a</sup>

<sup>a</sup>Transformers & Visitors — Lark documentation.

#### Definition 2.39: Visitor

Ein Programm, dass von **unten-nach-oben**, nach dem **Breadth First Search** Prinzip alle Knoten des **Derivation Tree** besucht und in Bezug zu Compilerbau, beim Antreffen eines bestimmten **Knoten** des **Derivation Tree**, diesen **in-place** mit anderen Knoten **tauscht** oder **manipuliert**, um den **Derivation Tree** für die weitere Verarbeitung durch z.B. einen **Transformer** zu vereinfachen.<sup>a,b</sup>

<sup>a</sup>Kann theoretisch auch zur Konstruktion eines **Abstract Syntax Tree** verwendet werden, wenn z.B. eine externe Klasse verwendet wird, welches für die Konstruktion des **Abstract Syntax Tree** verantwortlich ist. Aber dafür ist ein **Transformer** besser geeignet.

<sup>b</sup>Transformers & Visitors — Lark documentation.

**Definition 2.40: Abstrakte Syntax**

*Syntax*, die beschreibt, was für Arten von **Komposition** bei den **Knoten** eines **Abstract Syntax Trees** möglich sind.

Jene Produktionen, die in der **Konkreten Syntax** für die Umsetzung von **Präzidenz** notwendig waren, sind in der **Abstrakten Syntax** abgeflacht.<sup>a</sup>

<sup>a</sup>G. Siek, *Course Webpage for Compilers (P423, P523, E313, and E513)*.

**Definition 2.41: Abstract Syntax Tree**

*Compilerinterne Darstellung* eines Programs, in welcher sich anhand der Knoten auf dem **Pfad** von der **Wurzel** zu einem **Blatt** nicht mehr direkt nachvollziehen lässt, durch welche **Produktionen** dieses Blatt abgeleitet wurde.

Der **Abstract Syntax Tree** hat einmal den Zweck, dass die **Kompositionen**, die die Knoten bilden können **semantisch** näher an den **Instructions eines Assemblers** dran sind und, dass man mit einem **Abstract Syntax Tree** bei der Betrachtung eines **Knoten**, der für einen Teil des Programms steht, möglichst schnell die Fragen beantworten kann, welche **Funktionalität** der Sprache dieser umsetzt, welche **Bestandteile** er hat und welche **Funktionalität** der Sprache diese Bestandteile umsetzen usw.<sup>a</sup>

<sup>a</sup>G. Siek, *Course Webpage for Compilers (P423, P523, E313, and E513)*.

Die **Baumdatenstruktur** des **Derivation Tree** und **Abstract Syntax Tree** ermöglicht es die Operationen, die ein Compiler bzw. Interpreter bei der Weiterverarbeitung des Inputstrings ausführen muss möglichst **effizient** auszuführen und auf **unkomplizierte** Weise direkt zu erkennen, welche er ausführen muss.

Um eine Gesamtübersicht über die **Syntaktische Analyse** zu geben, ist in Abbildung 2.4 die Syntaktische mit dem Beispiel aus Subkapitel 2.3 fortgeführt.

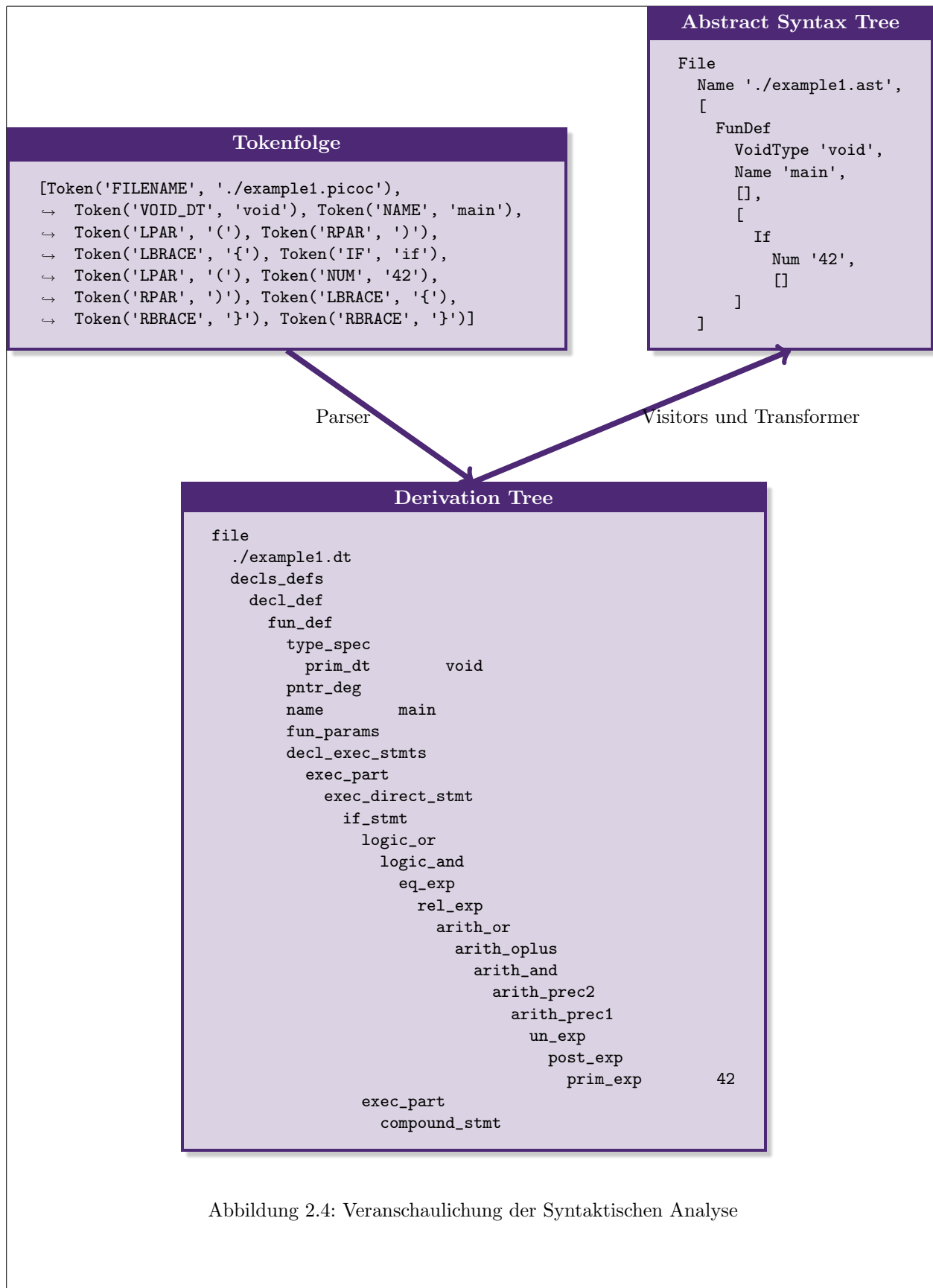


Abbildung 2.4: Veranschaulichung der Syntaktischen Analyse

## 2.5 Code Generierung

### Definition 2.42: Pass

*a*

<sup>a</sup>G. Siek, *Course Webpage for Compilers (P423, P523, E313, and E513)*.

### Definition 2.43: Monadische Normalform

*a*

<sup>a</sup>G. Siek, *Course Webpage for Compilers (P423, P523, E313, and E513)*.

Ein echter Compiler verwendet Graph Coloring ... Register ...

## 2.6 Fehlermeldungen

### Definition 2.44: Fehlermeldung

*Benachrichtigung* beliebiger Form, die darüber informiert, dass:

1. Ein Program beim **Kompilieren** von der **Konkreten Syntax** abweicht, also der **Inputstring** sich nicht mit der Konkreten Syntax **ableiten** lässt oder auf etwas **zugegriffen** werden soll, was noch **nicht** deklariert oder definiert wurde.
2. Beim Ausführen eine **verbotene** Operation ausgeführt wurde.<sup>a</sup>

<sup>a</sup>Errors in C/C++ - GeeksforGeeks.

### 2.6.1 Kategorien von Fehlermeldungen





# 3 Implementierung

## 3.1 Lexikalische Analyse

### 3.1.1 Konkrete Syntax für die Lexikalische Analyse

<i>COMMENT</i>	::=	"//"/[ $\backslash$ $\backslash$ $n$ ]* /   "/"* /( $\backslash$ $n$ )*?/ "*" /	<i>L_Comment</i>
<i>RETI_COMMENT.2</i>	::=	"//""?"#"/[ $\backslash$ $\backslash$ $n$ ]* /	
<i>DIG_NO_0</i>	::=	"1"   "2"   "3"   "4"   "5"   "6"   "7"   "8"   "9"	<i>L_Arith</i>
<i>DIG_WITH_0</i>	::=	"0"   <i>DIG_NO_0</i>	
<i>NUM</i>	::=	"0"   <i>DIG_NO_0</i> <i>DIG_WITH_0</i> *	
<i>ASCII_CHAR</i>	::=	"_".~"	
<i>CHAR</i>	::=	"'" <i>ASCII_CHAR</i> "'"	
<i>FILENAME</i>	::=	<i>ASCII_CHAR</i> + ".picoc"	
<i>LETTER</i>	::=	"a"..z"   "A"..Z"	
<i>NAME</i>	::=	( <i>LETTER</i>   "_") ( <i>LETTER</i> — <i>DIG_WITH_0</i> — "_")*	
<i>name</i>	::=	<i>NAME</i>   <i>INT_NAME</i>   <i>CHAR_NAME</i>   <i>VOID_NAME</i>	
<i>NOT</i>	::=	"~"	
<i>REF_AND</i>	::=	"&"	
<i>un_op</i>	::=	<i>SUB_MINUS</i>   <i>LOGIC_NOT</i>   <i>NOT</i>   <i>MUL_DEREF_PNTR</i>   <i>REF_AND</i>	
<i>MUL_DEREF_PNTR</i>	::=	"*"	
<i>DIV</i>	::=	"/"	
<i>MOD</i>	::=	"%"	
<i>prec1_op</i>	::=	<i>MUL_DEREF_PNTR</i>   <i>DIV</i>   <i>MOD</i>	
<i>ADD</i>	::=	"+"	
<i>SUB_MINUS</i>	::=	"-"	
<i>prec2_op</i>	::=	<i>ADD</i>   <i>SUB_MINUS</i>	
<i>LT</i>	::=	"<"	<i>L_Logic</i>
<i>LTE</i>	::=	"<="	
<i>GT</i>	::=	">"	
<i>GTE</i>	::=	">="	
<i>rel_op</i>	::=	<i>LT</i>   <i>LTE</i>   <i>GT</i>   <i>GTE</i>	
<i>EQ</i>	::=	"=="	
<i>NEQ</i>	::=	"!="	
<i>eq_op</i>	::=	<i>EQ</i>   <i>NEQ</i>	
<i>LOGIC_NOT</i>	::=	"!"	
<i>INT_DT.2</i>	::=	"int"	
<i>INT_NAME.3</i>	::=	"int" ( <i>LETTER</i>   <i>DIG_WITH_0</i>   "_")+	<i>L_Assign_Alloc</i>
<i>CHAR_DT.2</i>	::=	"char"	
<i>CHAR_NAME.3</i>	::=	"char" ( <i>LETTER</i>   <i>DIG_WITH_0</i>   "_")+	
<i>VOID_DT.2</i>	::=	"void"	
<i>VOID_NAME.3</i>	::=	"void" ( <i>LETTER</i>   <i>DIG_WITH_0</i>   "_")+	
<i>prim_dt</i>	::=	<i>INT_DT</i>   <i>CHAR_DT</i>   <i>VOID_DT</i>	

### 3.1.2 Basic Lexer

## 3.2 Syntaktische Analyse

### 3.2.1 Konkrete Syntax für die Syntaktische Analyse

In 3.2.1

<i>prim_exp</i>	::=	<i>name</i>   <i>NUM</i>   <i>CHAR</i>   "(" <i>logic_or</i> ")"	<i>L_Arith</i> +
<i>post_exp</i>	::=	<i>array_subscr</i>   <i>struct_attr</i>   <i>fun_call</i>	<i>L_Array</i> +
		<i>input_exp</i>   <i>print_exp</i>   <i>prim_exp</i>	<i>L_Pntr</i> +
<i>un_exp</i>	::=	<i>un_opun_exp</i>   <i>post_exp</i>	<i>L_Struct</i> + <i>L_Fun</i>
<i>input_exp</i>	::=	"input" "(" ")"	<i>L_Arith</i>
<i>print_exp</i>	::=	"print" "(" <i>logic_or</i> ")"	
<i>arith_prec1</i>	::=	<i>arith_prec1</i> <i>prec1_op</i> <i>un_exp</i>   <i>un_exp</i>	
<i>arith_prec2</i>	::=	<i>arith_prec2</i> <i>prec2_op</i> <i>arith_prec1</i>   <i>arith_prec1</i>	
<i>arith_and</i>	::=	<i>arith_and</i> "&" <i>arith_prec2</i>   <i>arith_prec2</i>	
<i>arith_oplus</i>	::=	<i>arith_oplus</i> "^" <i>arith_and</i>   <i>arith_and</i>	
<i>arith_or</i>	::=	<i>arith_or</i> " " <i>arith_oplus</i>   <i>arith_oplus</i>	
<i>rel_exp</i>	::=	<i>rel_exp</i> <i>rel_op</i> <i>arith_or</i>   <i>arith_or</i>	<i>L_Logic</i>
<i>eq_exp</i>	::=	<i>eq_exp</i> <i>eq_oprel_exp</i>   <i>rel_exp</i>	
<i>logic_and</i>	::=	<i>logic_and</i> "&&" <i>eq_exp</i>   <i>eq_exp</i>	
<i>logic_or</i>	::=	<i>logic_or</i> "  " <i>logic_and</i>   <i>logic_and</i>	
<i>type_spec</i>	::=	<i>prim_dt</i>   <i>struct_spec</i>	<i>L_Assign_Alloc</i>
<i>alloc</i>	::=	<i>type_spec</i> <i>pntr_decl</i>	
<i>assign_stmt</i>	::=	<i>un_exp</i> "=" <i>logic_or</i> ";"	
<i>initializer</i>	::=	<i>logic_or</i>   <i>array_init</i>   <i>struct_init</i>	
<i>init_stmt</i>	::=	<i>alloc</i> "=" <i>initializer</i> ";"	
<i>const_init_stmt</i>	::=	"const" <i>type_spec</i> <i>name</i> "=" <i>NUM</i> ";"	
<i>pntr_deg</i>	::=	"*" *	<i>L_Pntr</i>
<i>pntr_decl</i>	::=	<i>pntr_deg</i> <i>array_decl</i>   <i>array_decl</i>	
<i>array_dims</i>	::=	("[" <i>NUM</i> "]" ) *	<i>L_Array</i>
<i>array_decl</i>	::=	<i>name</i> <i>array_dims</i>   "(" <i>pntr_decl</i> ")" <i>array_dims</i>	
<i>array_init</i>	::=	"{" <i>initializer</i> ("," <i>initializer</i> ) * "}"	
<i>array_subscr</i>	::=	<i>post_exp</i> "[" <i>logic_or</i> "]"	
<i>struct_spec</i>	::=	"struct" <i>name</i>	<i>L_Struct</i>
<i>struct_params</i>	::=	( <i>alloc</i> ";" ) +	
<i>struct_decl</i>	::=	"struct" <i>name</i> "{" <i>struct_params</i> "}"	
<i>struct_init</i>	::=	"{" " " <i>name</i> "=" <i>initializer</i> ("," " " <i>name</i> "=" <i>initializer</i> ) * "}"	
<i>struct_attr</i>	::=	<i>post_exp</i> "." <i>name</i>	
<i>if_stmt</i>	::=	"if" "(" "(" <i>logic_or</i> ")" ")" <i>exec_part</i>	<i>L_If_Else</i>
<i>if_else_stmt</i>	::=	"if" "(" "(" <i>logic_or</i> ")" ")" <i>exec_part</i> "else" <i>exec_part</i>	
<i>while_stmt</i>	::=	"while" "(" "(" <i>logic_or</i> ")" ")" <i>exec_part</i>	<i>L_Loop</i>
<i>do_while_stmt</i>	::=	"do" <i>exec_part</i> "while" "(" "(" <i>logic_or</i> ")" ")" ";"	

Grammar 3.2.1: Konkrete Syntax Syntaktische Analyse in EBNF, Teil 1

<i>decl_exp_stmt</i>	::=	<i>alloc</i> ";"	<i>L_Stmt</i>
<i>decl_direct_stmt</i>	::=	<i>assign_stmt</i>   <i>init_stmt</i>   <i>const_init_stmt</i>	
<i>decl_part</i>	::=	<i>decl_exp_stmt</i>   <i>decl_direct_stmt</i>   <i>RETI_COMMENT</i>	
<i>compound_stmt</i>	::=	"{" <i>exec_part</i> * "}"	
<i>exec_exp_stmt</i>	::=	<i>logic_or</i> ";"	
<i>exec_direct_stmt</i>	::=	<i>if_stmt</i>   <i>if_else_stmt</i>   <i>while_stmt</i>   <i>do_while_stmt</i>   <i>assign_stmt</i>   <i>fun_return_stmt</i>	
<i>exec_part</i>	::=	<i>compound_stmt</i>   <i>exec_exp_stmt</i>   <i>exec_direct_stmt</i>   <i>RETI_COMMENT</i>	
<i>decl_exec_stmts</i>	::=	<i>decl_part</i> * <i>exec_part</i> *	
<i>fun_args</i>	::=	[ <i>logic_or</i> ("," <i>logic_or</i> )*]	<i>L_Fun</i>
<i>fun_call</i>	::=	<i>name</i> ("(" <i>fun_args</i> ")")	
<i>fun_return_stmt</i>	::=	"return" [ <i>logic_or</i> ] ";"	
<i>fun_params</i>	::=	[ <i>alloc</i> ("," <i>alloc</i> )*]	
<i>fun_decl</i>	::=	<i>type_spec</i> <i>pntr_deg</i> <i>name</i> ("(" <i>fun_params</i> ")")	
<i>fun_def</i>	::=	<i>type_spec</i> <i>pntr_deg</i> <i>name</i> ("(" <i>fun_params</i> ")") " {" <i>decl_exec_stmts</i> "}"	
<i>decl_def</i>	::=	( <i>struct_decl</i>   <i>fun_decl</i> ) ";"   <i>fun_def</i>	<i>L_File</i>
<i>decls_defs</i>	::=	<i>decl_def</i> *	
<i>file</i>	::=	<i>FILENAME</i> <i>decls_defs</i>	

Grammar 3.2.2: Konkrete Syntax für die Syntaktische Analyse in EBNF, Teil 2

### 3.2.2 Umsetzung von Präzidenz

Die **PicoC** Programmiersprache hat dieselben **Präzidenzregeln** implementiert, wie die Programmiersprache **C**<sup>1</sup>. Die **Präzidenzregeln** von **PicoC** sind in Tabelle 3.1 aufgelistet.

Präzidenz	Operator	Beschreibung	Assoziativität
1	<i>a()</i>	Funktionsaufruf	Links, dann rechts →
	<i>a[]</i>	Indezzugriff	
	<i>a.b</i>	Attributzugriff	
2	<i>-a</i>	Unäres Minus	Rechts, dann links ←
	<i>!a ~a</i>	Logisches NOT und Bitweise NOT	
	<i>*a &amp;a</i>	Dereferenz und Referenz, auch Adresse-von	
3	<i>a*b a/b a%b</i>	Multiplikation, Division und Modulo	Links, dann rechts →
4	<i>a+b a-b</i>	Addition und Subtraktion	
5	<i>a&lt;b a&lt;=b a&gt;b a&gt;=b</i>	Kleiner, Kleiner Gleich, Größer, Größer gleich	
6	<i>a==b a!=b</i>	Gleichheit und Ungleichheit	
7	<i>a&amp;b</i>	Bitweise UND	
8	<i>a^b</i>	Bitweise XOR (exclusive or)	
9	<i>a b</i>	Bitweise ODER (inclusive or)	
10	<i>a&amp;&amp;b</i>	Logisches UND	
11	<i>a  b</i>	Logisches ODER	
12	<i>a=b</i>	Zuweisung	Rechts, dann links ←
13	<i>a,b</i>	Komma	Links, dann rechts →

Tabelle 3.1: Präzidenzregeln von PicoC

<sup>1</sup>*C Operator Precedence - [cppreference.com](http://cppreference.com).*

### 3.2.3 Derivation Tree Generierung

#### 3.2.3.1 Early Parser

#### 3.2.3.2 Codebeispiel

```

1 struct st {int *(*attr)[5][6];};
2
3 void main() {
4     struct st *(*var)[3][2];
5 }

```

Code 3.1: PicoC Code für Derivation Tree Generierung

```

1 file
2   ./example_dt_simple_ast_gen_array_decl_and_alloc.dt
3 decls_defs
4   decl_def
5     struct_decl
6       name st
7       struct_params
8       alloc
9       type_spec
10      prim_dt int
11      pntr_decl
12      pntr_deg *
13      array_decl
14      pntr_decl
15      pntr_deg *
16      array_decl
17      name attr
18      array_dims
19      array_dims
20      5
21      6
22   decl_def
23   fun_def
24     type_spec
25     prim_dt void
26     pntr_deg
27     name main
28     fun_params
29     decl_exec_stmts
30     decl_part
31     decl_exp_stmt
32     alloc
33     type_spec
34     struct_spec
35     name st
36     pntr_decl
37     pntr_deg *
38     array_decl
39     pntr_decl
40     pntr_deg *

```

```

41         array_decl
42         name var
43         array_dims
44     array_dims
45     3
46     2

```

Code 3.2: Derivation Tree nach Derivation Tree Generierung

### 3.2.4 Derivation Tree Vereinfachung

#### 3.2.4.1 Visitor

#### 3.2.4.2 Codebeispiel

Beispiel aus Subkapitel 3.2.3.2 wird fortgeführt.

```

1 file
2   ./example_dt_simple_ast_gen_array_decl_and_alloc.dt_simple
3   decls_defs
4     decl_def
5       struct_decl
6         name st
7         struct_params
8         alloc
9         ptr_decl
10        ptr_deg *
11        array_decl
12        array_dims
13        5
14        6
15        ptr_decl
16        ptr_deg *
17        array_decl
18        array_dims
19        type_spec
20        prim_dt int
21    name attr
22  decl_def
23    fun_def
24      type_spec
25      prim_dt void
26      ptr_deg
27      name main
28      fun_params
29      decl_exec_stmts
30      decl_part
31      decl_exp_stmt
32      alloc
33      ptr_decl
34      ptr_deg *
35      array_decl
36      array_dims

```

```
37         3
38         2
39     pntr_decl
40     pntr_deg *
41     array_decl
42     array_dims
43     type_spec
44     struct_spec
45     name st
46 name var
```

Code 3.3: Derivation Tree nach Derivation Tree Vereinfachung

### 3.2.5 Abstrakt Syntax Tree Generierung

#### 3.2.5.1 PicoC-Knoten

PiocC-Knoten	Beschreibung
Name()	Ein <b>Bezeichner</b> , aber da es keine gute Kurzform für Identifier() (englisches Wort für Bezeichner) gibt, wurde dieser Knoten Name() genannt.
Num()	Eine <b>Zahl</b> .
Char()	Ein <b>Zeichen</b> der <b>ASCII-Zeichenkodierung</b> .
Minus(), Not(), DerefOp(), RefOp(), LogicNot()	Die <b>unären Operatoren</b> un_op: -a, ~a, *a, &a !a.
Add(), Sub(), Mul(), Div(), Mod(), Oplus(), And(), Or(), LogicAnd(), LogicOr()	Die <b>binären Operatoren</b> bin_op: a + b, a - b, a * b, a / b, a % b, a ^ b, a & b, a   b, a && b, a    b.
Eq(), NEq(), Lt(), LtE(), Gt(), GtE()	Die <b>Relationen</b> rel: a == b, a != b, a < b, a <= b, a > b, a >= b.
Const(), Writeable()	Die <b>Type Qualifier</b> type_qual: const, was für ein <b>nicht beschreibbare Konstante</b> steht und das <b>nicht</b> Angeben von const, was für einen <b>beschreibbare</b> Variable steht.
IntType(), CharType(), VoidType()	Die <b>Type Specifier</b> für <b>Primitiven Datentypen</b> , die in der Abstrakten Syntax, um eine intuitive Bezeichnung zu haben einfach nur unter <b>Datentypen</b> eingeordnet werden: int, char, void.
Placeholder()	<b>Platzhalter</b> für einen Knoten, der diesen später <b>ersetzt</b> .
BinOp(exp, bin_op, exp)	Container für eine <b>binäre Operation</b> mit 2 Expressions.
UnOp(un_op, exp)	Container für eine <b>unäre Operation</b> mit einer Expression.
Exit(num)	Container für einen <b>Exit Code</b> , der vor der Beendigung in das ACC Register geschrieben wird und steht für die <b>Beendigung</b> des laufenden Programmes.
Atom(exp, rel, exp)	Container für eine <b>binäre Relation</b> mit 2 Expressions.
ToBool(exp)	Container für einen <b>Arithmetischen Ausdruck</b> , wie z.B. 1 + 3 oder einfach nur 3, der nicht nur 1 oder 0 als Ergebnis haben kann und daher bei einem Ergebnis $x > 1$ auf 1 abgebildet wird.
Alloc(type_qual, datatype, name, local_var_or_param)	<b>Container</b> für eine <b>Allokation</b> mit den notwendigen Knoten type_qual, datatype und name, die alle für einen Eintrag in der <b>Symboltabelle</b> notwendigen Informationen enthalten. Zudem besitzt er ein verstecktes Attribut local_var_or_param, dass die Information trägt, ob es sich bei der <b>Variable</b> um eine <b>Lokale Variable</b> oder einen <b>Parameter</b> handelt.
Assign(lhs, exp)	Container für eine <b>Zuweisung</b> , wobei lhs ein Subscr(exp1, exp2), Deref(exp1, exp2), Attr(exp, name) oder Name('var') sein kann und exp ein beliebiger <b>Logischer Ausdruck</b> sein kann.
Exp(exp, datatype, error_data)	Container für einen <b>beliebigen Ausdruck</b> , dessen Ergebnis auf den <b>Stack</b> soll. Zudem besitzt er 2 versteckte Attribute, wobei datatype im <b>RETI Blocks Pass</b> wichtig ist und error_data für <b>Fehlermeldungen</b> wichtig ist.
Stack(num)	Container, der für das <b>temporäre</b> Ergebnis einer Berechnung, das num Speicherzellen relativ zum <b>Stackpointer Register</b> SP steht.

Tabelle 3.2: PicoC-Knoten Teil 1

PiocC-Knoten	Beschreibung
Stackframe(num)	Container, der für eine Variable steht, die num Speicherzellen relativ zum <b>Begin-Aktive-Funktion Register</b> BAF steht.
Global(num)	Container, der für eine Variable steht, die num Speicherzellen relativ zum <b>Datensegment Register</b> DS steht.
StackMalloc(num)	Container, der für das <b>Allokieren</b> von num Speicherzellen auf dem <b>Stack</b> steht.
PntrDecl(num, datatype)	Container, der für den <b>Pointerdatatype</b> steht, wobei das <b>Attribut</b> num die <b>Anzahl zusammengefasster Pointer</b> angibt und datatype der Datentyp ist, auf den der oder die <b>Pointer</b> zeigen.
Ref(exp, datatype, error_data)	Container, der für die Anwendung des <b>Referenz-Operators</b> steht. Zudem besitzt er 2 versteckte Attribute, wobei datatype im <b>RETI Blocks Pass</b> wichtig ist und error_data für <b>Fehlermeldungen</b> wichtig ist.
Deref(lhs, exp)	Container für den <b>Indexzugriff</b> auf einen <b>Array-</b> oder <b>Pointerdatatype</b> , wobei exp1 eine angehängte weitere Subscr(exp1, exp2), Deref(exp1, exp2), Attr(exp, name) oder ein Name('var') sein kann und exp2 der Index ist auf den zugegriffen werden soll.
ArrayDecl(nums, datatype)	Container, der für den <b>Arraydatatype</b> steht, wobei das <b>Attribut</b> nums eine Liste von Num('x') ist, die die <b>Dimensionen</b> des Arrays angibt und datatype der Datentyp ist, der über das Anwenden von Subscript() auf das Array zugreifbar ist.
Array(exps, datatype)	Container für den <b>Initializer</b> eines <b>Arrays</b> , dessen Einträge exps weitere Initializer für eine <b>Array-Dimension</b> oder ein Initializer für ein <b>Struct</b> oder ein <b>Logischer Ausdruck</b> sein können. Des Weiteren besitzt er ein verstecktes Attribut datatype, welches für den <b>PicoC-Mon Pass</b> Informationen transportiert, die für <b>Fehlermeldungen</b> wichtig sind.
Subscr(exp1, exp2)	Container für den <b>Indexzugriff</b> auf einen <b>Array-</b> oder <b>Pointerdatatype</b> , wobei exp1 eine angehängte weitere Subscr(exp1, exp2), Deref(exp1, exp2) oder Attr(exp, name) Operation sein kann oder ein Name('var') sein kann und exp2 der Index ist auf den zugegriffen werden soll.
StructSpec(name)	Container für einen selbst definierten <b>Structdatatype</b> , wobei das <b>Attribut</b> name festlegt, welchen <b>selbst definierte</b> Structdatatype dieser Container-Knoten repräsentiert.
Attr(exp, name)	Container für den <b>Attributzugriff</b> auf einen <b>Structdatatype</b> , wobei exp1 eine angehängte weitere Subscr(exp1, exp2), Deref(exp1, exp2) oder Attr(exp, name) Operation sein kann oder ein Name('var') sein kann und name das Attribut ist, auf das zugegriffen werden soll.
Struct(assigns, datatype)	Container für den <b>Initializer</b> eines <b>Structs</b> , dessen Eintrag assigns eine Liste von Assign(lhs, exp) ist mit einer Zuordnung eines <b>Attributezeichners</b> , zu einem weiteren Initializer für eine <b>Array-Dimension</b> oder zu einem Initializer für ein <b>Struct</b> oder zu einem <b>Logischen Ausdruck</b> . Des Weiteren besitzt er ein verstecktes Attribut datatype, welches für den <b>PicoC-Mon Pass</b> Informationen transportiert, die für <b>Fehlermeldungen</b> wichtig sind.

Tabelle 3.3: PicoC-Knoten Teil 2



Piocc-Knoten	Beschreibung
StructDecl(name, allocs)	Container für die Deklaration eines <b>selbstdefinierten Structdatentyps</b> , wobei <b>name</b> der <b>Bezeichner</b> des Structdatentyps ist und <b>allocs</b> eine Liste von Bezeichnern der <b>Attribute</b> des Structdatentyps mit dazugehörigem <b>Datentyp</b> , wofür sich der <b>Container-Knoten</b> Alloc(type_qual, datatype, name) sehr gut als <b>Container</b> eignet.
If(exp, stmts_goto)	Container für ein <b>If Statement</b> inklusive <b>Condition</b> exp und einem <b>Branch</b> stmts_goto, indem eine <b>Liste von Statements</b> stehen kann oder ein einzelnes GoTo(Name('block.xyz')).
IfElse(exp, stmts_goto1, stmts_goto2)	Container für ein <b>If-Else Statement</b> inklusive <b>Condition</b> exp und 2 <b>Branches</b> stmts_goto1 und stmts_goto2, die zwei Alternativen darstellen in denen jeweils <b>Liste von Statements</b> oder GoTo(Name('block.xyz'))'s stehen können.
While(exp, stmts_goto)	Container für ein <b>While-Statement</b> inklusive <b>Condition</b> exp und einem <b>Branch</b> stmts_goto, indem eine <b>Liste von Statements</b> stehen kann oder ein einzelnes GoTo(Name('block.xyz')).
DoWhile(exp, stmts_goto)	Container für ein <b>Do-While-Statement</b> inklusive <b>Condition</b> exp und einem <b>Branch</b> stmts_goto, indem eine <b>Liste von Statements</b> stehen kann oder ein einzelnes GoTo(Name('block.xyz')).
Call(name, exps)	Container für einen <b>Funktionsaufruf</b> : fun_name(exps), wobei <b>name</b> der <b>Bezeichner</b> der Funktion ist, die aufgerufen werden soll und <b>exps</b> eine <b>Liste von Argumenten</b> ist, die an die Funktion übergeben werden soll.
Return(exp)	Container für ein <b>Return-Statement</b> , wobei das <b>Attribut</b> exp einen <b>Logischen Ausdruck</b> darstellt, dessen Ergebnis vom <b>Return-Statement</b> zurückgegeben wird.
FunDecl(datatype, name, allocs)	Container für eine <b>Funktionsdeklaration</b> , wobei datatype der <b>Rückgabewert</b> der Funktion ist, name der <b>Bezeichner</b> der Funktion ist und allocs die <b>Parameter</b> der Funktion sind, wobei der <b>Container-Knoten</b> Alloc(type_spec, datatype, name) als Container für die Parameter dient.
FunDef(datatype, name, allocs, stmts_blocks)	Container für eine <b>Funktionsdefinition</b> , wobei datatype der <b>Rückgabewert</b> der Funktion ist, name der <b>Bezeichner</b> der Funktion ist, allocs die <b>Parameter</b> der Funktion sind, wobei der <b>Container-Knoten</b> Alloc(type_spec, datatype, name) als Container für die Parameter dient und stmts_blocks eine Liste von <b>Statements</b> bzw. <b>Blöcken</b> ist, welche diese Funktion beinhaltet.
NewStackframe(fun_name, goto_after_call)	Container für die <b>Erstellung</b> eines neuen <b>Stackframes</b> , wobei fun_name der Bezeichner der Funktion ist, für die ein neuer <b>Stackframe</b> erstellt werden soll und später dazu dient den <b>Block</b> dieser Funktion zu finden, weil dieser für den weiteren Kompilierungsvorgang wichtige Information in seinen <b>versteckten Attribute</b> angehängt hat und goto_after_call ein GoTo(Name('addr@next_instr')) ist, welches später durch die <b>Adresse</b> der Instruction, die direkt auf die <b>Jump Instruction</b> folgt, ersetzt wird.
RemoveStackframe()	Container für das <b>Entfernen</b> des aktuellen <b>Stackframes</b> .

Tabelle 3.4: PicoC-Knoten Teil 3

PiocC-Knoten	Beschreibung
<code>File(name, decls_defs_blocks)</code>	Container für alle <b>Funktionen</b> oder <b>Blöcke</b> , welche eine Datei als Ursprung haben, wobei <b>name</b> der <b>Dateiname</b> der Datei ist, die erstellt wird und <b>decls_defs_blocks</b> eine Liste von <b>Funktionen</b> bzw. <b>Blöcken</b> ist.
<code>Block(name, stmts_instrs, instrs_before, num_instrs, param_size, local_vars_size)</code>	Container für <b>Statements</b> , der auch als <b>Block</b> bezeichnet wird, wobei das Attribut <b>name</b> der Bezeichner des <b>Labels</b> des Blocks ist und <b>stmts_instrs</b> eine <b>Liste von Statements</b> oder <b>Instructions</b> . Zudem besitzt er noch 3 <b>versteckte Attribute</b> , wobei <b>instrs_before</b> die Zahl der <b>Instructions</b> vor diesem <b>Block</b> zählt, <b>num_instrs</b> die Zahl der Instructions ohne Kommentare in diesem Block zählt, <b>param_size</b> die voraussichtliche Anzahl an Speicherzellen aufaddiert, die für die <b>Parameter</b> der Funktion belegt werden müssen und <b>local_vars_size</b> die voraussichtliche Anzahl an Speicherzellen aufaddiert, die für die <b>lokalen Variablen</b> der Funktion belegt werden müssen.
<code>GoTo(name)</code>	Container für ein <b>Goto</b> zu einem anderen <b>Block</b> , wobei das Attribut <b>name</b> der Bezeichner des <b>Labels</b> des Blocks ist zu dem Gesprungen werden soll.
<code>SingleLineComment(prefix, content)</code>	Container für einen <b>Kommentar</b> , den der Compiler selber während des <b>Kompilervorgangs</b> erstellt, der im <b>RETI-Interpreter</b> selbst später <b>nicht</b> sichtbar sein wird, aber in den <b>Immediate-Dateien</b> , welche die <b>Abstract Syntax Trees</b> nach den verschiedenen <b>Passes</b> enthalten.
<code>RETIComment(value)</code>	Container für einen <b>Kommentar</b> im Code der Form: <code>// # comment</code> , der im <b>RETI-Interpreter</b> später sichtbar sein wird und zur Orientierung genutzt werden kann, allerdings in einer tatsächlichen Implementierung einer <b>RETI-CPU nicht umsetzbar</b> ist und auch nicht sinnvoll wäre umzusetzen. Der <b>Kommentar</b> ist im Attribut <b>value</b> , welches jeder Knoten besitzt gespeichert.

Tabelle 3.5: PicoC-Knoten Teil 4

Die ausgegrauten Attribute der PicoC-Nodes sind **versteckte Attribute**, die **nicht** direkt bei der Erstellung der **PicoC-Nodes** mit einem Wert **initialisiert** werden, sondern im **Verlauf der Kompilierung** beim Durchlaufen der verschiedenen Passes etwas zugewiesen bekommen, dass im weiteren Kompilierungsvorgang **Informationen** transportiert, die später im Kompilierungsvorgang nicht mehr so leicht zugänglich wären.

Jeder **Knoten** hat darüberhinaus auch noch 2 **Attribute** **value** und **position**, wobei **value** bei einem **Token-Knoten** (Definition 3.1) dem **Tokenwert** des Tokens, welches es ersetzt entspricht und bei **Container-Knoten** (Definition 3.2) unbesetzt ist. Das **Attribut** **position** wird später für Fehlermeldungen gebraucht.

**Definition 3.1: Token-Knoten**

Ersetzt ein **Token** bei der Generierung des **Abstract Syntax Tree**, damit der Zugriff auf Knoten des Abstract Syntax Tree möglichst **simpel** ist und keine vermeidbaren Fallunterscheidungen gemacht werden müssen.

**Token-Knoten** entsprechen im Abstract Syntax Tree **Blättern**.<sup>a</sup>

<sup>a</sup>Thiemann, „Compilerbau“.

**Definition 3.2: Container-Knoten**

Dient als **Container** für andere **Container-Knoten** und **Token-Knoten**. Die **Container-Knoten** werden optimalerweise immer so gewählt, dass sie **mehrere Produktionen** der **Konkreten Syntax** abdecken, die einen **gleichen Aufbau** haben und sich auch unter einem **Überbegriff** zusammenfassen lassen.<sup>a</sup>

**Container-Knoten** entsprechen im Abstract Syntax Tree **Inneren Knoten**.<sup>b</sup>

<sup>a</sup>Wie z.B. die verschiedenen **Arithmetischen Ausdrücke**, wie z.B.  $1 \% 3$  und **Logischen Ausdrücke**, wie z.B.  $1 \ \&\& \ 2 < 3$ , die einen gleichen Aufbau haben mit immer einer **Operation** in der Mitte haben und 2 **Operanden** auf beiden Seiten und sich unter dem Überbegriff **Binäre Operationen** zusammenfassen lassen.

<sup>b</sup>Thiemann, „Compilerbau“.

## 3.2.5.2 RETI-Knoten

RETI-Knoten	Beschreibung
Program(name, instrs)	Container für alle <b>Instructions</b> : <name> <instrs>, wobei <b>name</b> der <b>Dateiname</b> der Datei ist, die erstellt wird und <b>instrs</b> eine <b>Liste von Instructions</b> ist.
Instr(op, args)	Container für eine <b>Instruction</b> : <op> <args>, wobei <b>op</b> eine <b>Operation</b> ist und <b>args</b> eine <b>Liste von Argumenten</b> für dieser Operation.
Jump(rel, im_goto)	Container für eine <b>Jump-Instruction</b> : JUMP<rel> <im>, wobei <b>rel</b> eine <b>Relation</b> ist und <b>im_goto</b> ein <b>Immediate Value</b> <b>Im(val)</b> für die <b>Anzahl an Speicherzellen</b> , um die relativ zur <b>Jump-Instruction</b> gesprungen werden soll oder ein <b>GoTo(Name('block.xyz'))</b> , das später im <b>RETI-Patch Pass</b> durch einen passenden <b>Immediate Value</b> ersetzt wird.
Int(num)	Container für einen <b>Interruptaufruf</b> : INT <im>, wobei <b>num</b> die <b>Interruptvektornummer</b> (IVN) für die passende Speicherzelle in der <b>Interruptvektortabelle</b> ist, in der die Adresse der <b>Interrupt-Service-Routine</b> (ISR) steht.
Call(name, reg)	Container für einen <b>Prozeduraufruf</b> : CALL <name> <reg>, wobei <b>name</b> der <b>Bezeichner</b> der Prozedur, die aufgerufen werden soll ist und <b>reg</b> ein <b>Register</b> ist, das als <b>Argument</b> an die Prozedur dient. Diese <b>Operation</b> ist in der Betriebssysteme Vorlesung <sup>a</sup> nicht deklariert, sondern wurde dazuerfunden, um unkompliziert ein CALL PRINT ACC oder CALL INPUT ACC im RETI-Interpreter simulieren zu können.
Name(val)	Bezeichner für eine <b>Prozedur</b> , wie PRINT und INPUT oder den <b>Programmnamen</b> . Dieses <b>Argument</b> ist in der Betriebssysteme Vorlesung <sup>a</sup> nicht deklariert, sondern wurde dazuerfunden, um Bezeichner, wie PRINT, INPUT oder Programmname schreiben zu können.
Reg(reg)	Container für ein <b>Register</b> .
Im(val)	Ein <b>Immediate Value</b> , z.B. 42, -3 usw.
Add(), Sub(), Mult(), Div(), Mod(), Oplus(), Or(), And()	<b>Compute-Memory</b> oder <b>Compute-Register</b> Operationen: ADD, SUB, MULT, DIV, OPLUS, OR, AND.
Addi(), Subi(), Multi(), Divi(), Modi(), Oplusi(), Ori(), Andi()	<b>Compute-Immediate</b> Operationen: ADDI, SUBI, MULTI, DIVI, MODI, OPLUSI, ORI, ANDI.
Load(), Loadin(), Loadi()	<b>Load</b> Operationen: LOAD, LOADIN, LOADI.
Store(), Storein(), Move()	<b>Store</b> Operationen: STORE, STOREIN, MOVE.
Lt(), LtE(), Gt(), GtE(), Eq(), NEq(), Always(), NOP()	<b>Relationen</b> : <, <=, >, >=, ==, !=, _NOP.
Rti()	<b>Return-From-Interrupt</b> Operation: RTI.
Pc(), In1(), In2(), Acc(), Sp(), Baf(), Cs(), Ds()	<b>Register</b> : PC, IN1, IN2, ACC, SP, BAF, CS, DS.

<sup>a</sup> C. Scholl, „Betriebssysteme“

Tabelle 3.6: RETI-Knoten

**3.2.5.3 Kompositionen von PicoC-Knoten und RETI-Knoten mit besonderer Bedeutung**

Hier sind jegliche **Kompositionen** von **PicoC-Knoten** und **RETI-Knoten** aufgelistet, die eine **besondere Bedeutung** haben und nicht bereits in der **Abstrakten Syntax 3.2.1** enthalten sind.

Komposition	Beschreibung
Ref(Global(Num('addr')))	Speichert <b>Adresse</b> der Speicherzelle, die Num('addr') Speicherzellen <b>relativ</b> zum <b>Datensegment Register</b> DS steht auf den <b>Stack</b> .
Ref(Stackframe(Num('addr')))	Speichert <b>Adresse</b> der Speicherzelle, die Num('addr') Speicherzellen <b>relativ</b> zum <b>Begin-Aktive-Funktion Register</b> BAF steht auf den <b>Stack</b> .
Ref(Subscr(Stack(Num('addr1')), Stack(Num('addr2'))))	Berechnet die nächste <b>Adresse</b> aus der <b>Adresse</b> , die an Speicherzelle Stack(Num('addr1')) steht und dem <b>Subscript Index</b> , der an Speicherzelle Stack(Num('addr2')) steht und speichert diese auf den <b>Stack</b> . Die Berechnung ist abhängig davon ob der <b>Datentyp</b> ArrayDecl(datatype) oder PntrDecl(datatype) ist. Der <b>Datentyp</b> ist ein verstecktes Attribut von Ref(exp).
Ref(Attr(Stack(Num('addr1')), Name('attr')))	Berechnet die nächste <b>Adresse</b> aus der <b>Adresse</b> , die an Speicherzelle Stack(Num('addr1')) steht und dem <b>Attributnamen</b> Name('attr') und speichert diese auf den <b>Stack</b> . Zur Berechnung ist der Name des <b>Struct</b> in StructSpec(Name('st')) notwendig, dessen <b>Attribut</b> Name('attr') ist. StructSpec(Name('st')) ist ein verstecktes Attribut von Ref(exp).
Assign(Stack(Num('size')), Global(Num('addr')))	Schreibt Num('size') viele Speicherzellen, die ab Global(Num('addr')) relativ zum <b>Datensegment Register</b> DS stehen, versetzt genauso auf den <b>Stack</b> .
Assign(Stack(Num('size')), Stackframe(Num('addr')))	Schreibt Num('size') viele Speicherzellen, die ab Stackframe(Num('addr')) relativ zum <b>Begin-Aktive-Funktion Register</b> BAF stehen, versetzt genauso auf den <b>Stack</b> .
Exp(Global(Num('addr')))	Speichert <b>Inhalt</b> der Speicherzelle, die Num('addr') Speicherzellen <b>relativ</b> zum <b>Datensegment Register</b> DS steht auf den <b>Stack</b> .
Exp(Stackframe(Num('addr')))	Speichert <b>Inhalt</b> der Speicherzelle, die Num('addr') Speicherzellen <b>relativ</b> zum <b>Begin-Aktive-Funktion Register</b> BAF steht auf den <b>Stack</b> .
Exp(Stack(Num('addr')))	Speichert <b>Inhalt</b> der Speicherzelle, die Num('addr') Speicherzellen <b>relativ</b> zum <b>Stackpointer Register</b> SP steht auf den <b>Stack</b> .
Assign(Stack(Num('addr1')), Stack(Num('addr2')))	Speichert <b>Inhalt</b> der Speicherzelle Stack(Num('addr2')), die Num('addr2') Speicherzellen <b>relativ</b> zum <b>Stackpointer Register</b> SP steht an der <b>Adresse</b> in der Speicherzelle, die Num('addr1') Speicherzellen <b>relativ</b> zum <b>Stackpointer Register</b> SP steht.
Assign(Global(Num('addr')), Stack(Num('size')))	Schreibt Num('size') viele Speicherzellen, die auf dem <b>Stack</b> stehen, versetzt genauso auf die Speicherzellen ab Num('addr') <b>relativ</b> zum <b>Datensegment Register</b> DS.
Assign(Stackframe(Num('addr')), Stack(Num('size')))	Schreibt Num('size') viele Speicherzellen, die auf dem <b>Stack</b> stehen, versetzt genauso auf die Speicherzellen ab Num('addr') <b>relativ</b> zum <b>Begin-Aktive-Funktion Register</b> BAF.
Exp(Reg(reg))	Schreibt den aktuellen Wert des <b>Registers</b> reg auf den <b>Stack</b> .
Instr(Loadi(), [Reg(Acc()), GoTo(Name('addr@next_instr'))])	Lädt in das Register ACC die <b>Adresse</b> der Instruction, die in diesem Kontext direkt nach dem Sprung zum Block einer anderen Funktion steht.

Tabelle 3.7: Kompositionen von PicoC-Knoten und RETI-Knoten mit besonderer Bedeutung

Um die obige Tabelle 3.7 nicht mit unnötig viel repetitiven Inhalt zu füllen, wurden die zahlreichen Kompositionen **ausgelassen**, bei denen einfach nur **exp** durch  $\text{Stack}(\text{Num}('x')), x \in \mathbb{N}$  ersetzt wurde.

Zudem sind auch jegliche Kombinationen ausgelassen, bei denen einfach nur eine **Expression** an ein  $\text{Exp}(\text{exp})$  bzw.  $\text{Ref}(\text{exp})$  drangehängt wurde.

#### 3.2.5.4 Abstrakte Syntax

<i>un_op</i>	::=	<i>Minus()</i>   <i>Not()</i>	<i>L_Arith</i>
<i>bin_op</i>	::=	<i>Add()</i>   <i>Sub()</i>   <i>Mul()</i>   <i>Div()</i>   <i>Mod()</i>   <i>Oplus()</i>   <i>And()</i>   <i>Or()</i>	
<i>exp</i>	::=	<i>Name(str)</i>   <i>Num(str)</i>   <i>Char(str)</i>   <i>BinOp</i> ( <i>&lt;exp&gt;</i> , <i>&lt;bin_op&gt;</i> , <i>&lt;exp&gt;</i> )   <i>UnOp</i> ( <i>&lt;un_op&gt;</i> , <i>&lt;exp&gt;</i> )   <i>Call</i> ( <i>Name('input')</i> , <i>None</i> )	
<i>exp_stmts</i>	::=	<i>Alloc</i> ( <i>&lt;type_qual&gt;</i> , <i>&lt;datatype&gt;</i> , <i>Name(str)</i> )   <i>Call</i> ( <i>Name('print')</i> , <i>&lt;exp&gt;</i> )	
<i>un_op</i>	::=	<i>LogicNot()</i>	<i>L_Logic</i>
<i>rel</i>	::=	<i>Eq()</i>   <i>NEq()</i>   <i>Lt()</i>   <i>LtE()</i>   <i>Gt()</i>   <i>GtE()</i>	
<i>bin_op</i>	::=	<i>LogicAnd()</i>   <i>LogicOr()</i>	
<i>exp</i>	::=	<i>Atom</i> ( <i>&lt;exp&gt;</i> , <i>&lt;rel&gt;</i> , <i>&lt;exp&gt;</i> )   <i>ToBool</i> ( <i>&lt;exp&gt;</i> )	
<i>type_qual</i>	::=	<i>Const()</i>   <i>Writeable()</i>	<i>L_Assign_Alloc</i>
<i>datatype</i>	::=	<i>IntType()</i>   <i>CharType()</i>   <i>VoidType()</i>	
<i>lhs</i>	::=	<i>Alloc</i> ( <i>&lt;type_qual&gt;</i> , <i>&lt;datatype&gt;</i> , <i>Name(str)</i> )   <i>&lt;ref_loc&gt;</i>	
<i>exp_stmts</i>	::=	<i>Alloc</i> ( <i>&lt;type_qual&gt;</i> , <i>&lt;datatype&gt;</i> , <i>Name(str)</i> )	
<i>stmt</i>	::=	<i>Assign</i> ( <i>&lt;lhs&gt;</i> , <i>&lt;exp&gt;</i> )   <i>Exp</i> ( <i>&lt;exp_stmts&gt;</i> )	
<i>datatype</i>	::=	<i>PntrDecl</i> ( <i>Num(str)</i> , <i>&lt;datatype&gt;</i> )	<i>L_Pntr</i>
<i>deref_loc</i>	::=	<i>Ref</i> ( <i>&lt;ref_loc&gt;</i> )   <i>&lt;ref_loc&gt;</i>	
<i>ref_loc</i>	::=	<i>Name(str)</i>   <i>Deref</i> ( <i>&lt;deref_loc&gt;</i> , <i>&lt;exp&gt;</i> )   <i>Subscr</i> ( <i>&lt;deref_loc&gt;</i> , <i>&lt;exp&gt;</i> )   <i>Attr</i> ( <i>&lt;ref_loc&gt;</i> , <i>Name(str)</i> )	
<i>exp</i>	::=	<i>Deref</i> ( <i>&lt;deref_loc&gt;</i> , <i>&lt;exp&gt;</i> )   <i>Ref</i> ( <i>&lt;ref_loc&gt;</i> )	
<i>datatype</i>	::=	<i>ArrayDecl</i> ( <i>Num(str)</i> +, <i>&lt;datatype&gt;</i> )	<i>L_Array</i>
<i>exp</i>	::=	<i>Subscr</i> ( <i>&lt;deref_loc&gt;</i> , <i>&lt;exp&gt;</i> )   <i>Array</i> ( <i>&lt;exp&gt;</i> +)	
<i>datatype</i>	::=	<i>StructSpec</i> ( <i>Name(str)</i> )	<i>L_Struct</i>
<i>exp</i>	::=	<i>Attr</i> ( <i>&lt;ref_loc&gt;</i> , <i>Name(str)</i> )   <i>Struct</i> ( <i>Assign</i> ( <i>Name(str)</i> , <i>&lt;exp&gt;</i> ) +)	
<i>decl_def</i>	::=	<i>StructDecl</i> ( <i>Name(str)</i> , <i>Alloc</i> ( <i>Writeable()</i> , <i>&lt;datatype&gt;</i> , <i>Name(str)</i> ) +)	
<i>stmt</i>	::=	<i>If</i> ( <i>&lt;exp&gt;</i> , <i>&lt;stmt&gt;</i> *)   <i>IfElse</i> ( <i>&lt;exp&gt;</i> , <i>&lt;stmt&gt;</i> *, <i>&lt;stmt&gt;</i> *)	<i>L_If_Else</i>
<i>stmt</i>	::=	<i>While</i> ( <i>&lt;exp&gt;</i> , <i>&lt;stmt&gt;</i> *)   <i>DoWhile</i> ( <i>&lt;exp&gt;</i> , <i>&lt;stmt&gt;</i> *)	<i>L_Loop</i>
<i>exp</i>	::=	<i>Call</i> ( <i>Name(str)</i> , <i>&lt;exp&gt;</i> *)	<i>L_Fun</i>
<i>exp_stmts</i>	::=	<i>Call</i> ( <i>Name(str)</i> , <i>&lt;exp&gt;</i> *)	
<i>stmt</i>	::=	<i>Return</i> ( <i>&lt;exp&gt;</i> )	
<i>decl_def</i>	::=	<i>FunDecl</i> ( <i>&lt;datatype&gt;</i> , <i>Name(str)</i> , <i>Alloc</i> ( <i>Writeable()</i> , <i>&lt;datatype&gt;</i> , <i>Name(str)</i> )*)   <i>FunDef</i> ( <i>&lt;datatype&gt;</i> , <i>Name(str)</i> , <i>Alloc</i> ( <i>Writeable()</i> , <i>&lt;datatype&gt;</i> , <i>Name(str)</i> )*, <i>&lt;stmt&gt;</i> *)	
<i>file</i>	::=	<i>File</i> ( <i>Name(str)</i> , <i>&lt;decl_def&gt;</i> *)	<i>L_File</i>

Grammar 3.2.3: Abstrakte Syntax für *L\_Piocc*



### 3.2.5.5 Transformer

### 3.2.5.6 Codebeispiel

Beispiel welches in Subkapitel 3.2.3.2 angefangen wurde, wird hier fortgeführt.

```
1 File
2   Name './example_dt_simple_ast_gen_array_decl_and_alloc.ast',
3   [
4     StructDecl
5       Name 'st',
6       [
7         Alloc
8           Writeable,
9           PtrDecl
10            Num '1',
11            ArrayDecl
12              [
13                Num '5',
14                Num '6'
15              ],
16            PtrDecl
17              Num '1',
18              IntType 'int',
19            Name 'attr'
20          ],
21      FunDef
22        VoidType 'void',
23        Name 'main',
24        [],
25        [
26          Exp
27            Alloc
28              Writeable,
29              PtrDecl
30                Num '1',
31                ArrayDecl
32                  [
33                    Num '3',
34                    Num '2'
35                  ],
36                PtrDecl
37                  Num '1',
38                  StructSpec
39                    Name 'st',
40                  Name 'var'
41                ]
42          ]
43        ]
44      ]
45    ]
46  ]
```

Code 3.4: Abstract Syntax Tree aus vereinfachtem Derivation Tree generiert

### 3.3 Code Generierung

#### 3.3.1 Übersicht

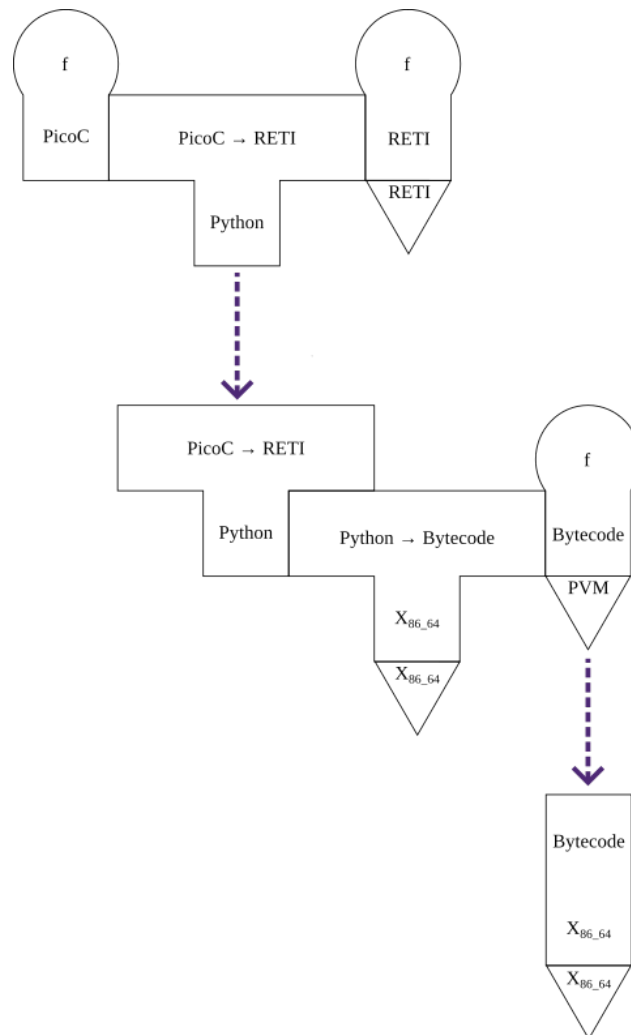


Abbildung 3.1: Cross-Compiler Kompiliervorgang ausgeschrieben

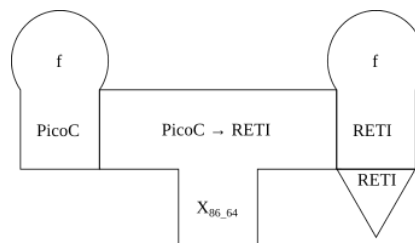


Abbildung 3.2: Cross-Compiler Kompiliervorgang Kurzform

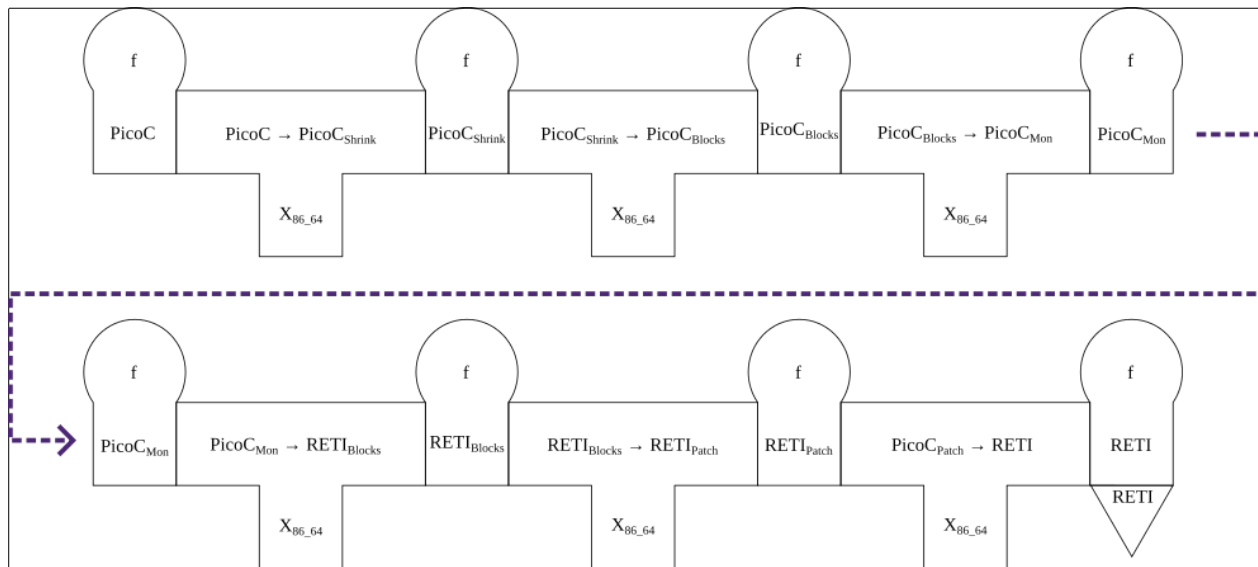


Abbildung 3.3: Architektur mit allen Passes ausgeschrieben

## 3.3.2 Passes

### 3.3.2.1 PicoC-Shrink Pass

#### 3.3.2.1.1 Codebeispiel

```

1 // Author: Christoph Scholl, from the Operating Systems Lecture
2
3 void main() {
4     int n = 4;
5     int res = 1;
6     while (1) {
7         if (n == 1) {
8             return;
9         }
10        res = n * res;
11        n = n - 1;
12    }
13 }

```

Code 3.5: PicoC Code für Codebeispiel

```

1 File
2   Name './example_faculty_it.ast',
3   [
4     FunDef
5       VoidType 'void',
6       Name 'main',
7       [],
8       [

```

```

9      Assign
10      Alloc
11      Writeable,
12      IntType 'int',
13      Name 'n',
14      Num '4',
15      Assign
16      Alloc
17      Writeable,
18      IntType 'int',
19      Name 'res',
20      Num '1',
21      While
22      Num '1',
23      [
24      If
25      Atom
26      Name 'n',
27      Eq '==',
28      Num '1',
29      [
30      Return
31      Empty
32      ],
33      Assign
34      Name 'res',
35      BinOp
36      Name 'n',
37      Mul '*',
38      Name 'res',
39      Assign
40      Name 'n',
41      BinOp
42      Name 'n',
43      Sub '-',
44      Num '1'
45      ]
46  ]
47  ]

```

Code 3.6: Abstract Syntax Tree für Codebeispiel

```

1 File
2   Name './example_faculty_it.picoc_shrink',
3   [
4     FunDef
5     VoidType 'void',
6     Name 'main',
7     [],
8     [
9       Assign
10      Alloc
11      Writeable,
12      IntType 'int',

```

```

13     Name 'n',
14     Num '4',
15   Assign
16     Alloc
17       Writeable,
18       IntType 'int',
19       Name 'res',
20     Num '1',
21   While
22     Num '1',
23   [
24     If
25       Atom
26         Name 'n',
27         Eq '==',
28         Num '1',
29       [
30         Return
31         Empty
32       ],
33     Assign
34       Name 'res',
35       BinOp
36         Name 'n',
37         Mul '*',
38         Name 'res',
39     Assign
40       Name 'n',
41       BinOp
42         Name 'n',
43         Sub '-',
44         Num '1'
45   ]
46 ]
47 ]

```

Code 3.7: PicoC Shrink Pass für Codebeispiel

### 3.3.2.2 PicoC-Blocks Pass

#### 3.3.2.2.1 Abstrakte Syntax

<i>decl_def</i>	$::=$	<i>FunDef</i> ( $\langle datatype \rangle$ , <i>Name</i> ( <i>str</i> ), <i>Alloc</i> ( <i>Writeable</i> ()), $\langle datatype \rangle$ , <i>Name</i> ( <i>str</i> ))* , $\langle block \rangle$ *)	<i>L_Fun</i>
<i>block</i>	$::=$	<i>Block</i> ( <i>Name</i> ( <i>str</i> ), $\langle stmt \rangle$ *)	<i>L_Blocks</i>
<i>stmt</i>	$::=$	<i>GoTo</i> ( <i>Name</i> ( <i>str</i> ))   <i>NewStackframe</i> ( <i>Name</i> ()), <i>GoTo</i> ( <i>str</i> )   <i>RemoveStackframe</i> ()   <i>SetScope</i> ( <i>Name</i> ( <i>str</i> ))   <i>SingleLineComment</i> ( <i>str</i> , <i>str</i> )	

Grammar 3.3.1: Abstrakte Syntax für  $L_{PicoC\_Blocks}$ 

#### 3.3.2.2.2 Codebeispiel

```

1 File
2   Name './example_faculty_it.picoc_blocks',
3   [
4     FunDef
5       VoidType 'void',
6       Name 'main',
7       [],
8       [
9         Block
10          Name 'main.5',
11          [
12            Assign
13              Alloc
14                Writeable,
15                IntType 'int',
16                Name 'n',
17                Num '4',
18            Assign
19              Alloc
20                Writeable,
21                IntType 'int',
22                Name 'res',
23                Num '1',
24            // While(Num('1'), []),
25            GoTo
26              Name 'condition_check.4'
27          ],
28          Block
29            Name 'condition_check.4',
30            [
31              IfElse
32                Num '1',
33                GoTo
34                  Name 'while_branch.3',
35                GoTo
36                  Name 'while_after.0'
37            ],
38          Block
39            Name 'while_branch.3',
40            [
41              // If(Atom(Name('n'), Eq('=='), Num('1')), []),
42              IfElse
43                Atom
44                  Name 'n',
45                  Eq '==',
46                  Num '1',
47                GoTo
48                  Name 'if.2',
49                GoTo
50                  Name 'if_else_after.1'
51            ],
52          Block
53            Name 'if.2',
54            [
55              Return
56                Empty
57            ],

```

```

58     Block
59     Name 'if_else_after.1',
60     [
61         Assign
62         Name 'res',
63         BinOp
64         Name 'n',
65         Mul '*',
66         Name 'res',
67         Assign
68         Name 'n',
69         BinOp
70         Name 'n',
71         Sub '-',
72         Num '1',
73         GoTo
74         Name 'condition_check.4'
75     ],
76     Block
77     Name 'while_after.0',
78     []
79 ]
80 ]

```

Code 3.8: PicoC-Blocks Pass für Codebeispiel

### 3.3.2.3 PicoC-Mon Pass

#### 3.3.2.3.1 Abstrakte Syntax

<i>ref_loc</i>	$::=$	$Stack(Num(str)) \mid Global(Num(str))$	$L\_Assign\_Alloc$
		$\mid Stackframe(Num(str))$	
<i>error_data</i>	$::=$	$\langle exp \rangle \mid Pos(Num(str), Num(str))$	
<i>exp</i>	$::=$	$Stack(Num(str)) \mid Ref(\langle ref\_loc \rangle, \langle datatype \rangle, \langle error\_data \rangle)$	
<i>stmt</i>	$::=$	$Exp(\langle exp \rangle)$	
		$\mid Assign(Alloc(Writable(), StructSpec(Name(str)), Name(str)),$	
		$Struct(Assign(Name(str), \langle exp \rangle +, \langle datatype \rangle))$	
		$\mid Assign(Alloc(Writable(), ArrayDecl(Num(str) +, \langle datatype \rangle),$	
		$Name(str), Array(\langle exp \rangle +, \langle datatype \rangle))$	
<i>symbol_table</i>	$::=$	$SymbolTable(\langle symbol \rangle)$	$L\_Symbol\_Table$
<i>symbol</i>	$::=$	$Symbol(\langle type\_qual \rangle, \langle datatype \rangle, \langle name \rangle, \langle val \rangle, \langle pos \rangle, \langle size \rangle)$	
<i>type_qual</i>	$::=$	$Empty()$	
<i>datatype</i>	$::=$	$BuiltIn() \mid SelfDefined()$	
<i>name</i>	$::=$	$Name(str)$	
<i>val</i>	$::=$	$Num(str) \mid Empty()$	
<i>pos</i>	$::=$	$Pos(Num(str), Num(str)) \mid Empty()$	
<i>size</i>	$::=$	$Num(str) \mid Empty()$	

Grammar 3.3.2: Abstrakte Syntax für  $L_{PicoC\_Mon}$ 

#### Definition 3.3: Symboltabelle

**3.3.2.3.2 Codebeispiel**

```

1 File
2   Name './example_faculty_it.picoc_mon',
3   [
4     Block
5       Name 'main.5',
6       [
7         // Assign(Name('n'), Num('4')),
8         Exp
9           Num '4',
10        Assign
11          GlobalWrite
12            Num '0',
13          Tmp
14            Num '1',
15          // Assign(Name('res'), Num('1')),
16        Exp
17          Num '1',
18        Assign
19          GlobalWrite
20            Num '1',
21          Tmp
22            Num '1',
23          // While(Num('1'), []),
24        Exp
25          GoTo
26            Name 'condition_check.4'
27      ],
28    Block
29      Name 'condition_check.4',
30      [
31        // IfElse(Num('1'), GoTo(Name('while_branch.3')), GoTo(Name('while_after.0'))),
32        Exp
33          Num '1',
34        IfElse
35          Tmp
36            Num '1',
37          GoTo
38            Name 'while_branch.3',
39          GoTo
40            Name 'while_after.0'
41      ],
42    Block
43      Name 'while_branch.3',
44      [
45        // If(Atom(Name('n'), Eq('=='), Num('1')), []),
46        // IfElse(Atom(Name('n'), Eq('=='), Num('1')), GoTo(Name('if.2')),
47        ↪ GoTo(Name('if_else_after.1'))),
48        Exp
49          GlobalRead
50            Num '0',
51        Exp
52          Num '1',
53        Exp
54          Atom
55            Tmp

```



```

55         Num '2',
56         Eq '==',
57         Tmp
58         Num '1',
59     IfElse
60     Tmp
61     Num '1',
62     GoTo
63     Name 'if.2',
64     GoTo
65     Name 'if_else_after.1'
66 ],
67 Block
68     Name 'if.2',
69     [
70         Return
71         Empty
72     ],
73 Block
74     Name 'if_else_after.1',
75     [
76         // Assign(Name('res'), BinOp(Name('n'), Mul('*'), Name('res'))),
77         Exp
78         GlobalRead
79         Num '0',
80         Exp
81         GlobalRead
82         Num '1',
83         Exp
84         BinOp
85         Tmp
86         Num '2',
87         Mul '*',
88         Tmp
89         Num '1',
90         Assign
91         GlobalWrite
92         Num '1',
93         Tmp
94         Num '1',
95         // Assign(Name('n'), BinOp(Name('n'), Sub('-'), Num('1'))),
96         Exp
97         GlobalRead
98         Num '0',
99         Exp
100        Num '1',
101        Exp
102        BinOp
103        Tmp
104        Num '2',
105        Sub '-',
106        Tmp
107        Num '1',
108        Assign
109        GlobalWrite
110        Num '0',
111        Tmp

```

```

112     Num '1',
113     Exp
114     GoTo
115     Name 'condition_check.4'
116 ],
117 Block
118   Name 'while_after.0',
119   [
120     Return
121     Empty
122   ]
123 ]

```

Code 3.9: PicoC-Mon Pass für Codebeispiel

### 3.3.2.4 RETI-Blocks Pass

#### 3.3.2.4.1 Abstrakte Syntax

<i>program</i>	$::=$	$Program(Name(str), \langle block \rangle^*)$	$L\_Program$
<i>exp_stmts</i>	$::=$	$GoTo(str)$	$L\_Blocks$
<i>instrs_before</i>	$::=$	$Num(str)$	
<i>num_instrs</i>	$::=$	$Num(str)$	
<i>block</i>	$::=$	$Block(Name(str), \langle instr \rangle^*, \langle instrs\_before \rangle, \langle num\_instrs \rangle)$	
<i>instr</i>	$::=$	$GoTo(Name(str))$	

Grammar 3.3.3: Abstrakte Syntax für  $L_{RETI\_Blocks}$ 

#### 3.3.2.4.2 Codebeispiel

```

1 File
2   Name './example_faculty_it.reti_blocks',
3   [
4     Block
5       Name 'main.5',
6       [
7         # // Assign(Name('n'), Num('4')),
8         # Exp(Num('4')),
9         SUBI SP 1,
10        LOADI ACC 4,
11        STOREIN SP ACC 1,
12        # Assign(GlobalWrite(Num('0')), Tmp(Num('1'))),
13        LOADIN SP ACC 1,
14        STOREIN DS ACC 0,
15        ADDI SP 1,
16        # // Assign(Name('res'), Num('1')),
17        # Exp(Num('1')),
18        SUBI SP 1,
19        LOADI ACC 1,
20        STOREIN SP ACC 1,
21        # Assign(GlobalWrite(Num('1')), Tmp(Num('1'))),
22        LOADIN SP ACC 1,

```

```

23     STOREIN DS ACC 1,
24     ADDI SP 1,
25     # // While(Num('1'), []),
26     Exp
27         GoTo
28             Name 'condition_check.4'
29 ],
30 Block
31     Name 'condition_check.4',
32     [
33         # // IfElse(Num('1'), GoTo(Name('while_branch.3')), GoTo(Name('while_after.0'))),
34         # Exp(Num('1')),
35         SUBI SP 1,
36         LOADI ACC 1,
37         STOREIN SP ACC 1,
38         # IfElse(Tmp(Num('1')), GoTo(Name('while_branch.3')), GoTo(Name('while_after.0'))),
39         LOADIN SP ACC 1,
40         ADDI SP 1,
41         JUMP== GoTo
42             Name 'while_after.0';,
43     Exp
44         GoTo
45             Name 'while_branch.3'
46 ],
47 Block
48     Name 'while_branch.3',
49     [
50         # // If(Atom(Name('n'), Eq('=='), Num('1')), []),
51         # // IfElse(Atom(Name('n'), Eq('=='), Num('1')), GoTo(Name('if.2')),
52         ↪ GoTo(Name('if_else_after.1'))),
53         # Exp(GlobalRead(Num('0'))),
54         SUBI SP 1,
55         LOADIN DS ACC 0,
56         STOREIN SP ACC 1,
57         # Exp(Num('1')),
58         SUBI SP 1,
59         LOADI ACC 1,
60         STOREIN SP ACC 1,
61         LOADIN SP ACC 2,
62         LOADIN SP IN2 1,
63         SUB ACC IN2,
64         JUMP== 3;,
65         LOADI ACC 0,
66         JUMP 2;,
67         LOADI ACC 1,
68         STOREIN SP ACC 2,
69         ADDI SP 1,
70         # IfElse(Tmp(Num('1')), GoTo(Name('if.2')), GoTo(Name('if_else_after.1'))),
71         LOADIN SP ACC 1,
72         ADDI SP 1,
73         JUMP== GoTo
74             Name 'if_else_after.1';,
75     Exp
76         GoTo
77             Name 'if.2'
78 ],
79 Block

```

```

79     Name 'if.2',
80     [
81         # Return(Empty()),
82         LOADIN BAF PC -1
83     ],
84     Block
85     Name 'if_else_after.1',
86     [
87         # // Assign(Name('res'), BinOp(Name('n'), Mul('*'), Name('res'))),
88         # Exp(GlobalRead(Num('0'))),
89         SUBI SP 1,
90         LOADIN DS ACC 0,
91         STOREIN SP ACC 1,
92         # Exp(GlobalRead(Num('1'))),
93         SUBI SP 1,
94         LOADIN DS ACC 1,
95         STOREIN SP ACC 1,
96         LOADIN SP ACC 2,
97         LOADIN SP IN2 1,
98         MULT ACC IN2,
99         STOREIN SP ACC 2,
100        ADDI SP 1,
101        # Assign(GlobalWrite(Num('1')), Tmp(Num('1'))),
102        LOADIN SP ACC 1,
103        STOREIN DS ACC 1,
104        ADDI SP 1,
105        # // Assign(Name('n'), BinOp(Name('n'), Sub('-'), Num('1'))),
106        # Exp(GlobalRead(Num('0'))),
107        SUBI SP 1,
108        LOADIN DS ACC 0,
109        STOREIN SP ACC 1,
110        # Exp(Num('1')),
111        SUBI SP 1,
112        LOADI ACC 1,
113        STOREIN SP ACC 1,
114        LOADIN SP ACC 2,
115        LOADIN SP IN2 1,
116        SUB ACC IN2,
117        STOREIN SP ACC 2,
118        ADDI SP 1,
119        # Assign(GlobalWrite(Num('0')), Tmp(Num('1'))),
120        LOADIN SP ACC 1,
121        STOREIN DS ACC 0,
122        ADDI SP 1,
123        Exp
124        GoTo
125        Name 'condition_check.4'
126    ],
127    Block
128    Name 'while_after.0',
129    [
130        # Return(Empty()),
131        LOADIN BAF PC -1
132    ]
133 ]

```

Code 3.10: RETI-Blocks Pass für Codebeispiel

### 3.3.2.5 RETI-Patch Pass

#### 3.3.2.5.1 Abstrakte Syntax

$$\text{stmt} ::= \text{Exit}(\text{Num}(\text{str}))$$

Grammar 3.3.4: Abstrakte Syntax für  $L_{\text{RETI\_Patch}}$

#### 3.3.2.5.2 Codebeispiel

```

1 File
2   Name './example_faculty_it.reti_patch',
3   [
4     Block
5       Name 'start.6',
6       [],
7     Block
8       Name 'main.5',
9       [
10        # // Assign(Name('n'), Num('4')),
11        # Exp(Num('4')),
12        SUBI SP 1,
13        LOADI ACC 4,
14        STOREIN SP ACC 1,
15        # Assign(GlobalWrite(Num('0')), Tmp(Num('1'))),
16        LOADIN SP ACC 1,
17        STOREIN DS ACC 0,
18        ADDI SP 1,
19        # // Assign(Name('res'), Num('1')),
20        # Exp(Num('1')),
21        SUBI SP 1,
22        LOADI ACC 1,
23        STOREIN SP ACC 1,
24        # Assign(GlobalWrite(Num('1')), Tmp(Num('1'))),
25        LOADIN SP ACC 1,
26        STOREIN DS ACC 1,
27        ADDI SP 1,
28        # // While(Num('1'), [])
29      ],
30    Block
31      Name 'condition_check.4',
32      [
33        # // IfElse(Num('1'), GoTo(Name('while_branch.3')), GoTo(Name('while_after.0'))),
34        # Exp(Num('1')),
35        SUBI SP 1,
36        LOADI ACC 1,
37        STOREIN SP ACC 1,
38        # IfElse(Tmp(Num('1')), GoTo(Name('while_branch.3')), GoTo(Name('while_after.0'))),
39        LOADIN SP ACC 1,
40        ADDI SP 1,
41        JUMP== GoTo
42          Name 'while_after.0';
43      ],
44    Block
45      Name 'while_branch.3',

```

```

46     [
47         # // If(Atom(Name('n'), Eq('=='), Num('1')), []),
48         # // IfElse(Atom(Name('n'), Eq('=='), Num('1')), GoTo(Name('if.2')),
49         ↪ GoTo(Name('if_else_after.1'))),
49         # Exp(GlobalRead(Num('0'))),
50         SUBI SP 1,
51         LOADIN DS ACC 0,
52         STOREIN SP ACC 1,
53         # Exp(Num('1')),
54         SUBI SP 1,
55         LOADI ACC 1,
56         STOREIN SP ACC 1,
57         LOADIN SP ACC 2,
58         LOADIN SP IN2 1,
59         SUB ACC IN2,
60         JUMP== 3;,
61         LOADI ACC 0,
62         JUMP 2;,
63         LOADI ACC 1,
64         STOREIN SP ACC 2,
65         ADDI SP 1,
66         # IfElse(Tmp(Num('1')), GoTo(Name('if.2')), GoTo(Name('if_else_after.1'))),
67         LOADIN SP ACC 1,
68         ADDI SP 1,
69         JUMP== GoTo
70             Name 'if_else_after.1';
71     ],
72     Block
73     Name 'if.2',
74     [
75         # Return(Empty()),
76         LOADIN BAF PC -1
77     ],
78     Block
79     Name 'if_else_after.1',
80     [
81         # // Assign(Name('res'), BinOp(Name('n'), Mul('*'), Name('res'))),
82         # Exp(GlobalRead(Num('0'))),
83         SUBI SP 1,
84         LOADIN DS ACC 0,
85         STOREIN SP ACC 1,
86         # Exp(GlobalRead(Num('1'))),
87         SUBI SP 1,
88         LOADIN DS ACC 1,
89         STOREIN SP ACC 1,
90         LOADIN SP ACC 2,
91         LOADIN SP IN2 1,
92         MULT ACC IN2,
93         STOREIN SP ACC 2,
94         ADDI SP 1,
95         # Assign(GlobalWrite(Num('1')), Tmp(Num('1'))),
96         LOADIN SP ACC 1,
97         STOREIN DS ACC 1,
98         ADDI SP 1,
99         # // Assign(Name('n'), BinOp(Name('n'), Sub('-'), Num('1'))),
100        # Exp(GlobalRead(Num('0'))),
101        SUBI SP 1,

```

```

102     LOADIN DS ACC 0,
103     STOREIN SP ACC 1,
104     # Exp(Num('1')),
105     SUBI SP 1,
106     LOADI ACC 1,
107     STOREIN SP ACC 1,
108     LOADIN SP ACC 2,
109     LOADIN SP IN2 1,
110     SUB ACC IN2,
111     STOREIN SP ACC 2,
112     ADDI SP 1,
113     # Assign(GlobalWrite(Num('0')), Tmp(Num('1'))),
114     LOADIN SP ACC 1,
115     STOREIN DS ACC 0,
116     ADDI SP 1,
117     Exp
118     GoTo
119     Name 'condition_check.4'
120 ],
121 Block
122     Name 'while_after.0',
123     [
124         # Return(Empty()),
125         LOADIN BAF PC -1
126     ]
127 ]

```

Code 3.11: RETI-Patch Pass für Codebeispiel

### 3.3.2.6 RETI Pass

#### 3.3.2.6.1 Konkrete und Abstrakte Syntax

<i>dig_no_0</i>	::=	"1"   "2"   "3"   "4"   "5"   "6"	<i>L_Program</i>
		"7"   "8"   "9"	
<i>dig_with_0</i>	::=	"0"   <i>dig_no_0</i>	
<i>num</i>	::=	"0"   <i>dig_no_0</i> <i>dig_with_0</i> *   "-" <i>dig_with_0</i> *	
<i>letter</i>	::=	"a"..."Z"	
<i>name</i>	::=	<i>letter</i> ( <i>letter</i>   <i>dig_with_0</i>   _)*	
<i>reg</i>	::=	"ACC"   "IN1"   "IN2"   "PC"   "SP"	
		"BAF"   "CS"   "DS"	
<i>arg</i>	::=	<i>reg</i>   <i>num</i>	
<i>rel</i>	::=	"=="   "!="   "<"   "<="   ">"	
		">="   "_NOP"	

Grammar 3.3.5: Konkrete Syntax für  $L_{RETI\_Lex}$

<i>instr</i>	::=	"ADD" reg arg   "ADDI" reg num   "SUB" reg arg	<i>L_Program</i>
		"SUBI" reg num   "MULT" reg arg   "MULTI" reg num	
		"DIV" reg arg   "DIVI" reg num   "MOD" reg arg	
		"MODI" reg num   "OPLUS" reg arg   "OPLUSI" reg num	
		"OR" reg arg   "ORI" reg num	
		"AND" reg arg   "ANDI" reg num	
		"LOAD" reg num   "LOADIN" arg arg num	
		"LOADI" reg num	
		"STORE" reg num   "STOREIN" arg argnum	
		"MOVE" reg reg	
		"JUMP" rel num   INT num   RTI	
		"CALL" "INPUT" reg   "CALL" "PRINT" reg	
<i>program</i>	::=	name (instr";")*	

Grammar 3.3.6: Konkrete Syntax für  $L_{RETI\_Parse}$

<i>reg</i>	::=	<i>ACC()</i>   <i>IN1()</i>   <i>IN2()</i>   <i>PC()</i>   <i>SP()</i>   <i>BAF()</i>	<i>L_Program</i>
		<i>CS()</i>   <i>DS()</i>	
<i>arg</i>	::=	<i>Reg</i> ( <i>&lt;reg&gt;</i> )   <i>Num</i> ( <i>str</i> )	
<i>rel</i>	::=	<i>Eq()</i>   <i>NEq()</i>   <i>Lt()</i>   <i>LtE()</i>   <i>Gt()</i>   <i>GtE()</i>	
		<i>Always()</i>   <i>NOp()</i>	
<i>op</i>	::=	<i>Add()</i>   <i>Addi()</i>   <i>Sub()</i>   <i>Subi()</i>   <i>Mult()</i>	
		<i>Multi()</i>   <i>Div()</i>   <i>Divi()</i>	
		<i>Mod()</i>   <i>Modi()</i>   <i>Oplus()</i>   <i>Oplusi()</i>   <i>Or()</i>	
		<i>Ori()</i>   <i>And()</i>   <i>Andi()</i>	
		<i>Load()</i>   <i>Loadin()</i>   <i>Loadi()</i>	
		<i>Store()</i>   <i>Storein()</i>   <i>Move()</i>	
<i>instr</i>	::=	<i>Instr</i> ( <i>&lt;op&gt;</i> , <i>&lt;arg&gt;</i> +)   <i>Jump</i> ( <i>&lt;rel&gt;</i> , <i>Num</i> ( <i>str</i> ))   <i>Int</i> ( <i>Num</i> ( <i>str</i> ))	
		<i>RTI()</i>   <i>Call</i> ( <i>Name</i> ('print'), <i>&lt;reg&gt;</i> )   <i>Call</i> ( <i>Name</i> ('input'), <i>&lt;reg&gt;</i> )	
		<i>SingleLineComment</i> ( <i>str</i> , <i>str</i> )	
<i>program</i>	::=	<i>Program</i> ( <i>Name</i> ( <i>str</i> ), <i>&lt;instr&gt;</i> *)	

Grammar 3.3.7: Abstrakte Syntax für  $L_{RETI}$

### 3.3.2.6.2 Codebeispiel

```
1 # // Assign(Name('n'), Num('4'))
2 # Exp(Num('4'))
3 SUBI SP 1;
4 LOADI ACC 4;
5 STOREIN SP ACC 1;
6 # Assign(GlobalWrite(Num('0')), Tmp(Num('1')))
7 LOADIN SP ACC 1;
8 STOREIN DS ACC 0;
9 ADDI SP 1;
10 # // Assign(Name('res'), Num('1'))
11 # Exp(Num('1'))
12 SUBI SP 1;
13 LOADI ACC 1;
14 STOREIN SP ACC 1;
15 # Assign(GlobalWrite(Num('1')), Tmp(Num('1')))
```



```

16 LOADIN SP ACC 1;
17 STOREIN DS ACC 1;
18 ADDI SP 1;
19 # // While(Num('1'), [])
20 # // IfElse(Num('1'), GoTo(Name('while_branch.3')), GoTo(Name('while_after.0')))
21 # Exp(Num('1'))
22 SUBI SP 1;
23 LOADI ACC 1;
24 STOREIN SP ACC 1;
25 # IfElse(Tmp(Num('1')), GoTo(Name('while_branch.3')), GoTo(Name('while_after.0')))
26 LOADIN SP ACC 1;
27 ADDI SP 1;
28 JUMP== 49;
29 # // If(Atom(Name('n'), Eq('=='), Num('1')), [])
30 # // IfElse(Atom(Name('n'), Eq('=='), Num('1')), GoTo(Name('if.2')),
    ↪ GoTo(Name('if_else_after.1')))
31 # Exp(GlobalRead(Num('0')))
32 SUBI SP 1;
33 LOADIN DS ACC 0;
34 STOREIN SP ACC 1;
35 # Exp(Num('1'))
36 SUBI SP 1;
37 LOADI ACC 1;
38 STOREIN SP ACC 1;
39 LOADIN SP ACC 2;
40 LOADIN SP IN2 1;
41 SUB ACC IN2;
42 JUMP== 3;
43 LOADI ACC 0;
44 JUMP 2;
45 LOADI ACC 1;
46 STOREIN SP ACC 2;
47 ADDI SP 1;
48 # IfElse(Tmp(Num('1')), GoTo(Name('if.2')), GoTo(Name('if_else_after.1')))
49 LOADIN SP ACC 1;
50 ADDI SP 1;
51 JUMP== 2;
52 # Return(Empty())
53 LOADIN BAF PC -1;
54 # // Assign(Name('res'), BinOp(Name('n'), Mul('*'), Name('res')))
55 # Exp(GlobalRead(Num('0')))
56 SUBI SP 1;
57 LOADIN DS ACC 0;
58 STOREIN SP ACC 1;
59 # Exp(GlobalRead(Num('1')))
60 SUBI SP 1;
61 LOADIN DS ACC 1;
62 STOREIN SP ACC 1;
63 LOADIN SP ACC 2;
64 LOADIN SP IN2 1;
65 MULT ACC IN2;
66 STOREIN SP ACC 2;
67 ADDI SP 1;
68 # Assign(GlobalWrite(Num('1')), Tmp(Num('1')))
69 LOADIN SP ACC 1;
70 STOREIN DS ACC 1;
71 ADDI SP 1;

```

```
72 # // Assign(Name('n'), BinOp(Name('n'), Sub('-', Num('1'))))
73 # Exp(GlobalRead(Num('0'))))
74 SUBI SP 1;
75 LOADIN DS ACC 0;
76 STOREIN SP ACC 1;
77 # Exp(Num('1'))
78 SUBI SP 1;
79 LOADI ACC 1;
80 STOREIN SP ACC 1;
81 LOADIN SP ACC 2;
82 LOADIN SP IN2 1;
83 SUB ACC IN2;
84 STOREIN SP ACC 2;
85 ADDI SP 1;
86 # Assign(GlobalWrite(Num('0')), Tmp(Num('1'))))
87 LOADIN SP ACC 1;
88 STOREIN DS ACC 0;
89 ADDI SP 1;
90 JUMP -53;
91 # Return(Empty())
92 LOADIN BAF PC -1;
```

Code 3.12: RETI Pass für Codebeispiel

### 3.3.3 Umsetzung von Pointern

#### 3.3.3.1 Referenzierung

Die **Referenzierung** `&var` wird im Folgenden anhand des Beispiels Code 3.13 erklärt.

```

1 void main() {
2     int var = 42;
3     int *pntr = &var;
4 }
```

Code 3.13: PicoC Code für Pointer Referenzierung

Der Node `Ref(Name('var'))` repräsentiert in der **Abstrakten Syntax** in Code 3.14 eine **Referenzierung** `&var`.

```

1 File
2   Name './example_pntr_ref.ast',
3   [
4     FunDef
5       VoidType 'void',
6       Name 'main',
7       [],
8       [
9         Assign
10          Alloc
11            Writeable,
12            IntType 'int',
13            Name 'var',
14            Num '42',
15          Assign
16            Alloc
17              Writeable,
18              PntrDecl
19                Num '1',
20                IntType 'int',
21                Name 'pntr',
22              Ref
23                Name 'var'
24            ]
25          ]
26   ]
```

Code 3.14: Abstract Syntax Tree für Pointer Referenzierung

Im **PicoC-Mon Pass** in Code 3.15 wird der Node `Ref(Name('var'))` durch die Nodes `Ref(GlobalRead(Num('0')))` und `Assign(GlobalWrite(Num('1')), Tmp(Num('1')))` ersetzt. Im Fall, dass in `Ref(exp)` das `exp` vielleicht nicht direkt ein `Name('var')` enthält und `exp` vielleicht ein `Subscr(Attr(Name('var')))` ist, sind noch weitere Anweisungen zwischen Zeile 14 und 15 nötig, die sich in diesem Beispiel um das Übersetzen von `Subscr(exp)` und `Attr(exp)` kümmern. Die Vorgehen hierfür ist in Subkapitel 3.3.6.2 erklärt.

```

1 File
2   Name './example_pntr_ref.picoc_mon',
3   [
4     Block
5       Name 'main.O',
6       [
7         // Assign(Name('var'), Num('42')),
8         Exp
9           Num '42',
10        Assign
11          GlobalWrite
12            Num '0',
13          Tmp
14            Num '1',
15        // Assign(Name('pntr'), Ref(Name('var'))),
16        Ref
17          GlobalRead
18            Num '0',
19        Assign
20          GlobalWrite
21            Num '1',
22          Tmp
23            Num '1',
24        Return
25          Empty
26      ]
27  ]

```

Code 3.15: PicoC Mon Pass für Pointer Referenzierung

Im **PicoC-Blocks Pass** in Code 3.16 wird der die Nodes

```

1 File
2   Name './example_pntr_ref.reti_blocks',
3   [
4     Block
5       Name 'main.O',
6       [
7         # // Assign(Name('var'), Num('42')),
8         # Exp(Num('42')),
9         SUBI SP 1,
10        LOADI ACC 42,
11        STOREIN SP ACC 1,
12        # Assign(GlobalWrite(Num('0')), Tmp(Num('1'))),
13        LOADIN SP ACC 1,
14        STOREIN DS ACC 0,
15        ADDI SP 1,
16        # // Assign(Name('pntr'), Ref(Name('var'))),
17        # Ref(GlobalRead(Num('0'))),
18        SUBI SP 1,
19        LOADI IN1 0,
20        ADD IN1 DS,
21        STOREIN SP IN1 1,
22        # Assign(GlobalWrite(Num('1')), Tmp(Num('1'))),

```

```

23     LOADIN SP ACC 1,
24     STOREIN DS ACC 1,
25     ADDI SP 1,
26     # Return(Empty()),
27     LOADIN BAF PC -1
28 ]
29 ]

```

Code 3.16: RETI Blocks Pass für Pointer Referenzierung

### 3.3.3.2 Pointer Dereferenzierung durch Zugriff auf Arrayindex ersetzen

```

1 void main() {
2     int var = 42;
3     int *pntr = &var;
4     *pntr;
5 }

```

Code 3.17: PicoC Code für Pointer Dereferenzierung

```

1 File
2   Name './example_pntr_deref.ast',
3   [
4     FunDef
5       VoidType 'void',
6       Name 'main',
7       [],
8       [
9         Assign
10          Alloc
11            Writeable,
12            IntType 'int',
13            Name 'var',
14            Num '42',
15          Assign
16            Alloc
17              Writeable,
18              PntrDecl
19                Num '1',
20                IntType 'int',
21                Name 'pntr',
22              Ref
23                Name 'var',
24            Exp
25              Deref
26                Name 'pntr',
27                Num '0'
28          ]
29 ]

```

Code 3.18: Abstract Syntax Tree für Pointer Dereferenzierung

```
1 File
2   Name './example_pntr_deref.picoc_shrink',
3   [
4     FunDef
5       VoidType 'void',
6       Name 'main',
7       [],
8       [
9         Assign
10          Alloc
11            Writeable,
12            IntType 'int',
13            Name 'var',
14            Num '42',
15          Assign
16            Alloc
17              Writeable,
18              PntrDecl
19                Num '1',
20                IntType 'int',
21                Name 'pntr',
22              Ref
23                Name 'var',
24            Exp
25              Subscr
26                Name 'pntr',
27                Num '0'
28      ]
29   ]
```

Code 3.19: PicoC Shrink Pass für Pointer Dereferenzierung

### 3.3.4 Umsetzung von Arrays

#### 3.3.4.1 Initialisierung von Arrays

```
1 void main() {
2   int ar[2][1] = {{4}, {2}};
3 }
```

Code 3.20: PicoC Code für Array Initialisierung

```
1 File
2   Name './example_array_init.ast',
3   [
4     FunDef
5       VoidType 'void',
6       Name 'main',
7       [],
```

```

8      [
9      Assign
10     Alloc
11     Writeable,
12     ArrayDecl
13     [
14     Num '2',
15     Num '1'
16     ],
17     IntType 'int',
18     Name 'ar',
19     Array
20     [
21     Array
22     [
23     Num '4'
24     ],
25     Array
26     [
27     Num '2'
28     ]
29     ]
30 ]
31 ]

```

Code 3.21: Abstract Syntax Tree für Array Initialisierung

```

1 SymbolTable
2 [
3   Symbol(
4     {
5       type qualifier:      Empty()
6       datatype:            FunDecl(VoidType('void'), Name('main'), [])
7       name:                Name('main')
8       value or address:    Empty()
9       position:            Pos(Num('1'), Num('5'))
10      size:                 Empty()
11    },
12   Symbol(
13     {
14       type qualifier:      Writeable()
15       datatype:            ArrayDecl([Num('2'), Num('1')], IntType('int'))
16       name:                Name('ar@main')
17       value or address:    Num('0')
18       position:            Pos(Num('2'), Num('6'))
19       size:                 Num('2')
20     }
21 ]

```

Code 3.22: Symboltabelle für Array Initialisierung

```

1 File
2   Name './example_array_init.picoc_mon',
3   [
4     Block
5       Name 'main.0',
6       [
7         // Assign(Name('ar'), Array([Array([Num('4')]), Array([Num('2')])])),
8         Exp
9           Num '4',
10        Exp
11          Num '2',
12        Assign
13          GlobalWrite
14            Num '0',
15          Tmp
16            Num '2',
17        Return
18          Empty
19      ]
20  ]

```

Code 3.23: PicoC Mon Pass für Array Initialisierung

```

1 File
2   Name './example_array_init.reti_blocks',
3   [
4     Block
5       Name 'main.0',
6       [
7         # // Assign(Name('ar'), Array([Array([Num('4')]), Array([Num('2')])])),
8         # Exp(Num('4')),
9         SUBI SP 1,
10        LOADI ACC 4,
11        STOREIN SP ACC 1,
12        # Exp(Num('2')),
13        SUBI SP 1,
14        LOADI ACC 2,
15        STOREIN SP ACC 1,
16        # Assign(GlobalWrite(Num('0')), Tmp(Num('2'))),
17        LOADIN SP ACC 1,
18        STOREIN DS ACC 1,
19        LOADIN SP ACC 2,
20        STOREIN DS ACC 0,
21        ADDI SP 2,
22        # Return(Empty()),
23        LOADIN BAF PC -1
24      ]
25  ]

```

Code 3.24: RETI Blocks Pass für Array Initialisierung

### 3.3.4.2 Zugriff auf Arrayindex

Der Zugriff auf einen bestimmten Index eines Arrays ist wie folgt umgesetzt:



```
1 void main() {  
2   int ar[2] = {1, 2};  
3   ar[2];  
4 }
```

Code 3.25: PicoC Code für Zugriff auf Arrayindex

```
1 File  
2   Name './example_array_access.ast',  
3   [  
4     FunDef  
5       VoidType 'void',  
6       Name 'main',  
7       [],  
8       [  
9         Assign  
10          Alloc  
11            Writeable,  
12            ArrayDecl  
13              [  
14                Num '2'  
15              ],  
16              IntType 'int',  
17              Name 'ar',  
18              Array  
19                [  
20                  Num '1',  
21                  Num '2'  
22                ],  
23              Exp  
24                Subscr  
25                  Name 'ar',  
26                  Num '2'  
27              ]  
28    ]
```

Code 3.26: Abstract Syntax Tree für Zugriff auf Arrayindex

```
1 File  
2   Name './example_array_access.picoc_mon',  
3   [  
4     Block  
5       Name 'main.0',  
6       [  
7         // Assign(Name('ar'), Array([Num('1'), Num('2')])),  
8         Exp  
9           Num '1',  
10        Exp  
11          Num '2',
```

```

12      Assign
13      GlobalWrite
14      Num '0',
15      Tmp
16      Num '2',
17      // Exp(Subscr(Name('ar'), Num('2'))),
18      Ref
19      GlobalRead
20      Num '0',
21      Exp
22      Num '2',
23      Ref
24      Subscr
25      Tmp
26      Num '2',
27      Tmp
28      Num '1',
29      Exp
30      Subscr
31      Tmp
32      Num '1',
33      Num '0',
34      Return
35      Empty
36  ]
37 ]

```

Code 3.27: PicoC Mon Pass für Zugriff auf Arrayindex

```

1 File
2   Name './example_array_access.reti_blocks',
3   [
4     Block
5       Name 'main.0',
6       [
7         # // Assign(Name('ar'), Array([Num('1'), Num('2')])),
8         # Exp(Num('1')),
9         SUBI SP 1,
10        LOADI ACC 1,
11        STOREIN SP ACC 1,
12        # Exp(Num('2')),
13        SUBI SP 1,
14        LOADI ACC 2,
15        STOREIN SP ACC 1,
16        # Assign(GlobalWrite(Num('0')), Tmp(Num('2'))),
17        LOADIN SP ACC 1,
18        STOREIN DS ACC 1,
19        LOADIN SP ACC 2,
20        STOREIN DS ACC 0,
21        ADDI SP 2,
22        # // Exp(Subscr(Name('ar'), Num('2'))),
23        # Ref(GlobalRead(Num('0'))),
24        SUBI SP 1,
25        LOADI IN1 0,

```

```

26      ADD IN1 DS,
27      STOREIN SP IN1 1,
28      # Exp(Num('2')),
29      SUBI SP 1,
30      LOADI ACC 2,
31      STOREIN SP ACC 1,
32      # Ref(Subscr(Tmp(Num('2')), Tmp(Num('1')))),
33      LOADIN SP IN1 2,
34      LOADIN SP IN2 1,
35      MULTI IN2 1,
36      ADD IN1 IN2,
37      ADDI SP 1,
38      STOREIN SP IN1 1,
39      # Exp(Subscr(Tmp(Num('1')), Num('0'))),
40      LOADIN SP IN1 1,
41      LOADIN IN1 ACC 0,
42      STOREIN SP ACC 1,
43      # Return(Empty()),
44      LOADIN BAF PC -1
45  ]
46 ]

```

Code 3.28: RETI Blocks Pass für Zugriff auf Arrayindex

### 3.3.4.3 Zuweisung an Arrayindex

```

1 void main() {
2     int ar[2];
3     ar[2] = 42;
4 }

```

Code 3.29: PicoC Code für Zuweisung an Arrayindex

```

1 File
2   Name './example_array_assignment.ast',
3   [
4     FunDef
5       VoidType 'void',
6       Name 'main',
7       [],
8       [
9         Exp
10          Alloc
11          Writeable,
12          ArrayDecl
13            [
14              Num '2'
15            ],
16          IntType 'int',
17          Name 'ar',
18        Assign
19        Subscr

```

```

20         Name 'ar',
21         Num '2',
22         Num '42'
23     ]
24 ]

```

Code 3.30: Abstract Syntax Tree für Zuweisung an Arrayindex

```

1 File
2   Name './example_array_assignment.picoc_mon',
3   [
4     Block
5       Name 'main.0',
6       [
7         // Assign(Subscr(Name('ar'), Num('2')), Num('42')),
8         Exp
9           Num '42',
10        Ref
11          GlobalRead
12            Num '0',
13        Exp
14          Num '2',
15        Ref
16          Subscr
17            Tmp
18              Num '2',
19            Tmp
20              Num '1',
21        Assign
22          Subscr
23            Tmp
24              Num '1',
25            Num '0',
26          Tmp
27            Num '2',
28        Return
29          Empty
30      ]
31 ]

```

Code 3.31: PicoC Mon Pass für Zuweisung an Arrayindex

```

1 File
2   Name './example_array_assignment.reti_blocks',
3   [
4     Block
5       Name 'main.0',
6       [
7         # // Assign(Subscr(Name('ar'), Num('2')), Num('42')),
8         # Exp(Num('42')),
9         SUBI SP 1,

```

```

10     LOADI ACC 42,
11     STOREIN SP ACC 1,
12     # Ref(GlobalRead(Num('0'))),
13     SUBI SP 1,
14     LOADI IN1 0,
15     ADD IN1 DS,
16     STOREIN SP IN1 1,
17     # Exp(Num('2')),
18     SUBI SP 1,
19     LOADI ACC 2,
20     STOREIN SP ACC 1,
21     # Ref(Subscr(Tmp(Num('2')), Tmp(Num('1')))),
22     LOADIN SP IN1 2,
23     LOADIN SP IN2 1,
24     MULTI IN2 1,
25     ADD IN1 IN2,
26     ADDI SP 1,
27     STOREIN SP IN1 1,
28     # Assign(Subscr(Tmp(Num('1')), Num('0')), Tmp(Num('2'))),
29     LOADIN SP IN1 1,
30     LOADIN SP ACC 2,
31     ADDI SP 2,
32     STOREIN IN1 ACC 0,
33     # Return(Empty()),
34     LOADIN BAF PC -1
35 ]
36 ]

```

Code 3.32: RETI Blocks Pass für Zuweisung an Arrayindex

### 3.3.5 Umsetzung von Structs

#### 3.3.5.1 Deklaration von Structs

```

1 struct st1 {int *ar[3];};
2
3 struct st2 {struct st1 st;};
4
5 void main() {
6 }

```

Code 3.33: PicoC Code für Deklaration von Structs

```

1 SymbolTable
2 [
3     Symbol(
4     {
5         type qualifier:    Empty()
6         datatype:         ArrayDecl([Num('3')], PtrDecl(Num('1'), IntType('int')))
7         name:              Name('ar@st1')
8         value or address:  Empty()

```

```

9      position:      Pos(Num('1'), Num('17'))
10     size:          Num('3')
11   },
12   Symbol(
13   {
14     type qualifier:  Empty()
15     datatype:        StructDecl(Name('st1'), [Alloc(Writeable(),
16     ↪ ArrayDecl([Num('3')], PtrDecl(Num('1'), IntType('int'))), Name('ar'))])
17     name:            Name('st1')
18     value or address: [Name('ar@st1')]
19     position:        Pos(Num('1'), Num('7'))
20     size:            Num('3')
21   },
22   Symbol(
23   {
24     type qualifier:  Empty()
25     datatype:        StructSpec(Name('st1'))
26     name:            Name('st@st2')
27     value or address: Empty()
28     position:        Pos(Num('3'), Num('23'))
29     size:            Num('3')
30   },
31   Symbol(
32   {
33     type qualifier:  Empty()
34     datatype:        StructDecl(Name('st2'), [Alloc(Writeable(),
35     ↪ StructSpec(Name('st1')), Name('st'))])
36     name:            Name('st2')
37     value or address: [Name('st@st2')]
38     position:        Pos(Num('3'), Num('7'))
39     size:            Num('3')
40   },
41   Symbol(
42   {
43     type qualifier:  Empty()
44     datatype:        FunDecl(VoidType('void'), Name('main'), [])
45     name:            Name('main')
46     value or address: Empty()
47     position:        Pos(Num('5'), Num('5'))
48     size:            Empty()
49   }
50 ]

```

Code 3.34: Symboltabelle für Deklaration von Structs

### 3.3.5.2 Initialisierung von Structs

```

1 struct st1 {int *pntr[1];};
2
3 struct st2 {struct st1 st;};
4
5 void main() {
6     int var = 42;
7     struct st1 st = {.st={.pntr={{&var}}}};
8 }

```

Code 3.35: PicoC Code für Initialisierung von Structs

```
1 File
2   Name './example_struct_init.ast',
3   [
4     StructDecl
5       Name 'st1',
6       [
7         Alloc
8           Writeable,
9           ArrayDecl
10            [
11              Num '1'
12            ],
13            PtrDecl
14              Num '1',
15              IntType 'int',
16            Name 'pntr'
17          ],
18      StructDecl
19        Name 'st2',
20        [
21          Alloc
22            Writeable,
23            StructSpec
24              Name 'st1',
25              Name 'st'
26          ],
27      FunDef
28        VoidType 'void',
29        Name 'main',
30        [],
31        [
32          Assign
33            Alloc
34              Writeable,
35              IntType 'int',
36              Name 'var',
37              Num '42',
38          Assign
39            Alloc
40              Writeable,
41              StructSpec
42                Name 'st1',
43                Name 'st',
44            Struct
45              [
46                Assign
47                  Name 'st',
48                  Struct
49                    [
50                      Assign
51                        Name 'pntr',
52                        Array
```

```

53         [
54             Array
55             [
56                 Ref
57                 Name 'var'
58             ]
59         ]
60     ]
61 ]
62 ]
63 ]

```

Code 3.36: Abstract Syntax Tree für Initialisierung von Structs

```

1 SymbolTable
2 [
3     Symbol(
4     {
5         type qualifier:      Empty()
6         datatype:            ArrayDecl([Num('1')], PtrDecl(Num('1'), IntType('int')))
7         name:                Name('ptr@st1')
8         value or address:    Empty()
9         position:            Pos(Num('1'), Num('17'))
10        size:                Num('1')
11    },
12    Symbol(
13    {
14        type qualifier:      Empty()
15        datatype:            StructDecl(Name('st1'), [Alloc(Writeable(),
16        ↪ ArrayDecl([Num('1')], PtrDecl(Num('1'), IntType('int'))), Name('ptr'))])
17        name:                Name('st1')
18        value or address:    [Name('ptr@st1')]
19        position:            Pos(Num('1'), Num('7'))
20        size:                Num('1')
21    },
22    Symbol(
23    {
24        type qualifier:      Empty()
25        datatype:            StructSpec(Name('st1'))
26        name:                Name('st@st2')
27        value or address:    Empty()
28        position:            Pos(Num('3'), Num('23'))
29        size:                Num('1')
30    },
31    Symbol(
32    {
33        type qualifier:      Empty()
34        datatype:            StructDecl(Name('st2'), [Alloc(Writeable(),
35        ↪ StructSpec(Name('st1')), Name('st'))])
36        name:                Name('st2')
37        value or address:    [Name('st@st2')]
38        position:            Pos(Num('3'), Num('7'))
39        size:                Num('1')
40    },

```



```

39 Symbol(
40     {
41         type qualifier:      Empty()
42         datatype:            FunDecl(VoidType('void'), Name('main'), [])
43         name:                Name('main')
44         value or address:    Empty()
45         position:            Pos(Num('5'), Num('5'))
46         size:                Empty()
47     },
48     Symbol(
49         {
50             type qualifier:    Writeable()
51             datatype:          IntType('int')
52             name:              Name('var@main')
53             value or address:   Num('0')
54             position:           Pos(Num('6'), Num('6'))
55             size:              Num('1')
56         },
57         Symbol(
58             {
59                 type qualifier:    Writeable()
60                 datatype:          StructSpec(Name('st1'))
61                 name:              Name('st@main')
62                 value or address:   Num('1')
63                 position:           Pos(Num('7'), Num('13'))
64                 size:              Num('1')
65             }
66         ]

```

Code 3.37: Symboltabelle für Initialisierung von Structs

```

1 File
2   Name './example_struct_init.picoc_mon',
3   [
4       Block
5         Name 'main.0',
6         [
7             // Assign(Name('var'), Num('42')),
8             Exp
9               Num '42',
10            Assign
11              GlobalWrite
12                Num '0',
13              Tmp
14                Num '1',
15            // Assign(Name('st'), Struct([Assign(Name('st'), Struct([Assign(Name('pntr'),
16            ↪   Array([Array([Ref(Name('var'))]))]))]),
17            Ref
18              GlobalRead
19                Num '0',
20            Assign
21              GlobalWrite
22                Num '1',
23            Tmp

```

```

23     Num '1',
24     Return
25     Empty
26   ]
27 ]

```

Code 3.38: PicoC Mon Pass für Initialisierung von Structs

```

1 File
2   Name './example_struct_init.reti_blocks',
3   [
4     Block
5     Name 'main.0',
6     [
7       # // Assign(Name('var'), Num('42')),
8       # Exp(Num('42')),
9       SUBI SP 1,
10      LOADI ACC 42,
11      STOREIN SP ACC 1,
12      # Assign(GlobalWrite(Num('0')), Tmp(Num('1'))),
13      LOADIN SP ACC 1,
14      STOREIN DS ACC 0,
15      ADDI SP 1,
16      # // Assign(Name('st'), Struct([Assign(Name('st')), Struct([Assign(Name('pntr')),
17      ↪   Array([Array([Ref(Name('var'))])])])]),
18      # Ref(GlobalRead(Num('0'))),
19      SUBI SP 1,
20      LOADI IN1 0,
21      ADD IN1 DS,
22      STOREIN SP IN1 1,
23      # Assign(GlobalWrite(Num('1')), Tmp(Num('1'))),
24      LOADIN SP ACC 1,
25      STOREIN DS ACC 1,
26      ADDI SP 1,
27      # Return(Empty()),
28      LOADIN BAF PC -1
29    ]
30  ]

```

Code 3.39: RETI Blocks Pass für Initialisierung von Structs

### 3.3.5.3 Zugriff auf Structattribut

```

1 struct pos {int x; int y;};
2
3 void main() {
4   struct pos st = {.x=4, .y=2};
5   st.y;
6 }

```

Code 3.40: PicoC Code für Zugriff auf Structattribut

```

1 File
2   Name './example_struct_attr_access.ast',
3   [
4     StructDecl
5       Name 'pos',
6       [
7         Alloc
8           Writeable,
9           IntType 'int',
10          Name 'x',
11         Alloc
12           Writeable,
13           IntType 'int',
14           Name 'y'
15       ],
16     FunDef
17       VoidType 'void',
18       Name 'main',
19       [],
20       [
21         Assign
22           Alloc
23             Writeable,
24             StructSpec
25               Name 'pos',
26               Name 'st',
27             Struct
28               [
29                 Assign
30                   Name 'x',
31                   Num '4',
32                 Assign
33                   Name 'y',
34                   Num '2'
35               ],
36             Exp
37               Attr
38                 Name 'st',
39                 Name 'y'
40             ]
41       ]

```

Code 3.41: Abstract Syntax Tree für Zugriff auf Structattribut

```

1 File
2   Name './example_struct_attr_access.picoc_mon',
3   [
4     Block
5       Name 'main.0',
6       [
7         // Assign(Name('st'), Struct([Assign(Name('x'), Num('4')), Assign(Name('y'),
8           ↪ Num('2'))])),
9       ]
10      Exp

```

```

9      Num '4',
10     Exp
11     Num '2',
12     Assign
13     GlobalWrite
14     Num '0',
15     Tmp
16     Num '2',
17     Ref
18     GlobalRead
19     Num '0',
20     Ref
21     Attr
22     Tmp
23     Num '1',
24     Name 'y',
25     Exp
26     Subscr
27     Tmp
28     Num '1',
29     Num '0',
30     Return
31     Empty
32 ]
33 ]

```

Code 3.42: PicoC Mon Pass für Zugriff auf Structattribut

```

1 File
2 Name './example_struct_attr_access.reti_blocks',
3 [
4   Block
5     Name 'main.0',
6     [
7       # // Assign(Name('st'), Struct([Assign(Name('x'), Num('4')), Assign(Name('y'),
8         ↪ Num('2'))])),
9       # Exp(Num('4')),
10      SUBI SP 1,
11      LOADI ACC 4,
12      STOREIN SP ACC 1,
13      # Exp(Num('2')),
14      SUBI SP 1,
15      LOADI ACC 2,
16      STOREIN SP ACC 1,
17      # Assign(GlobalWrite(Num('0')), Tmp(Num('2'))),
18      LOADIN SP ACC 1,
19      STOREIN DS ACC 1,
20      LOADIN SP ACC 2,
21      STOREIN DS ACC 0,
22      ADDI SP 2,
23      # Ref(GlobalRead(Num('0'))),
24      SUBI SP 1,
25      LOADI IN1 0,
26      ADD IN1 DS,

```

```

26     STOREIN SP IN1 1,
27     # Ref(Attr(Tmp(Num('1')), Name('y'))),
28     LOADIN SP IN1 1,
29     ADDI IN1 1,
30     STOREIN SP IN1 1,
31     # Exp(Subscr(Tmp(Num('1')), Num('0'))),
32     LOADIN SP IN1 1,
33     LOADIN IN1 ACC 0,
34     STOREIN SP ACC 1,
35     # Return(Empty()),
36     LOADIN BAF PC -1
37 ]
38 ]

```

Code 3.43: RETI Blocks Pass für Zugriff auf Structattribut

### 3.3.5.4 Zuweisung an Structattribut

```

1 struct pos {int x; int y;};
2
3 void main() {
4     struct pos st = {.x=4, .y=2};
5     st.y = 42;
6 }

```

Code 3.44: PicoC Code für Zuweisung an Structattribut

```

1 File
2   Name './example_struct_attr_assignment.ast',
3   [
4     StructDecl
5       Name 'pos',
6       [
7         Alloc
8           Writeable,
9           IntType 'int',
10          Name 'x',
11         Alloc
12           Writeable,
13           IntType 'int',
14           Name 'y'
15       ],
16     FunDef
17       VoidType 'void',
18       Name 'main',
19       [],
20       [
21         Assign
22           Alloc
23             Writeable,
24             StructSpec
25             Name 'pos',

```

```

26     Name 'st',
27     Struct
28     [
29         Assign
30         Name 'x',
31         Num '4',
32         Assign
33         Name 'y',
34         Num '2'
35     ],
36     Assign
37     Attr
38     Name 'st',
39     Name 'y',
40     Num '42'
41 ]
42 ]

```

Code 3.45: Abstract Syntax Tree für Zuweisung an Structattribut

```

1 File
2   Name './example_struct_attr_assignment.picoc_mon',
3   [
4     Block
5     Name 'main.0',
6     [
7       // Assign(Name('st'), Struct([Assign(Name('x'), Num('4')), Assign(Name('y'),
8       ↪ Num('2'))])),
9     Exp
10    Num '4',
11    Exp
12    Num '2',
13    Assign
14    GlobalWrite
15    Num '0',
16    Tmp
17    Num '2',
18    // Assign(Attr(Name('st'), Name('y')), Num('42')),
19    Exp
20    Num '42',
21    Ref
22    GlobalRead
23    Num '0',
24    Ref
25    Attr
26    Tmp
27    Num '1',
28    Name 'y',
29    Assign
30    Subscr
31    Tmp
32    Num '1',
33    Num '0',
34    Tmp

```

```

34     Num '2',
35     Return
36     Empty
37 ]
38 ]

```

Code 3.46: PicoC Mon Pass für Zuweisung an Structattribut

```

1 File
2   Name './example_struct_attr_assignment.reti.blocks',
3   [
4     Block
5       Name 'main.0',
6       [
7         # // Assign(Name('st'), Struct([Assign(Name('x'), Num('4')), Assign(Name('y'),
8         ↪   Num('2'))])),
9         # Exp(Num('4')),
10        SUBI SP 1,
11        LOADI ACC 4,
12        STOREIN SP ACC 1,
13        # Exp(Num('2')),
14        SUBI SP 1,
15        LOADI ACC 2,
16        STOREIN SP ACC 1,
17        # Assign(GlobalWrite(Num('0')), Tmp(Num('2'))),
18        LOADIN SP ACC 1,
19        STOREIN DS ACC 1,
20        LOADIN SP ACC 2,
21        STOREIN DS ACC 0,
22        ADDI SP 2,
23        # // Assign(Attr(Name('st'), Name('y')), Num('42')),
24        # Exp(Num('42')),
25        SUBI SP 1,
26        LOADI ACC 42,
27        STOREIN SP ACC 1,
28        # Ref(GlobalRead(Num('0'))),
29        SUBI SP 1,
30        LOADI IN1 0,
31        ADD IN1 DS,
32        STOREIN SP IN1 1,
33        # Ref(Attr(Tmp(Num('1')), Name('y'))),
34        LOADIN SP IN1 1,
35        ADDI IN1 1,
36        STOREIN SP IN1 1,
37        # Assign(Subscr(Tmp(Num('1')), Num('0')), Tmp(Num('2'))),
38        LOADIN SP IN1 1,
39        LOADIN SP ACC 2,
40        ADDI SP 2,
41        STOREIN IN1 ACC 0,
42        # Return(Empty()),
43        LOADIN BAF PC -1
44      ]
45    ]

```

Code 3.47: RETI Blocks Pass für Zuweisung an Structattribut

### 3.3.6 Umsetzung der Derived Datatypes im Zusammenspiel

#### 3.3.6.1 Einleitungsteil für Globale Statische Daten und Stackframe

```

1 struct ar_with_len {int len; int ar[2];};
2
3 void main() {
4     struct ar_with_len st_ar[3];
5     int (*pntr2)[3];
6     pntr2;
7 }
8
9 void fun() {
10    struct ar_with_len st_ar[3];
11    int (*pntr1)[3];
12    pntr1;
13 }

```

Code 3.48: PicoC Code für den Einleitungsteil

```

1 File
2   Name './example_derived_dts_introduction_part.ast',
3   [
4     StructDecl
5       Name 'ar_with_len',
6       [
7         Alloc
8           Writeable,
9           IntType 'int',
10          Name 'len',
11          Alloc
12            Writeable,
13            ArrayDecl
14              [
15                Num '2'
16              ],
17            IntType 'int',
18            Name 'ar'
19          ],
20      FunDef
21        VoidType 'void',
22        Name 'main',
23        [],
24        [
25          Exp
26            Alloc
27              Writeable,
28              ArrayDecl
29                [
30                  Num '3'
31                ],

```



```

32     StructSpec
33         Name 'ar_with_len',
34     Name 'st_ar',
35 Exp
36     Alloc
37         Writeable,
38     PtrDecl
39         Num '1',
40     ArrayDecl
41         [
42             Num '3'
43         ],
44     PtrDecl
45         Num '1',
46         IntType 'int',
47     Name 'pntr2',
48 Exp
49     Name 'pntr2'
50 ],
51 FunDef
52     VoidType 'void',
53     Name 'fun',
54     [],
55     [
56         Exp
57             Alloc
58                 Writeable,
59             ArrayDecl
60                 [
61                     Num '3'
62                 ],
63             StructSpec
64                 Name 'ar_with_len',
65             Name 'st_ar',
66         Exp
67             Alloc
68                 Writeable,
69             PtrDecl
70                 Num '1',
71             ArrayDecl
72                 [
73                     Num '3'
74                 ],
75             IntType 'int',
76             Name 'pntr1',
77         Exp
78             Name 'pntr1'
79     ]
80 ]

```

Code 3.49: Abstract Syntax Tree für den Einleitungsteil

```

1 File
2     Name './example_derived_dts_introduction_part.picoc_mon',

```

```

3  [
4    Block
5      Name 'main.1',
6      [
7        Exp
8          GlobalRead
9            Num '9',
10       Return
11         Empty
12     ],
13   Block
14     Name 'fun.0',
15     [
16       Exp
17         StackRead
18           Num '9',
19       Return
20         Empty
21     ]
22 ]

```

Code 3.50: PicoC Mon Pass für den Einleitungsteil

```

1 File
2   Name './example_derived_dts_introduction_part.reti_blocks',
3   [
4     Block
5       Name 'main.1',
6       [
7         # Exp(GlobalRead(Num('9'))),
8         SUBI SP 1,
9         LOADIN DS ACC 9,
10        STOREIN SP ACC 1,
11        # Return(Empty()),
12        LOADIN BAF PC -1
13      ],
14    Block
15      Name 'fun.0',
16      [
17        # Exp(StackRead(Num('9'))),
18        SUBI SP 1,
19        LOADIN BAF ACC -11,
20        STOREIN SP ACC 1,
21        # Return(Empty()),
22        LOADIN BAF PC -1
23      ]
24    ]

```

Code 3.51: RETI Blocks Pass für den Einleitungsteil

**3.3.6.2 Mittelteil für die verschiedenen Derived Datatypes**

```

1 struct st1 {int (*ar)[1];};
2
3 void main() {
4     int var[1] = {42};
5     struct st1 st_first = {.ar=&var};
6     (*st_first.ar)[0];
7 }

```

Code 3.52: PicoC Code für den Mittelteil

```

1 File
2   Name './example_derived_dts_main_part.ast',
3   [
4     StructDecl
5       Name 'st1',
6       [
7         Alloc
8           Writeable,
9           PtrDecl
10            Num '1',
11            ArrayDecl
12              [
13                Num '1'
14              ],
15            IntType 'int',
16            Name 'ar'
17          ],
18      FunDef
19        VoidType 'void',
20        Name 'main',
21        [],
22        [
23          Assign
24            Alloc
25              Writeable,
26              ArrayDecl
27                [
28                  Num '1'
29                ],
30              IntType 'int',
31              Name 'var',
32          Array
33            [
34              Num '42'
35            ],
36          Assign
37            Alloc
38              Writeable,
39              StructSpec
40                Name 'st1',
41                Name 'st_first',
42          Struct

```

```

43      [
44          Assign
45              Name 'ar',
46              Ref
47                  Name 'var'
48      ],
49      Exp
50      Subscr
51      Deref
52      Attr
53          Name 'st_first',
54          Name 'ar',
55          Num '0',
56          Num '0'
57  ]
58 ]

```

Code 3.53: Abstract Syntax Tree für den Mittelteil

```

1 File
2   Name './example_derived_dts_main_part.picoc_mon',
3   [
4       Block
5       Name 'main.0',
6       [
7           // Assign(Name('var'), Array([Num('42')]])),
8           Exp
9           Num '42',
10          Assign
11          GlobalWrite
12          Num '0',
13          Tmp
14          Num '1',
15          // Assign(Name('st_first'), Struct([Assign(Name('ar'), Ref(Name('var')))])),
16          Ref
17          GlobalRead
18          Num '0',
19          Assign
20          GlobalWrite
21          Num '1',
22          Tmp
23          Num '1',
24          // Exp(Subscr(Subscr(Attr(Name('st_first'), Name('ar')), Num('0')), Num('0'))),
25          Ref
26          GlobalRead
27          Num '1',
28          Ref
29          Attr
30          Tmp
31          Num '1',
32          Name 'ar',
33          Exp
34          Num '0',
35          Ref

```

```

36         Subscr
37         Tmp
38         Num '2',
39         Tmp
40         Num '1',
41     Exp
42     Num '0',
43     Ref
44     Subscr
45     Tmp
46     Num '2',
47     Tmp
48     Num '1',
49     Exp
50     Subscr
51     Tmp
52     Num '1',
53     Num '0',
54     Return
55     Empty
56 ]
57 ]

```

Code 3.54: PicoC Mon Pass für den Mittelteil

```

1 File
2   Name './example_derived_dts_main_part.reti_blocks',
3   [
4     Block
5     Name 'main.0',
6     [
7       # // Assign(Name('var'), Array([Num('42')])),
8       # Exp(Num('42')),
9       SUBI SP 1,
10      LOADI ACC 42,
11      STOREIN SP ACC 1,
12      # Assign(GlobalWrite(Num('0')), Tmp(Num('1'))),
13      LOADIN SP ACC 1,
14      STOREIN DS ACC 0,
15      ADDI SP 1,
16      # // Assign(Name('st_first'), Struct([Assign(Name('ar'), Ref(Name('var')))])),
17      # Ref(GlobalRead(Num('0'))),
18      SUBI SP 1,
19      LOADI IN1 0,
20      ADD IN1 DS,
21      STOREIN SP IN1 1,
22      # Assign(GlobalWrite(Num('1')), Tmp(Num('1'))),
23      LOADIN SP ACC 1,
24      STOREIN DS ACC 1,
25      ADDI SP 1,
26      # // Exp(Subscr(Subscr(Attr(Name('st_first'), Name('ar')), Num('0')), Num('0'))),
27      # Ref(GlobalRead(Num('1'))),
28      SUBI SP 1,
29      LOADI IN1 1,

```

```

30     ADD IN1 DS,
31     STOREIN SP IN1 1,
32     # Ref(Attr(Tmp(Num('1')), Name('ar'))),
33     LOADIN SP IN1 1,
34     ADDI IN1 0,
35     STOREIN SP IN1 1,
36     # Exp(Num('0')),
37     SUBI SP 1,
38     LOADI ACC 0,
39     STOREIN SP ACC 1,
40     # Ref(Subscr(Tmp(Num('2')), Tmp(Num('1')))),
41     LOADIN SP IN2 2,
42     LOADIN IN2 IN1 0,
43     LOADIN SP IN2 1,
44     MULTI IN2 1,
45     ADD IN1 IN2,
46     ADDI SP 1,
47     STOREIN SP IN1 1,
48     # Exp(Num('0')),
49     SUBI SP 1,
50     LOADI ACC 0,
51     STOREIN SP ACC 1,
52     # Ref(Subscr(Tmp(Num('2')), Tmp(Num('1')))),
53     LOADIN SP IN1 2,
54     LOADIN SP IN2 1,
55     MULTI IN2 1,
56     ADD IN1 IN2,
57     ADDI SP 1,
58     STOREIN SP IN1 1,
59     # Exp(Subscr(Tmp(Num('1')), Num('0'))),
60     LOADIN SP IN1 1,
61     LOADIN IN1 ACC 0,
62     STOREIN SP ACC 1,
63     # Return(Empty()),
64     LOADIN BAF PC -1
65 ]
66 ]

```

Code 3.55: RETI Blocks Pass für den Mittelteil

### 3.3.6.3 Schlussteil für die verschiedenen Derived Datatypes

```

1 struct st {int attr[2];};
2
3 void main() {
4     int ar1[1][2] = {{42, 314}};
5     struct st ar2[1] = {.attr={42, 314}};
6     int var = 42;
7     int *pntr1 = &var;
8     int **pntr2 = &pntr1;
9
10    ar1[0];
11    ar2[0];
12    *pntr2;
13 }

```

Code 3.56: PicoC Code für den Schlussteil

```
1 File
2   Name './example_derived_dts_final_part.ast',
3   [
4     StructDecl
5       Name 'st',
6       [
7         Alloc
8           Writeable,
9           ArrayDecl
10            [
11              Num '2'
12            ],
13            IntType 'int',
14            Name 'attr'
15        ],
16      FunDef
17        VoidType 'void',
18        Name 'main',
19        [],
20        [
21          Assign
22            Alloc
23              Writeable,
24              ArrayDecl
25                [
26                  Num '1',
27                  Num '2'
28                ],
29                IntType 'int',
30                Name 'ar1',
31              Array
32                [
33                  Array
34                    [
35                      Num '42',
36                      Num '314'
37                    ]
38                ],
39            Assign
40              Alloc
41                Writeable,
42                ArrayDecl
43                  [
44                    Num '1'
45                  ],
46                StructSpec
47                  Name 'st',
48                  Name 'ar2',
49                Struct
50                  [
51                    Assign
52                      Name 'attr',
```

```

53         Array
54         [
55             Num '42',
56             Num '314'
57         ]
58     ],
59     Assign
60     Alloc
61     Writeable,
62     IntType 'int',
63     Name 'var',
64     Num '42',
65     Assign
66     Alloc
67     Writeable,
68     PtrDecl
69     Num '1',
70     IntType 'int',
71     Name 'ptr1',
72     Ref
73     Name 'var',
74     Assign
75     Alloc
76     Writeable,
77     PtrDecl
78     Num '2',
79     IntType 'int',
80     Name 'ptr2',
81     Ref
82     Name 'ptr1',
83     Exp
84     Subscr
85     Name 'ar1',
86     Num '0',
87     Exp
88     Subscr
89     Name 'ar2',
90     Num '0',
91     Exp
92     Deref
93     Name 'ptr2',
94     Num '0'
95 ]
96 ]

```

Code 3.57: Abstract Syntax Tree für den Schlussteil

```

1 File
2   Name './example_derived_dts_final_part.picoc_mon',
3   [
4     Block
5     Name 'main.0',
6     [
7       // Assign(Name('ar1'), Array(Array([Num('42'), Num('314')]))),

```



```

8      Exp
9      Num '42',
10     Exp
11     Num '314',
12     Assign
13     GlobalWrite
14     Num '0',
15     Tmp
16     Num '2',
17     // Assign(Name('ar2'), Struct([Assign(Name('attr'), Array([Num('42'),
18     ↪ Num('314')]))])),
19     Exp
20     Num '42',
21     Exp
22     Num '314',
23     Assign
24     GlobalWrite
25     Num '2',
26     Tmp
27     Num '2',
28     // Assign(Name('var'), Num('42')),
29     Exp
30     Num '42',
31     Assign
32     GlobalWrite
33     Num '4',
34     Tmp
35     Num '1',
36     // Assign(Name('pntr1'), Ref(Name('var'))),
37     Ref
38     GlobalRead
39     Num '4',
40     Assign
41     GlobalWrite
42     Num '5',
43     Tmp
44     Num '1',
45     // Assign(Name('pntr2'), Ref(Name('pntr1'))),
46     Ref
47     GlobalRead
48     Num '5',
49     Assign
50     GlobalWrite
51     Num '6',
52     Tmp
53     Num '1',
54     // Exp(Subscr(Name('ar1'), Num('0'))),
55     Ref
56     GlobalRead
57     Num '0',
58     Exp
59     Num '0',
60     Ref
61     Subscr
62     Tmp
63     Num '2',
64     Tmp

```

```

64         Num '1',
65     Exp
66     Subscr
67     Tmp
68     Num '1',
69     Num '0',
70     // Exp(Subscr(Name('ar2'), Num('0'))),
71 Ref
72     GlobalRead
73     Num '2',
74 Exp
75     Num '0',
76 Ref
77     Subscr
78     Tmp
79     Num '2',
80     Tmp
81     Num '1',
82 Exp
83     Subscr
84     Tmp
85     Num '1',
86     Num '0',
87     // Exp(Subscr(Name('pntr2'), Num('0'))),
88 Ref
89     GlobalRead
90     Num '6',
91 Exp
92     Num '0',
93 Ref
94     Subscr
95     Tmp
96     Num '2',
97     Tmp
98     Num '1',
99 Exp
100    Subscr
101    Tmp
102    Num '1',
103    Num '0',
104 Return
105     Empty
106 ]
107 ]

```

Code 3.58: PicoC Mon Pass für den Schlussteil

```

1 File
2   Name './example_derived_dts_final_part.reti_blocks',
3   [
4     Block
5     Name 'main.0',
6     [
7       # // Assign(Name('ar1'), Array([Array([Num('42'), Num('314')])))),

```

```

8      # Exp(Num('42')),
9      SUBI SP 1,
10     LOADI ACC 42,
11     STOREIN SP ACC 1,
12     # Exp(Num('314')),
13     SUBI SP 1,
14     LOADI ACC 314,
15     STOREIN SP ACC 1,
16     # Assign(GlobalWrite(Num('0')), Tmp(Num('2'))),
17     LOADIN SP ACC 1,
18     STOREIN DS ACC 1,
19     LOADIN SP ACC 2,
20     STOREIN DS ACC 0,
21     ADDI SP 2,
22     # // Assign(Name('ar2'), Struct([Assign(Name('attr'), Array([Num('42'),
23     ↪ Num('314')]))])),
24     # Exp(Num('42')),
25     SUBI SP 1,
26     LOADI ACC 42,
27     STOREIN SP ACC 1,
28     # Exp(Num('314')),
29     SUBI SP 1,
30     LOADI ACC 314,
31     STOREIN SP ACC 1,
32     # Assign(GlobalWrite(Num('2')), Tmp(Num('2'))),
33     LOADIN SP ACC 1,
34     STOREIN DS ACC 3,
35     LOADIN SP ACC 2,
36     STOREIN DS ACC 2,
37     ADDI SP 2,
38     # // Assign(Name('var'), Num('42')),
39     # Exp(Num('42')),
40     SUBI SP 1,
41     LOADI ACC 42,
42     STOREIN SP ACC 1,
43     # Assign(GlobalWrite(Num('4')), Tmp(Num('1'))),
44     LOADIN SP ACC 1,
45     STOREIN DS ACC 4,
46     ADDI SP 1,
47     # // Assign(Name('pntr1'), Ref(Name('var'))),
48     # Ref(GlobalRead(Num('4'))),
49     SUBI SP 1,
50     LOADI IN1 4,
51     ADD IN1 DS,
52     STOREIN SP IN1 1,
53     # Assign(GlobalWrite(Num('5')), Tmp(Num('1'))),
54     LOADIN SP ACC 1,
55     STOREIN DS ACC 5,
56     ADDI SP 1,
57     # // Assign(Name('pntr2'), Ref(Name('pntr1'))),
58     # Ref(GlobalRead(Num('5'))),
59     SUBI SP 1,
60     LOADI IN1 5,
61     ADD IN1 DS,
62     STOREIN SP IN1 1,
63     # Assign(GlobalWrite(Num('6')), Tmp(Num('1'))),
64     LOADIN SP ACC 1,

```

```
64     STOREIN DS ACC 6,
65     ADDI SP 1,
66     # // Exp(Subscr(Name('ar1'), Num('0'))),
67     # Ref(GlobalRead(Num('0'))),
68     SUBI SP 1,
69     LOADI IN1 0,
70     ADD IN1 DS,
71     STOREIN SP IN1 1,
72     # Exp(Num('0')),
73     SUBI SP 1,
74     LOADI ACC 0,
75     STOREIN SP ACC 1,
76     # Ref(Subscr(Tmp(Num('2')), Tmp(Num('1')))),
77     LOADIN SP IN1 2,
78     LOADIN SP IN2 1,
79     MULTI IN2 2,
80     ADD IN1 IN2,
81     ADDI SP 1,
82     STOREIN SP IN1 1,
83     # Exp(Subscr(Tmp(Num('1')), Num('0'))),
84     # // Exp(Subscr(Name('ar2'), Num('0'))),
85     # Ref(GlobalRead(Num('2'))),
86     SUBI SP 1,
87     LOADI IN1 2,
88     ADD IN1 DS,
89     STOREIN SP IN1 1,
90     # Exp(Num('0')),
91     SUBI SP 1,
92     LOADI ACC 0,
93     STOREIN SP ACC 1,
94     # Ref(Subscr(Tmp(Num('2')), Tmp(Num('1')))),
95     LOADIN SP IN1 2,
96     LOADIN SP IN2 1,
97     MULTI IN2 2,
98     ADD IN1 IN2,
99     ADDI SP 1,
100    STOREIN SP IN1 1,
101    # Exp(Subscr(Tmp(Num('1')), Num('0'))),
102    LOADIN SP IN1 1,
103    LOADIN IN1 ACC 0,
104    STOREIN SP ACC 1,
105    # // Exp(Subscr(Name('pntr2'), Num('0'))),
106    # Ref(GlobalRead(Num('6'))),
107    SUBI SP 1,
108    LOADI IN1 6,
109    ADD IN1 DS,
110    STOREIN SP IN1 1,
111    # Exp(Num('0')),
112    SUBI SP 1,
113    LOADI ACC 0,
114    STOREIN SP ACC 1,
115    # Ref(Subscr(Tmp(Num('2')), Tmp(Num('1')))),
116    LOADIN SP IN2 2,
117    LOADIN IN2 IN1 0,
118    LOADIN SP IN2 1,
119    MULTI IN2 1,
120    ADD IN1 IN2,
```

```

121     ADDI SP 1,
122     STOREIN SP IN1 1,
123     # Exp(Subscr(Tmp(Num('1')), Num('0'))),
124     # Return(Empty()),
125     LOADIN BAF PC -1
126 ]
127 ]

```

Code 3.59: RETI Blocks Pass für den Schlussteil

### 3.3.7 Umsetzung von Funktionen

#### 3.3.7.1 Funktionen auflösen zu RETI Code

```

1 void main() {
2     return;
3 }
4
5 void fun1() {
6 }
7
8 int fun2() {
9     return 1;
10 }

```

Code 3.60: PicoC Code für 3 Funktionen

```

1 File
2   Name './example_3_funs.ast',
3   [
4     FunDef
5       VoidType 'void',
6       Name 'main',
7       [],
8       [
9         Return
10        Empty
11      ],
12    FunDef
13      VoidType 'void',
14      Name 'fun1',
15      [],
16      [],
17    FunDef
18      IntType 'int',
19      Name 'fun2',
20      [],
21      [
22        Return
23        Num '1'
24      ]
25  ]

```

25 ]

Code 3.61: Abstract Syntax Tree für 3 Funktionen

```
1 File
2   Name './example_3_funs.picoc_blocks',
3   [
4     FunDef
5       VoidType 'void',
6       Name 'main',
7       [],
8       [
9         Block
10          Name 'main.2',
11          [
12            Return
13              Empty
14          ]
15      ],
16     FunDef
17       VoidType 'void',
18       Name 'fun1',
19       [],
20       [
21         Block
22          Name 'fun1.1',
23          []
24      ],
25     FunDef
26       IntType 'int',
27       Name 'fun2',
28       [],
29       [
30         Block
31          Name 'fun2.0',
32          [
33            Return
34              Num '1'
35          ]
36      ]
37   ]
```

Code 3.62: PicoC Blocks Pass für 3 Funktionen

```
1 File
2   Name './example_3_funs.picoc_mon',
3   [
4     Block
5       Name 'main.2',
6       [
7         Return
```

```

8      Empty
9    ],
10   Block
11     Name 'fun1.1',
12     [
13       Return
14       Empty
15     ],
16   Block
17     Name 'fun2.0',
18     [
19       // Return(Num('1')),
20       Exp
21       Num '1',
22       Return
23       Tmp
24       Num '1'
25     ]
26 ]

```

Code 3.63: PicoC Mon Pass für 3 Funktionen

```

1 File
2   Name './example_3_funs.reti_blocks',
3   [
4     Block
5       Name 'main.2',
6       [
7         # Return(Empty()),
8         LOADIN BAF PC -1
9       ],
10    Block
11      Name 'fun1.1',
12      [
13        # Return(Empty()),
14        LOADIN BAF PC -1
15      ],
16    Block
17      Name 'fun2.0',
18      [
19        # // Return(Num('1')),
20        # Exp(Num('1')),
21        SUBI SP 1,
22        LOADI ACC 1,
23        STOREIN SP ACC 1,
24        # Return(Tmp(Num('1'))),
25        LOADIN SP ACC 1,
26        ADDI SP 1,
27        LOADIN BAF PC -1
28      ]
29 ]

```

Code 3.64: RETI Blocks Pass für 3 Funktionen

### 3.3.7.1.1 Sprung zur Main Funktion

```
1 void fun1() {  
2 }  
3  
4 int fun2() {  
5     return 1;  
6 }  
7  
8 void main() {  
9     return;  
10 }
```

Code 3.65: PicoC Code für Funktionen, wobei die main Funktion nicht die erste Funktion ist

```
1 File  
2   Name './example_3_funs_main.picoc_mon',  
3   [  
4       Block  
5         Name 'fun1.2',  
6         [  
7             Return  
8             Empty  
9         ],  
10      Block  
11        Name 'fun2.1',  
12        [  
13            // Return(Num('1')),  
14            Exp  
15              Num '1',  
16            Return  
17              Tmp  
18              Num '1'  
19        ],  
20      Block  
21        Name 'main.0',  
22        [  
23            Return  
24            Empty  
25        ]  
26    ]
```

Code 3.66: PicoC Mon Pass für Funktionen, wobei die main Funktion nicht die erste Funktion ist

```
1 File  
2   Name './example_3_funs_main.reti_blocks',  
3   [  
4       Block  
5         Name 'fun1.2',  
6         [  
7             Return  
8             Empty  
9         ],  
10      Block  
11        Name 'main.0',  
12        [  
13            Return  
14            Empty  
15        ]  
16    ]
```



```

7      # Return(Empty()),
8      LOADIN BAF PC -1
9  ],
10 Block
11   Name 'fun2.1',
12   [
13     # // Return(Num('1')),
14     # Exp(Num('1')),
15     SUBI SP 1,
16     LOADI ACC 1,
17     STOREIN SP ACC 1,
18     # Return(Tmp(Num('1'))),
19     LOADIN SP ACC 1,
20     ADDI SP 1,
21     LOADIN BAF PC -1
22   ],
23 Block
24   Name 'main.0',
25   [
26     # Return(Empty()),
27     LOADIN BAF PC -1
28   ]
29 ]

```

Code 3.67: PicoC Blocks Pass für Funktionen, wobei die main Funktion nicht die erste Funktion ist

```

1 File
2   Name './example_3_funs_main.reti_patch',
3   [
4     Block
5       Name 'start.3',
6       [
7         Exp
8           GoTo
9             Name 'main.0'
10        ],
11     Block
12       Name 'fun1.2',
13       [
14         # Return(Empty()),
15         LOADIN BAF PC -1
16       ],
17     Block
18       Name 'fun2.1',
19       [
20         # // Return(Num('1')),
21         # Exp(Num('1')),
22         SUBI SP 1,
23         LOADI ACC 1,
24         STOREIN SP ACC 1,
25         # Return(Tmp(Num('1'))),
26         LOADIN SP ACC 1,
27         ADDI SP 1,
28         LOADIN BAF PC -1

```

```

29     ],
30     Block
31     Name 'main.0',
32     [
33         # Return(Empty()),
34         LOADIN BAF PC -1
35     ]
36 ]

```

Code 3.68: PicoC Patch Pass für Funktionen, wobei die main Funktion nicht die erste Funktion ist

### 3.3.7.2 Funktionsdeklaration und -definition

```

1 int fun2(int var);
2
3 void fun1() {
4 }
5
6 void main() {
7     int var = fun2(42);
8     return;
9 }
10
11 int fun2(int var) {
12     return var;
13 }

```

Code 3.69: PicoC Code für Funktionen, wobei eine Funktion vorher deklariert werden muss

```

1 SymbolTable
2 [
3     Symbol(
4         {
5             type qualifier:      Empty()
6             datatype:            FunDecl(IntType('int'), Name('fun2'), [Alloc(Writeable(),
7                                     ↪ IntType('int'), Name('var'))])
8             name:                Name('fun2')
9             value or address:     Empty()
10            position:             Pos(Num('1'), Num('4'))
11            size:                 Empty()
12        },
13        Symbol(
14            {
15                type qualifier:      Empty()
16                datatype:            FunDecl(VoidType('void'), Name('fun1'), [])
17                name:                Name('fun1')
18                value or address:     Empty()
19                position:             Pos(Num('3'), Num('5'))
20                size:                 Empty()
21            },
22        Symbol(
23            {

```

```

23     type qualifier:      Empty()
24     datatype:           FunDecl(VoidType('void'), Name('main'), [])
25     name:                Name('main')
26     value or address:    Empty()
27     position:            Pos(Num('6'), Num('5'))
28     size:                Empty()
29 },
30 Symbol(
31 {
32     type qualifier:      Writeable()
33     datatype:           IntType('int')
34     name:                Name('var@main')
35     value or address:    Num('0')
36     position:            Pos(Num('7'), Num('6'))
37     size:                Num('1')
38 },
39 Symbol(
40 {
41     type qualifier:      Writeable()
42     datatype:           IntType('int')
43     name:                Name('var@fun2')
44     value or address:    Num('0')
45     position:            Pos(Num('11'), Num('13'))
46     size:                Num('1')
47 }
48 ]

```

Code 3.70: Symboltabelle für Funktionen, wobei eine Funktion vorher deklariert werden muss

### 3.3.7.3 Funktionsaufruf

#### 3.3.7.3.1 Ohne Rückgabewert

```

1 struct st {int attr1; int attr2[2];};
2
3 void stack_fun(struct st param[2][3]);
4
5 void main() {
6     struct st local_var[2][3];
7     stack_fun(local_var);
8     return;
9 }
10
11 void stack_fun(struct st param[2][3]) {
12     int local_var;
13 }

```

Code 3.71: PicoC Code für Funktionsaufruf ohne Rückgabewert

```

1 File
2     Name './example_fun_call_no_return_value.picoc_mon',

```

```

3  [
4    Block
5      Name 'main.1',
6      [
7        StackMalloc
8          Num '2',
9        Ref
10         GlobalRead
11         Num '0',
12        NewStackframe
13         Name 'stack_fun',
14        GoTo
15         Name 'addr@next_instr',
16      Exp
17        GoTo
18         Name 'stack_fun.0',
19      RemoveStackframe,
20      Return
21        Empty
22    ],
23    Block
24      Name 'stack_fun.0',
25      [
26        Return
27        Empty
28      ]
29  ]

```

Code 3.72: PicoC Mon Pass für Funktionsaufruf ohne Rückgabewert

```

1 File
2   Name './example_fun_call_no_return_value.reti.blocks',
3   [
4     Block
5       Name 'main.1',
6       [
7         # StackMalloc(Num('2')),
8         SUBI SP 2,
9         # Ref(GlobalRead(Num('0'))),
10        SUBI SP 1,
11        LOADI IN1 0,
12        ADD IN1 DS,
13        STOREIN SP IN1 1,
14        # NewStackframe(Name('stack_fun'), GoTo(Name('addr@next_instr'))),
15        MOVE BAF ACC,
16        ADDI SP 3,
17        MOVE SP BAF,
18        SUBI SP 4,
19        STOREIN BAF ACC 0,
20        LOADI ACC GoTo
21          Name 'addr@next_instr',
22        ADD ACC CS,
23        STOREIN BAF ACC -1,
24      Exp

```

```

25         GoTo
26         Name 'stack_fun.0',
27         # RemoveStackframe(),
28         MOVE BAF IN1,
29         LOADIN IN1 BAF 0,
30         MOVE IN1 SP,
31         # Return(Empty()),
32         LOADIN BAF PC -1
33     ],
34     Block
35     Name 'stack_fun.0',
36     [
37         # Return(Empty()),
38         LOADIN BAF PC -1
39     ]
40 ]

```

Code 3.73: RETI Blocks Pass für Funktionsaufruf ohne Rückgabewert

```

1 # StackMalloc(Num('2'))
2 SUBI SP 2;
3 # Ref(GlobalRead(Num('0')))
4 SUBI SP 1;
5 LOADI IN1 0;
6 ADD IN1 DS;
7 STOREIN SP IN1 1;
8 # NewStackframe(Name('stack_fun'), GoTo(Name('addr@next_instr')))
9 MOVE BAF ACC;
10 ADDI SP 3;
11 MOVE SP BAF;
12 SUBI SP 4;
13 STOREIN BAF ACC 0;
14 LOADI ACC 14;
15 ADD ACC CS;
16 STOREIN BAF ACC -1;
17 JUMP 5;
18 # RemoveStackframe()
19 MOVE BAF IN1;
20 LOADIN IN1 BAF 0;
21 MOVE IN1 SP;
22 # Return(Empty())
23 LOADIN BAF PC -1;
24 # Return(Empty())
25 LOADIN BAF PC -1;

```

Code 3.74: RETI Pass für Funktionsaufruf ohne Rückgabewert

### 3.3.7.3.2 Mit Rückgabewert

```

1 void stack_fun() {
2     return 42;

```

```

3 }
4
5 void main() {
6     int var = stack_fun();
7 }

```

Code 3.75: PicoC Code für Funktionsaufruf mit Rückgabewert

```

1 File
2   Name './example_fun_call_with_return_value.picoc_mon',
3   [
4     Block
5       Name 'stack_fun.1',
6       [
7         // Return(Num('42')),
8         Exp
9           Num '42',
10        Return
11          Tmp
12            Num '1'
13      ],
14     Block
15       Name 'main.0',
16       [
17         // Assign(Name('var'), Call(Name('stack_fun'), [])),
18         StackMalloc
19           Num '2',
20         NewStackframe
21           Name 'stack_fun',
22           GoTo
23             Name 'addr@next_instr',
24         Exp
25           GoTo
26             Name 'stack_fun.1',
27         RemoveStackframe,
28         Assign
29           GlobalWrite
30             Num '0',
31           Tmp
32             Num '1',
33         Return
34           Empty
35      ]
36   ]

```

Code 3.76: PicoC Mon Pass für Funktionsaufruf mit Rückgabewert

```

1 File
2   Name './example_fun_call_with_return_value.reti_blocks',
3   [
4     Block

```

```

5      Name 'stack_fun.1',
6      [
7          # // Return(Num('42')),
8          # Exp(Num('42')),
9          SUBI SP 1,
10         LOADI ACC 42,
11         STOREIN SP ACC 1,
12         # Return(Tmp(Num('1'))),
13         LOADIN SP ACC 1,
14         ADDI SP 1,
15         LOADIN BAF PC -1
16     ],
17     Block
18     Name 'main.0',
19     [
20         # // Assign(Name('var'), Call(Name('stack_fun'), [])),
21         # StackMalloc(Num('2')),
22         SUBI SP 2,
23         # NewStackframe(Name('stack_fun'), GoTo(Name('addr@next_instr'))),
24         MOVE BAF ACC,
25         ADDI SP 2,
26         MOVE SP BAF,
27         SUBI SP 2,
28         STOREIN BAF ACC 0,
29         LOADI ACC GoTo
30             Name 'addr@next_instr',
31         ADD ACC CS,
32         STOREIN BAF ACC -1,
33         Exp
34             GoTo
35             Name 'stack_fun.1',
36         # RemoveStackframe(),
37         MOVE BAF IN1,
38         LOADIN IN1 BAF 0,
39         MOVE IN1 SP,
40         # Assign(GlobalWrite(Num('0')), Tmp(Num('1'))),
41         LOADIN SP ACC 1,
42         STOREIN DS ACC 0,
43         ADDI SP 1,
44         # Return(Empty()),
45         LOADIN BAF PC -1
46     ]
47 ]

```

Code 3.77: RETI Blocks Pass für Funktionsaufruf mit Rückgabewert

```

1 JUMP 7;
2 # // Return(Num('42'))
3 # Exp(Num('42'))
4 SUBI SP 1;
5 LOADI ACC 42;
6 STOREIN SP ACC 1;
7 # Return(Tmp(Num('1')))
8 LOADIN SP ACC 1;

```

```

9 ADDI SP 1;
10 LOADIN BAF PC -1;
11 # // Assign(Name('var'), Call(Name('stack_fun'), []))
12 # StackMalloc(Num('2'))
13 SUBI SP 2;
14 # NewStackframe(Name('stack_fun'), GoTo(Name('addr@next_instr')))
15 MOVE BAF ACC;
16 ADDI SP 2;
17 MOVE SP BAF;
18 SUBI SP 2;
19 STOREIN BAF ACC 0;
20 LOADI ACC 17;
21 ADD ACC CS;
22 STOREIN BAF ACC -1;
23 JUMP -15;
24 # RemoveStackframe()
25 MOVE BAF IN1;
26 LOADIN IN1 BAF 0;
27 MOVE IN1 SP;
28 # Assign(GlobalWrite(Num('0')), Tmp(Num('1')))
29 LOADIN SP ACC 1;
30 STOREIN DS ACC 0;
31 ADDI SP 1;
32 # Return(Empty())
33 LOADIN BAF PC -1;

```

Code 3.78: RETI Pass für Funktionsaufruf mit Rückgabewert

### 3.3.7.3.3 Umsetzung von Call by Sharing für Arrays

```

1 void stack_fun(int (*param1)[3], int param2[2][3]) {
2 }
3
4 void main() {
5     int local_var1[2][3];
6     int local_var2[2][3];
7     stack_fun(local_var1, local_var2);
8 }

```

Code 3.79: PicoC Code für Call by Sharing für Arrays

```

1 File
2   Name './example_fun_call_by_sharing_array.piloc_mon',
3   [
4     Block
5       Name 'stack_fun.1',
6       [
7         Return
8         Empty
9       ],
10    Block

```



```

11     Name 'main.0',
12     [
13         StackMalloc
14         Num '2',
15         Ref
16         GlobalRead
17         Num '0',
18         Ref
19         GlobalRead
20         Num '6',
21     NewStackframe
22         Name 'stack_fun',
23         GoTo
24         Name 'addr@next_instr',
25     Exp
26         GoTo
27         Name 'stack_fun.1',
28     RemoveStackframe,
29     Return
30     Empty
31 ]
32 ]

```

Code 3.80: PicoC Mon Pass für Call by Sharing für Arrays

```

1 SymbolTable
2 [
3     Symbol(
4     {
5         type qualifier:      Empty()
6         datatype:            FunDecl(VoidType('void'), Name('stack_fun'),
7                               ↪ [Alloc(Writeable(), PntrDecl(Num('1'), ArrayDecl([Num('3')], IntType('int'))),
8                               ↪ Name('param1')), Alloc(Writeable(), ArrayDecl([Num('3')], IntType('int')),
9                               ↪ Name('param2'))])
10        name:                Name('stack_fun')
11        value or address:     Empty()
12        position:             Pos(Num('1'), Num('5'))
13        size:                 Empty()
14    },
15    Symbol(
16    {
17        type qualifier:      Writeable()
18        datatype:            PntrDecl(Num('1'), ArrayDecl([Num('3')], IntType('int')))
19        name:                Name('param1@stack_fun')
20        value or address:     Num('0')
21        position:             Pos(Num('1'), Num('21'))
22        size:                 Num('1')
23    },
24    Symbol(
25    {
26        type qualifier:      Writeable()
27        datatype:            PntrDecl(Num('1'), ArrayDecl([Num('3')], IntType('int')))
28        name:                Name('param2@stack_fun')
29        value or address:     Num('1')

```

```

27     position:      Pos(Num('1'), Num('37'))
28     size:          Num('1')
29   },
30   Symbol(
31   {
32     type qualifier: Empty()
33     datatype:       FunDecl(VoidType('void'), Name('main'), [])
34     name:           Name('main')
35     value or address: Empty()
36     position:       Pos(Num('4'), Num('5'))
37     size:           Empty()
38   },
39   Symbol(
40   {
41     type qualifier: Writeable()
42     datatype:       ArrayDecl([Num('2'), Num('3')], IntType('int'))
43     name:           Name('local_var1@main')
44     value or address: Num('0')
45     position:       Pos(Num('5'), Num('6'))
46     size:           Num('6')
47   },
48   Symbol(
49   {
50     type qualifier: Writeable()
51     datatype:       ArrayDecl([Num('2'), Num('3')], IntType('int'))
52     name:           Name('local_var2@main')
53     value or address: Num('6')
54     position:       Pos(Num('6'), Num('6'))
55     size:           Num('6')
56   }
57 ]

```

Code 3.81: Symboltabelle für Call by Sharing für Arrays

```

1 File
2   Name './example_fun_call_by_sharing_array.reti_blocks',
3   [
4     Block
5       Name 'stack_fun.1',
6       [
7         # Return(Empty()),
8         LOADIN BAF PC -1
9       ],
10    Block
11      Name 'main.0',
12      [
13        # StackMalloc(Num('2')),
14        SUBI SP 2,
15        # Ref(GlobalRead(Num('0'))),
16        SUBI SP 1,
17        LOADI IN1 0,
18        ADD IN1 DS,
19        STOREIN SP IN1 1,
20        # Ref(GlobalRead(Num('6'))),

```

```

21     SUBI SP 1,
22     LOADI IN1 6,
23     ADD IN1 DS,
24     STOREIN SP IN1 1,
25     # NewStackframe(Name('stack_fun'), GoTo(Name('addr@next_instr'))),
26     MOVE BAF ACC,
27     ADDI SP 4,
28     MOVE SP BAF,
29     SUBI SP 4,
30     STOREIN BAF ACC 0,
31     LOADI ACC GoTo
32         Name 'addr@next_instr',
33     ADD ACC CS,
34     STOREIN BAF ACC -1,
35     Exp
36     GoTo
37     Name 'stack_fun.1',
38     # RemoveStackframe(),
39     MOVE BAF IN1,
40     LOADIN IN1 BAF 0,
41     MOVE IN1 SP,
42     # Return(Empty()),
43     LOADIN BAF PC -1
44 ]
45 ]

```

Code 3.82: RETI Block Pass für Call by Sharing für Arrays

### 3.3.7.3.4 Umsetzung von Call by Value für Structs

```

1 struct st {int attr1; int attr2[2];};
2
3 void stack_fun(struct st param) {
4 }
5
6 void main() {
7     struct st local_var;
8     stack_fun(local_var);
9 }

```

Code 3.83: PicoC Code für Call by Value für Structs

```

1 File
2     Name './example_fun_call_by_value_struct.picoc_mon',
3     [
4         Block
5             Name 'stack_fun.1',
6             [
7                 Return
8                 Empty
9             ],

```

```

10  Block
11      Name 'main.0',
12      [
13          StackMalloc
14              Num '2',
15          Assign
16              Tmp
17              Num '3',
18              GlobalRead
19              Num '0',
20          NewStackframe
21              Name 'stack_fun',
22              GoTo
23                  Name 'addr@next_instr',
24          Exp
25              GoTo
26                  Name 'stack_fun.1',
27          RemoveStackframe,
28          Return
29              Empty
30      ]
31 ]

```

Code 3.84: PicoC Mon Pass für Call by Value für Structs

```

1 File
2   Name './example_fun_call_by_value_struct.reti.blocks',
3   [
4       Block
5           Name 'stack_fun.1',
6           [
7               # Return(Empty()),
8               LOADIN BAF PC -1
9           ],
10      Block
11          Name 'main.0',
12          [
13              # StackMalloc(Num('2')),
14              SUBI SP 2,
15              # Assign(Tmp(Num('3')), GlobalRead(Num('0'))),
16              SUBI SP 3,
17              LOADIN DS ACC 0,
18              STOREIN SP ACC 1,
19              LOADIN DS ACC 1,
20              STOREIN SP ACC 2,
21              LOADIN DS ACC 2,
22              STOREIN SP ACC 3,
23              # NewStackframe(Name('stack_fun'), GoTo(Name('addr@next_instr'))),
24              MOVE BAF ACC,
25              ADDI SP 5,
26              MOVE SP BAF,
27              SUBI SP 5,
28              STOREIN BAF ACC 0,
29              LOADI ACC GoTo

```

```
30         Name 'addr@next_instr',
31     ADD ACC CS,
32     STOREIN BAF ACC -1,
33     Exp
34     GoTo
35     Name 'stack_fun.1',
36     # RemoveStackframe(),
37     MOVE BAF IN1,
38     LOADIN IN1 BAF 0,
39     MOVE IN1 SP,
40     # Return(Empty()),
41     LOADIN BAF PC -1
42 ]
43 ]
```

Code 3.85: RETI Block Pass für Call by Value für Structs

### 3.3.8 Umsetzung kleinerer Details

## 3.4 Fehlermeldungen

### 3.4.1 Error Handler

### 3.4.2 Arten von Fehlermeldungen

#### 3.4.2.1 Syntaxfehler

#### 3.4.2.2 Laufzeitfehler

# 4 Ergebnisse und Ausblick

## 4.1 Compiler

### 4.1.1 Überblick über Funktionen

### 4.1.2 Vergleich mit GCC

### 4.1.3 Showmode

## 4.2 Qualitätssicherung

## 4.3 Erweiterungsideen

Wenn eines Tages eine **RETI-CPU** auf einem **FPGA** implementiert werden sollte, sodass ein **provisorisches Betriebssystem** darauf laufen könnte, dann wäre der nächste Schritt einen **Self-Compiling Compiler**  $C_{RETI\_PicoC}^{PicoC}$  (Definition 4.1) zu schreiben. Dadurch kann die **Unabhängigkeit** von der Programmiersprache  $L_{Python}$ , in der der momentane Compiler  $C_{PicoC}$  für  $L_{PicoC}$  implementiert ist und die Unabhängigkeit von einer **anderen Maschine**, die bisher immer für das Cross-Compiling notwendig war erreicht werden.

### Definition 4.1: Self-compiling Compiler

Compiler  $C_w^w$ , der in der Sprache  $L_w$  **geschrieben** ist, die er **selbst** kompiliert. Also ein Compiler, der sich **selbst** kompilieren kann.<sup>a</sup>

<sup>a</sup>Earley und Sturgis, „A formalism for translator interactions“.

Will man nun für eine Maschine  $M_{RETI}$ , auf der bisher keine anderen Programmiersprachen mittels **Bootstrapping** (Definition 4.4) zum laufen gebracht wurden, den gerade beschriebenen **Self-compiling Compiler**  $C_{RETI\_PicoC}^{PicoC}$  implementieren und hat bereits den gesamten **Self-compiling Compiler**  $C_{RETI\_PicoC}^{PicoC}$  in der Sprache  $L_{PicoC}$  geschrieben, so stösst man auf ein Problem, dass auf das **Henne-Ei-Problem**<sup>1</sup> reduziert werden kann. Man bräuchte, um den **Self-compiling Compiler**  $C_{RETI\_PicoC}^{PicoC}$  auf der **Maschine**  $M_{RETI}$  zu kompilieren bereits einen kompilierten **Self-compiling Compiler**  $C_{RETI\_PicoC}^{PicoC}$ , der mit der Maschinensprache  $B_{RETI}$  läuft. Es liegt eine **zirkulare Abhängigkeit** vor, die man nur auflösen kann, indem eine **externe Entität** zur Hilfe nimmt.

Da man den gesamten **Self-compiling Compiler**  $C_{RETI\_PicoC}^{PicoC}$  nicht selbst komplett in der Maschinensprache  $B_{RETI}$  schreiben will, wäre eine Möglichkeit, dass man den **Cross-Compiler**  $C_{PicoC}^{Python}$ , den man bereits in der Programmiersprache  $L_{Python}$  implementiert hat, der in diesem Fall einen **Bootstrapping Compiler**

<sup>1</sup>Beschreibt die Situation, wenn ein System sich selbst als **Abhängigkeit** hat, damit es überhaupt einen **Anfang** für dieses System geben kann. Dafür steht das Problem mit der **Henne** und dem **Ei** sinnbildlich, da hier die Frage ist, wie das ganze seinen Anfang genommen hat, da beides **zirkular** voneinander abhängt.

(Definition 4.3) darstellt, auf einer anderen Maschine  $M_{other}$  dafür nutzt, damit dieser den **Self-compiling Compiler**  $C_{RETI\_PicoC}^{PicoC}$  für die Maschine  $M_{RETI}$  kompiliert bzw. **bootstrapped** und man den kompilierten **RETI-Maschiendencode** dann einfach von der Maschine  $M_{other}$  auf die Maschine  $M_{RETI}$  kopiert.<sup>2</sup>

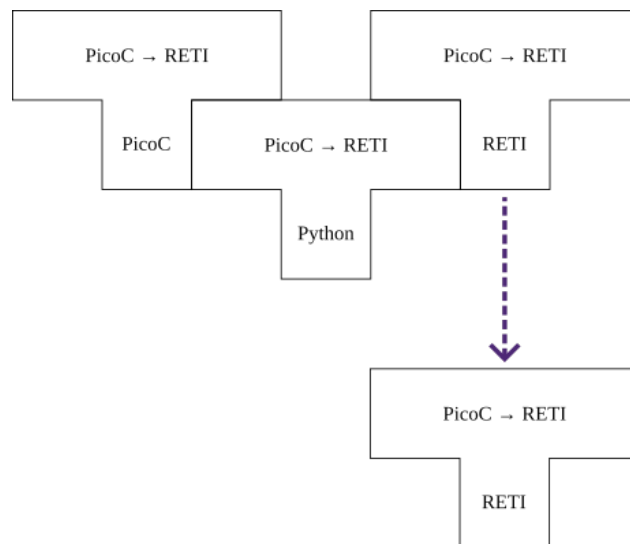


Abbildung 4.1: Cross-Compiler als Bootstrap Compiler

Einen ersten **minimalen Compiler**  $C_{2\_w\_min}$  für eine Maschine  $M_2$  und Wunschsprache  $L_w$  kann man entweder mittels eines **externen Bootstrap Compilers**  $C_w^o$  kompilieren<sup>a</sup> oder man schreibt ihn direkt in der **Maschinensprache**  $B_2$  bzw. wenn ein **Assembler** vorhanden ist, in der **Assemblesprache**  $A_2$ .

Die letzte Option wäre allerdings nur beim allerersten Compiler  $C_{first}$  für eine allererste **abstraktere Programmiersprache**  $L_{first}$  mit Schleifen, Verzweigungen usw. notwendig gewesen. Ansonsten hätte man immer eine Kette, die beim allerersten Compiler  $C_{first}$  anfängt fortführen können, in der ein Compiler einen anderen Compiler kompiliert bzw. einen ersten minimalen Compiler kompiliert und dieser minimale Compiler dann eine umfangreichere Version von sich kompiliert usw.

<sup>a</sup>In diesem Fall, dem **Cross-Compiler**  $C_{PicoC}^{Python}$ .

#### Definition 4.2: Minimaler Compiler

Compiler  $C_{w\_min}$ , der nur die **notwendigsten Funktionalitäten** einer Wunschsprache  $L_w$ , wie **Schleifen, Verzweigungen** kompiliert, die für die Implementierung eines **Self-compiling Compilers**  $C_w^o$  oder einer **ersten Version**  $C_{w_i}^{w_i}$  des Self-compiling Compilers  $C_w^o$  wichtig sind.<sup>a,b</sup>

<sup>a</sup>Den **PicoC-Compiler** könnte man auch als einen **minimalen Compiler** ansehen.

<sup>b</sup>Thiemann, „Compilerbau“.

<sup>2</sup>Im Fall, dass auf der Maschine  $M_{RETI}$  die Programmiersprache  $L_{Python}$  bereits mittels **Bootstrapping** zum Laufen gebracht wurde, könnte der **Self-compiling Compiler**  $C_{RETI\_PicoC}^{PicoC}$  auch mithilfe des **Cross-Compilers**  $C_{PicoC}^{Python}$  als **externe Entität** und der Programmiersprache  $L_{Python}$  auf der Maschine  $M_{RETI}$  selbst kompiliert werden.

**Definition 4.3: Bootstrap Compiler**

Compiler  $C_w^o$ , der es ermöglicht einen **Self-compiling Compiler**  $C_w^w$  zu **bootstrappen**, indem der **Self-compiling Compiler**  $C_w^w$  mit dem **Bootstrap Compiler**  $C_w^o$  **kompiliert** wird<sup>a</sup>. Der **Bootstrapping Compiler** stellt die **externe Entität** dar, die es ermöglicht die **zirkulare Abhängigkeit**, dass initial ein **Self-compiling Compiler**  $C_w^w$  bereits kompiliert vorliegen müsste, um sich selbst kompilieren zu können, zu brechen.<sup>b</sup>

<sup>a</sup>Dabei kann es sich um einen **lokal** auf der Maschine selbst laufenden Compiler oder auch um einen **Cross-Compiler** handeln.

<sup>b</sup>Thiemann, „Compilerbau“.

Aufbauend auf dem **Self-compiling Compiler**  $C_{RETI\_PicoC}^{PicoC}$ , der einen **minimalen Compiler** (Definition 4.2) für eine Teilmenge der **Programmiersprache**  $C$  bzw.  $L_C$  darstellt, könnte man auch noch weitere Teile der Programmiersprache  $C$  bzw.  $L_C$  für die Maschine  $M_{RETI}$  mittels **Bootstrapping** implementieren.<sup>3</sup>

Das bewerkstelligt man, indem man **iterativ** auf der Zielmaschine  $M_{RETI}$  selbst, aufbauend auf diesem **minimalen Compiler**  $C_{RETI\_PicoC}^{PicoC}$ , wie in Subdefinition 4.4.1 den minimalen Compiler schrittweise zu einem immer vollständigeren **C-Compiler**  $C_C$  weiterentwickelt.

**Definition 4.4: Bootstrapping**

Wenn man einen **Self-compiling Compiler**  $C_w^w$  einer Wunschsprache  $L_w$  auf einer **Zielmaschine**  $M$  zum laufen bringt<sup>a,b,c,d</sup>. Dabei ist die Art von **Bootstrapping** in 4.4.1 nochmal gesondert hervorzuheben:

**4.4.1:** Wenn man die **aktuelle Version** eines **Self-compiling Compilers**  $C_{w_i}^{w_i}$  der Wunschsprache  $L_{w_i}$  mithilfe von **früheren Versionen** seiner selbst kompiliert. Man schreibt also z.B. die **aktuelle Version** des **Self-compiling Compilers** in der Sprache  $L_{w_{i-1}}$ , welche von der **früheren Version** des **Compilers**, dem **Self-compiling Compiler**  $C_{w_{i-1}}^{w_{i-1}}$  **kompiliert** wird und schafft es so **iterativ** immer **umfangreichere Compiler** zu bauen.<sup>e,f,g</sup>

<sup>a</sup>Z.B. mithilfe eines **Bootstrap Compilers**.

<sup>b</sup>Der Begriff hat seinen Ursprung in der englischen **Redewendung** „pulling yourself up by your own bootstraps“, was im deutschen ungefähr der aus den **Lügend Geschichten des Freiherrn von Münchhausen** bekannten Redewendung „sich am eigenen Schopf aus dem Sumpf ziehen“ entspricht.

<sup>c</sup>Hat man einmal einen solchen **Self-compiling Compiler**  $C_w^w$  auf der Maschine  $M$  zum laufen gebracht, so kann man den Compiler auf der Maschine  $M$  weiterentwickeln, ohne von externen Entitäten, wie einer bestimmten Sprache  $L_o$ , in der der Compiler oder eine frühere Version des Compilers ursprünglich geschrieben war abhängig zu sein.

<sup>d</sup>Einen Compiler in der Sprache zu schreiben, die er selbst kompiliert und diesen Compiler dann sich selbst kompilieren zu lassen, kann eine gute **Probe aufs Exempel** darstellen, dass der Compiler auch wirklich funktioniert.

<sup>e</sup>Es ist hierbei theoretisch nicht notwendig den **letzten** **Self-compiling Compiler**  $C_{w_{i-1}}^{w_{i-1}}$  für das Kompilieren des **neuen** **Self-compiling Compilers**  $C_{w_i}^{w_i}$  zu verwenden, wenn z.B. der **Self-compiling Compiler**  $C_{w_{i-3}}^{w_{i-3}}$  auch bereits alle Funktionalitäten, die beim Schreiben des **Self-compiling Compilers**  $C_w^w$  verwendet werden kompilieren kann.

<sup>f</sup>Der Begriff ist sinnverwandt mit dem **Booten** eines Computers, wo die wichtigste Software, der **Kernel** zuerst in den Speicher geladen wird und darauf aufbauend von diesem dann das Betriebssysteme, welches bei Bedarf dann **Systemsoftware**, Software, die das Ausführen von Anwendungssoftware ermöglicht oder unterstützt, wie z.B. Treiber und **Anwendungssoftware**, Software, deren Anwendung darin besteht, dass sie dem Benutzer unmittelbar eine Dienstleistung zur Verfügung stellt, lädt.

<sup>g</sup>Earley und Sturgis, „A formalism for translator interactions“.

<sup>3</sup>Natürlich könnte man aber auch einfach den **Cross-Compiler**  $C_{PicoC}^{Python}$  um weitere Funktionalitäten von  $L_C$  erweitern, hat dann aber weiterhin eine **Abhängigkeit** von der Programmiersprache  $L_{Python}$ .



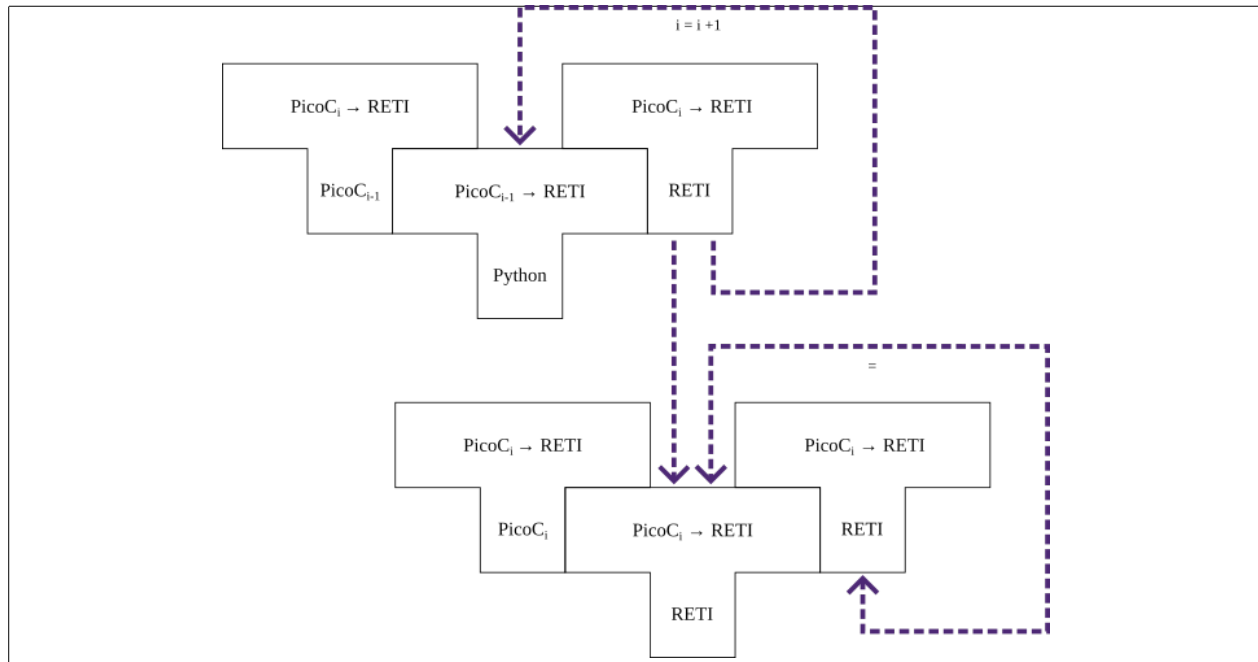


Abbildung 4.2: Iteratives Bootstrapping

Auch wenn ein **Self-compiling Compiler**  $C_{w_i}^{w_i}$  in der Subdefinition 4.4.1 selbst in einer früheren Version  $L_{w_{i-1}}$  der Programmiersprache  $L_{w_i}$  geschrieben wird, wird dieser nicht mit  $C_{w_i}^{w_{i-1}}$  bezeichnet, sondern mit  $C_{w_i}^{w_i}$ , da es bei **Self-compiling Compilern** darum geht, dass diese zwar in der Subdefinition 4.4.1 eine frühere Version  $C_{w_{i-1}}^{w_{i-1}}$  nutzen, um sich selbst kompilieren zu lassen, aber sie auch in der Lage sind sich selber zu kompilieren.

---

---

# A Appendix

## A.1 Konkrete und Abstrakte Syntax

## A.2 Bedienungsanleitungen

### A.2.1 PicoC-Compiler

### A.2.2 Showmode

### A.2.3 Entwicklertools

---

---

# Literatur

## Online

- *C Operator Precedence* - *cppreference.com*. URL: [https://en.cppreference.com/w/c/language/operator\\_precedence](https://en.cppreference.com/w/c/language/operator_precedence) (besucht am 27.04.2022).
- *Errors in C/C++* - *GeeksforGeeks*. URL: <https://www.geeksforgeeks.org/errors-in-cc/> (besucht am 10.05.2022).
- *JSON parser - Tutorial* — *Lark documentation*. URL: [https://lark-parser.readthedocs.io/en/latest/json\\_tutorial.html](https://lark-parser.readthedocs.io/en/latest/json_tutorial.html) (besucht am 09.07.2022).
- Ljohhuh. *What is an immediate value?* 4. Apr. 2018. URL: <https://reverseengineering.stackexchange.com/q/17671> (besucht am 13.04.2022).
- *Parsing Expressions · Crafting Interpreters*. URL: <https://www.craftinginterpreters.com/parsing-expressions.html> (besucht am 09.07.2022).
- *Transformers & Visitors* — *Lark documentation*. URL: <https://lark-parser.readthedocs.io/en/latest/visitors.html> (besucht am 09.07.2022).
- *What is Bottom-up Parsing?* URL: <https://www.tutorialspoint.com/what-is-bottom-up-parsing> (besucht am 22.06.2022).
- *What is Top-Down Parsing?* URL: <https://www.tutorialspoint.com/what-is-top-down-parsing> (besucht am 22.06.2022).

## Bücher

- G. Siek, Jeremy. *Course Webpage for Compilers (P423, P523, E313, and E513)*. 28. Jan. 2022. URL: <https://iucompilercourse.github.io/IU-Fall-2021/> (besucht am 28.01.2022).

## Artikel

- Earley, J. und Howard E. Sturgis. „A formalism for translator interactions“. In: *CACM* (1970). DOI: [10.1145/355598.362740](https://doi.org/10.1145/355598.362740).

## Vorlesungen

- Bast, Hannah. „Programmieren in C“. Vorlesung. Vorlesung. Universität Freiburg, 2020. URL: <https://ad-wiki.informatik.uni-freiburg.de/teaching/ProgrammierenCplusplusSS2020> (besucht am 09.07.2022).

- Nebel, Prof. Dr. Bernhard. „Theoretische Informatik“. Vorlesung. Vorlesung. Universität Freiburg, 2020. URL: [http://gki.informatik.uni-freiburg.de/teaching/ss20/info3/index\\_de.html](http://gki.informatik.uni-freiburg.de/teaching/ss20/info3/index_de.html) (besucht am 09.07.2022).
- Scholl, Christoph. „Betriebssysteme“. Vorlesung. Vorlesung. Universität Freiburg, 2020. URL: [https://abs.informatik.uni-freiburg.de/src/teach\\_main.php?id=157](https://abs.informatik.uni-freiburg.de/src/teach_main.php?id=157) (besucht am 09.07.2022).
- Scholl, Philipp. „Einführung in Embedded Systems“. Vorlesung. Vorlesung. Universität Freiburg, 2021. URL: <https://earth.informatik.uni-freiburg.de/uploads/es-2122/> (besucht am 09.07.2022).
- Thiemann, Peter. „Compilerbau“. Vorlesung. Vorlesung. Universität Freiburg, 2021. URL: <http://proglang.informatik.uni-freiburg.de/teaching/compilerbau/2021ws/> (besucht am 09.07.2022).
- — „Einführung in die Programmierung“. Vorlesung. Vorlesung. Universität Freiburg, 2018. URL: <http://proglang.informatik.uni-freiburg.de/teaching/info1/2018/> (besucht am 09.07.2022).

## Sonstige Quellen

- *Lark - a parsing toolkit for Python*. 26. Apr. 2022. URL: <https://github.com/lark-parser/lark> (besucht am 28.04.2022).