## Albert Ludwigs Universität Freiburg

### TECHNISCHE FAKULTÄT

### PicoC-Compiler

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

BACHELORARBEIT

Abgabedatum: 13. September 2022

Autor: Jürgen Mattheis

Gutachter: Prof. Dr. Scholl

Betreung: 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.

### Danksagungen

Bevor der Inhalt dieser Schrifftlichen Ausarbeitung der Bachelorarbeit anfängt, will ich einigen Personen noch meinen Dank aussprechen.

Ich schreibe die folgenden Danksagungen nicht auf eine bestimmte Weise, wie es sich vielleicht etabliert haben sollte Danksagungen zu schreiben und verwende auch keine künstlichen Floskeln, wie "mein aufrichtigster Dank" oder "aus tiefstem Herzen", sondern drücke im Folgenden die Dinge nur so aus, wie ich sie auch wirklich meine.

Estmal, ich hatte selten im Studium das Gefühl irgendwo Kunde zu sein, aber bei dieser Bachelorarbeit und dem vorangegangenen Bachelorprojekt hatte ich genau diese Gefühl, obwohl die Verhältnisse eigentlich genau umgekehrt sein sollten. Die Umgang mit mir wahr echt unglaublich nett und unbürokratisch, was ich als keine Selbverständlichkeit ansehe und sehr wertgeschätzt habe.

An erster Stelle will ich zu meinem Betreuer M.Sc. Tobias Seufert kommen, der netterweise auch bereits die Betreuung meines Bachelorprojektes übernommen hatte. Wie auch während des Bachelorprojektes, haben wir uns auch bei den Meetings während der Bachelorarbeit hervorragend verstanden. Dabei ging die Freundlichkeit und das Engagement seitens Tobias weit über das heraus, was man bereits als eine gute Betreuung bezeichnen würde.

Es gibt verschiedene Typen von Menschen, es gibt Leute, die nur genauso viel tun, wie es die Anforderungen verlangen und nichts darüberhinaus tun, wenn es nicht einen eigenen Vorteil für sie hat und es gibt Personen, die sich für nichts zu Schade sind und dies aus einer Philanthropie oder Leidenschafft heraus tun, auch wenn es für sie keine Vorteile hat. Tobias¹ konnte ich während der langen Zeit, die er mein Bachelorprojekt und dann meine Bachelorarbeit betreut hat eindeutig als letzteren Typ Mensch einordnen.

Er war sich nie zu Schade für meine vielen Fragen während der Meetings, auch wenn ich meine Zeit ziemlich oft überzogen habe<sup>2</sup>, er hat sich bei der Korrektur dieser Schrifftlichen Ausarbeitung sogar die Mühe gemacht bei den einzelnen Problemstellen längere, wirklich hilfreiche Textkommentare zu verfassen und obendrauf auch noch Tippfehler usw. angemerkt und war sich nicht zu Schade die Rolle des Nachrichtenübermittlers zwischen mir und Prof. Dr. Scholl zu übernehmen. All dies war absolut keine Selbverständlichkeit, vor allem wenn ich die Betreuung anderer Studenten, die ich kenne mit der vergleiche, die mir zu Teil wurde.

An den Kommentar zu meinem Betreuer Tobias will ich einen Kommentar zu meinem Gutachter Prof. Dr. Scholl anschließen. Wofür ich meinem Gutachter Prof. Dr. Scholl sehr dankbar bin, ist, dass er meine damals sehr ambitionierten Ideen für mögliche Funktionalitäten, die ich in den PicoC-Compiler für die Bachelorarbeit implementierten wollte runtergeschraubt hat. Man erlebt es äußerst selten im Studium, dass Studenten freiwillig weniger Arbeit gegeben wird.

Bei den für die Bachelorarbeit zu implementierenden Funktionalitäten gab es bei der Implementierung viele unerwartete kleine Details, die ich vorher garnicht bedacht hatte, die in ihrer Masse unerwartet viel Zeit zum Implementieren gebraucht haben. Mit den von Prof. Dr. Scholl festgelegten Funktionalitäten für die Bachelorarbeit ist der Zeitplan jedoch ziemlich perfekt aufgegangen. Mit meinen ambitionierten Plänen wäre es bei der Bachelorarbeit dagegeben wohl mit der Zeit äußerst kritisch geworden. Das Prof. Dr. Scholl mir zu

<sup>&</sup>lt;sup>1</sup>Wie auch Prof. Dr. Scholl. Hier geht es aber erstmal um Tobias.

<sup>&</sup>lt;sup>2</sup>Wofür ich mich auch nochmal Entschuldigen will.

seinem eigenen Nachteil $^3$  weniger Arbeit aufgebrummt hat empfand ich als ich eine äußerst nette Geste, die ich sehr geschätzt habe.

Wie mein Betreuer M.Sc. Tobias Seufert und wahrscheinlich auch mein Gutachter Prof. Dr. Scholl im Verlauf dieser Bachelorarbeit und des vorangegangenen Bachelorprojektes gemerkt haben, kann ich schon manchmal ziemlich eigensinnigen sein, bei der Weise, wie ich bestimmte Dinge umsetzen will. Ich habe es sehr geschätzt, dass mir das durchgehen gelassen wurde. Es ist, wie ich die Universitätswelt als Student erlebe bei Arbeitsvorgaben keine Selbverständlichkeit, dass dem Studenten überhaupt die Freiheit und das Vertrauen gegeben wird diese auf seine eigenen Weise umzusetzen.

Vor allem, da mein eigenes Vorgehen größtenteils Vorteile für mich hatte, da ich auf diese Weise am meisten über Compilerbau gelernt hab und eher Nachteile für Prof. Dr. Scholl, da mein eigenes Vorgehen entsprechend mehr Zeit brauchte und ich daher als Bachelorarbeit keinen dazu passenden RETI-Emulator mit Graphischer Anzeige implementieren konnte, da die restlichen Funktionalitäten des PicoC-Compilers noch implementiert werden mussten.

Glücklicherweise gibt es aber doch noch einen passenden RETI-Emulator, der den PicoC-Compiler über seine Kommandozeilenargumente aufruft, um ein PicoC-Programm visuell auf einer RETI-CPU auszuführen. Für dessen Implementierung hat sich Michel Giehl netterweise zur Verfügung gestellt. Daher Danke auch an Michel Giehl, dass er sich mit meinem PicoC-Compiler ausgeinandergesetzt hat und diesen in seinen RETI-Emulator integriert hat, sodass am Ende durch unsere beiden Arbeiten ein anschauliches Lerntool für die kommenden Studentengenerationen entstehen konnte. Vor allem da er auch mir darin vertrauen musste, dass ich mit meinem PicoC-Compiler nicht irgendeinen Misst baue. Der RETI-Emulator von Michel Giehl ist unter Link<sup>5</sup> zu finden.

Mir hat die Implementierung des PicoC-Compilers tatsächlich ziemlich viel Spaß gemacht, da Compilerbau auch in mein perönliches Interessengebiet fällt<sup>6</sup>. Das Aufschreiben dieser Schrifftlichen Ausarbeitung hat mir dagegen eher weniger Spaß gemacht<sup>7</sup>. Wobei ich allerdings sagen muss, dass ich eine große Erleichterung verspüre das ganze Wissen über Compilerbau mal aufgeschrieben zu haben, damit ich mir keine Sorgen machen muss dieses ziemlich nützliche Wissen irgendwann wieder zu vergessen. Es hilft einem auch als Programmierer ungemein weiter zu wissen, wie ein Compiler unter der Haube funktioniert, da man sich so viel besser merken, wie eine bestimmte Funktionalität einer Programmiersprache zu verwenden ist. Manch eine Funktionalität einer Programmiesprache kann in der Verwendung ziemlich wilkürlich erscheinen, wenn man die technische Umsetzung dahinter im Compiler nicht kennt.

Ich wollte mich daher auch noch dafür Bedanken, dass mir ein so ergiebiges und interessantes Thema als Bachelorarbeit vorgeschlagen wurde und vor allem, dass auch das Vertrauen in mich gesteckt wurde, dass ich am Ende auch einen funktionsfähigen, sauber programmierten und gut durchdachten Compiler implementiere.

Zum Schluss nochmal ein abschließendes Danke an meinen Betreuer M.Sc Seufert und meinen Gutachter Prof. Dr. Scholl für die Betreuung und Bereitstellung dieser interessanten Bachelorarbeit und des vorangegangenen Bachelorprojektes und Michel Giehl für das Integrieren des PicoC-Compilers in seinen RETI-Emulator.

<sup>&</sup>lt;sup>3</sup>Der PicoC-Compiler hätte schließlich mehr Funktionalitäten haben können.

<sup>&</sup>lt;sup>4</sup>Vielleicht finde ich ja noch im nächsten Semester während des Betriebssysteme Tutorats noch etwas Zeit einige weitere Features einzubauen oder möglicherweise im Rahmen eines Masterprojektes <sup>3</sup>.

 $<sup>^5</sup>$ https://github.com/michel-giehl/Reti-Emulator.

<sup>&</sup>lt;sup>6</sup>Womit nicht alle Studenten so viel Glück haben.

<sup>&</sup>lt;sup>7</sup>Dieses ständige überlegen, wo man möglicherweise eine Erklärlücke hat, ob man nicht was wichtiges ausgelassen hat usw.

### Inhaltsverzeichnis

Abbildungsv	erzeichn	is	I
Codeverzeich	nnis		II
Tabellenverz	eichnis		III
Definitionsve	erzeichni	s	IV
Grammatikv	erzeichn	is	$\mathbf{V}$
0.0.1	Umsetz	ung von Zeigern	1
	0.0.1.1	Referenzierung	1
	0.0.1.2	Referenzierung	3
0.0.2	Umsetz	ung von Feldern	5
	0.0.2.1	Initialisierung eines Feldes	5
	0.0.2.2	Zugriff auf einen Feldindex	
	0.0.2.3	Zuweisung an Feldindex	14
Literatur			$\mathbf{A}$

# Abbildungsverzeichnis

### Codeverzeichnis

0.1	PicoC-Code für Zeigerreferenzierung
0.2	Abstrakter Syntaxbaum für Zeigerreferenzierung
0.3	Symboltabelle für Zeigerreferenzierung
0.4	PicoC-ANF Pass für Zeigerreferenzierung
0.5	RETI-Blocks Pass für Zeigerreferenzierung.
0.6	PicoC-Code für Zeigerdereferenzierung.
0.7	Abstrakter Syntaxbaum für Zeigerdereferenzierung
0.8	PicoC-Shrink Pass für Zeigerdereferenzierung
0.9	PicoC-Code für die Initialisierung eines Feldes.
0.10	Abstrakter Syntaxbaum für die Initialisierung eines Feldes
0.11	Symboltabelle für die Initialisierung eines Feldes
0.12	PicoC-ANF Pass für die Initialisierung eines Feldes
0.13	RETI-Blocks Pass für die Initialisierung eines Feldes
0.14	PicoC-Code für Zugriff auf einen Feldindex
0.15	Abstrakter Syntaxbaum für Zugriff auf einen Feldindex
0.16	PicoC-ANF Pass für Zugriff auf einen Feldindex
0.17	RETI-Blocks Pass für Zugriff auf einen Feldindex
0.18	PicoC-Code für Zuweisung an Feldindex
0.19	Abstrakter Syntaxbaum für Zuweisung an Feldindex
0.20	PicoC-ANF Pass für Zuweisung an Feldindex
0.21	RETI-Blocks Pass für Zuweisung an Feldindex

### **Tabellenverzeichnis**

1	Datensegment nach o	der Initialisierung beider	Felder

### Definitionsverzeichnis

### Grammatikverzeichnis

#### 0.0.1 Umsetzung von Zeigern

Die Umsetzung von Zeigern ist in diesem Unterkapitel schnell erklärt, auch Dank eines kleinen Taschenspielertricks<sup>8</sup>. Hierbei sind nur die Operationen für Referenzierung und Dereferenzierung in den Unterkapiteln 0.0.1.1 und 0.0.1.2 zu erläutern. Referenzierung kann dazu genutzt werden einen Zeiger zu initialisieren und Dereferenzierung kann dazu genutzt werden, um auf diesen später zuzugreifen.

#### 0.0.1.1 Referenzierung

Referenzierung (z.B. &var) ist eine Operation bei der ein Zeiger auf eine Location in Form der Anfangsadresse dieser Location als Ergebnis zurückgegeben wird. Die Implementierung der Referenzierung wird im Folgenden anhand des Beispiels in Code 0.1 erklärt.

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

Code 0.1: PicoC-Code für Zeigerreferenzierung.

Der Knoten Ref(Name('var'))) repräsentiert im Abstrakten Syntaxbaum in Code 0.2 eine Referenzierung &var und der Knoten PntrDecl(Num('1'), IntType('int')) repräsentiert einen Zeiger \*pntr.

```
File
Name './example_pntr_ref.ast',

[
FunDef
VoidType 'void',
Name 'main',
[],
[],
[]
Assign(Alloc(Writeable(), IntType('int'), Name('var')), Num('42'))
Assign(Alloc(Writeable(), PntrDecl(Num('1'), IntType('int')), Name('pntr')),
AREF(Name('var')))

11  ]
12 ]
```

Code 0.2: Abstrakter Syntaxbaum für Zeigerreferenzierung.

Bevor man einem Zeiger eine Adresse (z.B. &var) zuweisen kann, muss dieser erstmal definiert sein. Dafür braucht es einen Eintrag in der Symboltabelle in Code 0.3.

```
Anmerkung Q
```

Die Anzahl Speicherzellen, die ein Zeiger<sup>a</sup> (z.B. eines Zeigers auf ein Feld von Integern: pntr = int \*pntr[3]) belegt ist dabei immer: size(type(pntr)) = 1 Speicherzelle.  $^{bcd}$ 

<sup>&</sup>lt;sup>a</sup>Die im size-Attribut der Symboltabelle eingetragen ist.

 $<sup>^</sup>b$ Eine Speicherzelle ist in der RETI-Architektur, wie in Unterkapitel  $\ref{eq:condition}$  erklärt 4 Byte breit.

 $<sup>^</sup>c$ Die Funktion size berechnet die Anzahl Speicherzellen, die ein Datentyp belegt.

<sup>&</sup>lt;sup>8</sup>Später mehr dazu.

 $^d$ Die Funktion type ordnet einer Variable ihren Datentyp zu. Das ist notwendig, weil die Funktion size als Definitionsmenge Datentypen hat.

```
SymbolTable
 2
     Ε
       Symbol
         {
           type qualifier:
                                     Empty()
           datatype:
                                     FunDecl(VoidType('void'), Name('main'), [])
 7
8
                                     Name('main')
           name:
           value or address:
                                     Empty()
 9
           position:
                                     Pos(Num('1'), Num('5'))
10
                                     Empty()
           size:
11
         },
12
       Symbol
13
14
                                     Writeable()
           type qualifier:
15
           datatype:
                                     IntType('int')
16
                                     Name('var@main')
           name:
17
                                     Num('0')
           value or address:
18
                                     Pos(Num('2'), Num('6'))
           position:
19
           size:
                                     Num('1')
20
         },
21
       Symbol
22
         {
23
           type qualifier:
                                     Writeable()
                                     PntrDecl(Num('1'), IntType('int'))
24
           datatype:
25
           name:
                                     Name('pntr@main')
26
                                     Num('1')
           value or address:
27
           position:
                                     Pos(Num('3'), Num('7'))
28
           size:
                                     Num('1')
29
         }
30
     ]
```

Code 0.3: Symboltabelle für Zeigerreferenzierung.

Im PicoC-ANF Pass in Code 0.4 wird der Knoten Ref(Name('var'))) durch die Knoten Ref(GlobalRead(Num('0'))) und Assign(GlobalWrite(Num('1')), Tmp(Num('1'))) ersetzt, deren Bedeutung in Unterkapitel ?? erklärt wurde. Im Fall, dass in Ref(exp)) das exp vielleicht nicht direkt ein Name('var') enthält und exp z.B. ein Subscr(Attr(Name('var'), Name('attr'))) ist, sind noch weitere Anweisungen zwischen den Zeilen 11 und 12 nötig, die sich in diesem Beispiel um das Übersetzen von Subscr(exp) und Attr(exp, name) nach dem Schema in Subkapitel ?? kümmern.

Code 0.4: PicoC-ANF Pass für Zeigerreferenzierung.

Im RETI-Blocks Pass in Code 0.5 werden die PicoC-Knoten Ref(Global(Num('0'))) und Assign(Global(Num('1')), Stack(Num('1'))) durch ihre semantisch entsprechenden RETI-Knoten ersetzt, deren Bedeutungen im ?? in Unterkapitel ?? erklärt sind.

```
1 File
     Name './example_pntr_ref.reti_blocks',
     Γ
       Block
         Name 'main.0',
 6
7
8
9
           # // Assign(Name('var'), Num('42'))
           # Exp(Num('42'))
           SUBI SP 1;
10
           LOADI ACC 42;
           STOREIN SP ACC 1;
12
           # Assign(Global(Num('0')), Stack(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(Global(Num('0')))
18
           SUBI SP 1;
19
           LOADI IN1 0;
20
           ADD IN1 DS;
           STOREIN SP IN1 1;
22
           # Assign(Global(Num('1')), Stack(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
         ]
     ]
```

Code 0.5: RETI-Blocks Pass für Zeigerreferenzierung.

#### 0.0.1.2 Dereferenzierung durch Zugriff auf Feldindex ersetzen

Dereferenzierung (z.B. \*var) ist eine Operation bei der einem Zeiger zur Location hin gefolgt wird, auf welche dieser zeigt und das Ergebnis z.B. der Inhalt der ersten Speicherzelle der referenzierten Location ist. Die Implementierung von Dereferenzierung wird im Folgenden anhand des Beispiels in Code 0.6 erklärt.

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

Code 0.6: PicoC-Code für Zeigerdereferenzierung.

Der Knoten Deref(Name('var'), Num('0'))) repräsentiert im Abstrakten Syntaxbaum in Code 0.7 eine Dereferenzierung \*var. Es gibt hierbei zwei Fälle. Bei der Anwendung von Zeigerarithmetik, wie z.B. \*(var + 2 - 1) übersetzt sich diese zu Deref(Name('var'), BinOp(Num('2'), Sub(), Num('1'))). Bei einer normalen Dereferenzierung, wie z.B. \*var, übersetzt sich diese zu Deref(Name('var'), Num('0'))<sup>9</sup>.

```
File
    Name './example_pntr_deref.ast',
4
      FunDef
        VoidType 'void',
        Name 'main',
         [],
8
          Assign(Alloc(Writeable(), IntType('int'), Name('var')), Num('42'))
10
          Assign(Alloc(Writeable(), PntrDecl(Num('1'), IntType('int')), Name('pntr')),

→ Ref(Name('var')))
          Exp(Deref(Name('pntr'), Num('0')))
11
12
        ٦
13
    ]
```

Code 0.7: Abstrakter Syntaxbaum für Zeigerdereferenzierung.

Im PicoC-Shrink Pass in Code 0.8 wird ein Trick angewandet, bei dem jeder Knoten Deref(Name('pntr'), Num('0')) einfach durch den Knoten Subscr(Name('pntr'), Num('0')) ersetzt wird. Die Bedeutung des letzteren Knoten wurde in Unterkapitel ?? erklärt. Der Trick besteht darin, dass der Dereferenzierungsoperator (z.B. \*(var + 1)) sich identisch zum Operator für den Zugriff auf einen Feldindex (z.B. var[1]) verhält, wie es bereits im Unterkapitel ?? erläutert wurde. Damit spart man sich viele vermeidbare Fallunterscheidungen und doppelten Code und kann die Derefenzierung (z.B. \*(var + 1)) einfach von den Routinen für einen Zugriff auf einen Feldindex (z.B. var[1]) übernehmen lassen.

<sup>&</sup>lt;sup>9</sup>Das Num('0') steht dafür, dass dem Zeiger gefolgt wird, aber danach nicht noch mit einem Versatz von der Größe des Datentyps in diesem Kontext auf eine nebenliegende Location zugegriffen wird.

```
Assign(Alloc(Writeable(), PntrDecl(Num('1'), IntType('int')), Name('pntr')),

Ref(Name('var')))

Exp(Subscr(Name('pntr'), Num('0')))

[ ]

]
```

Code 0.8: PicoC-Shrink Pass für Zeigerdereferenzierung.

#### 0.0.2 Umsetzung von Feldern

Bei Feldern ist in diesem Unterkapitel die Umsetzung der Innitialisierung eines Feldes 0.0.2.1, des Zugriffs auf einen Feldindex 0.0.2.2 und der Zuweisung an einen Feldindex 0.0.2.3 zu klären.

#### 0.0.2.1 Initialisierung eines Feldes

Die Umsetzung der Initialisierung eines Feldes (z.B. int ar[2][1] = {{3+1}, {5}}) wird im Folgenden anhand des Beispiels in Code 0.9 erklärt.

```
void fun() {
  int ar[2][2] = {{3, 4}, {5, 6}};
}

void main() {
  int ar[2][1] = {{3+1}, {5}};
}
```

Code 0.9: PicoC-Code für die Initialisierung eines Feldes.

Die Initialisierung eines Feldes intar[2][1]={{3+1},{5}} wird im Abstrakten Syntaxbaum in Code 0.10 mithilfe den Knoten Assign(Alloc(Writeable(),ArrayDecl([Num('2'),Num('1')],IntType('int')),Name('ar')),Array([Array([BinOp(Num('3'),Add('+'),Num('1'))]),Array([Num('5')])))) dargestellt.

```
1 File
    Name './example_array_init.ast',
    Γ
4
      FunDef
5
         VoidType 'void',
        Name 'fun',
         [],
         Γ
9
           Assign(Alloc(Writeable(), ArrayDecl([Num('2'), Num('2')], IntType('int')),
              Name('ar')), Array([Array([Num('3'), Num('4')]), Array([Num('5'), Num('6')])]))
10
        ],
11
      FunDef
12
         VoidType 'void',
13
        Name 'main',
         [],
         [
```

```
Assign(Alloc(Writeable(), ArrayDecl([Num('2'), Num('1')], IntType('int')),

→ Name('ar')), Array([Array([BinOp(Num('3'), Add('+'), Num('1'))]),

→ Array([Num('5')])]))

17  ]

18 ]
```

Code 0.10: Abstrakter Syntaxbaum für die Initialisierung eines Feldes.

Bei der Initialisierung eines Feldes wird zuerst Alloc(Writeable(), ArrayDecl([Num('2'), Num('1')], IntType('int'))) ausgewertet, da eine Variable zuerst definiert sein muss, bevor man sie verwenden kann<sup>10</sup>. Das Definieren der Variable ar erfolgt mittels der Symboltabelle, die in Code 0.11 dargestellt ist.

Auf dem Stackframe wird ein Feld verglichen zur Wachstumrichtung des Stacks rückwärts in den Stackframe geschrieben und die relative Adresse des ersten Elements als Adresse des Feldes in der Symboltabelle in Code 0.11 genommen. Dies ist in Tabelle 1 für ein Datensegment der Größe 8 dargstellt. Der Stack wächst zwar verglichen zu den Globalen Statischen Daten in die entgegengesetzte Richtung, aber Felder in den Globalen Statischen Daten und in einem Stackframe haben die gleiche Ausrichtung. Das macht den Zugriff auf einen Feldindex in Subkapitel 0.0.2.2 deutlich unkomplizierter. Auf diese Weise muss beim Zugriff auf einen Feldindex nicht zwischen Stackframe und Globalen Statischen Daten unterschieden werden.

Relativ-	$\mathbf{Wert}$
adresse	
0	4
1	5
3	3
2	4
1	5
0	6

Tabelle 1: Datensegment nach der Initialisierung beider Felder.

#### Anmerkung Q

Die Anzahl Speicherzellen, die ein Feld<sup>a</sup> datatype  $\operatorname{ar}[\dim_{k}]$  belegt berechnet sich aus der Mächtigkeit der einzelnen Dimensionen des Feldes, multipliziert mit der Größe des grundlegenden Datentyps der einzelnen Feldelemente:  $size(type(\operatorname{ar})) = \left(\prod_{i=1}^{n} \dim_{j}\right) \cdot size(\operatorname{datatype})^{bc}$ 

```
1 SymbolTable
2 [
```

<sup>&</sup>lt;sup>a</sup>Die im size-Attribut des Symboltabelleneintrags eingetragen ist.

<sup>&</sup>lt;sup>b</sup>Die Funktion size berechnet die Anzahl Speicherzellen, die ein Datentyp belegt.

 $<sup>^</sup>c$ Die Funktion type ordnet einer Variable ihren Datentyp zu. Das ist notwendig, weil die Funktion size als Definitionsmenge Datentypen hat.

<sup>&</sup>lt;sup>10</sup>Das widerspricht der üblichen Auswertungsreihenfolge beim Zuweisungsoperator =, der rechtsassoziativ ist. Der Zuweisungsoperator = tritt allerdings erst später in Aktion.

```
Symbol
 4
         {
 5
           type qualifier:
                                     Empty()
           datatype:
                                     FunDecl(VoidType('void'), Name('fun'), [])
 7
8
                                     Name('fun')
           name:
           value or address:
                                     Empty()
 9
                                     Pos(Num('1'), Num('5'))
           position:
10
                                     Empty()
           size:
11
         },
12
       Symbol
13
         {
14
           type qualifier:
                                     Writeable()
15
                                     ArrayDecl([Num('2'), Num('2')], IntType('int'))
           datatype:
16
                                     Name('ar@fun')
           name:
17
                                     Num('3')
           value or address:
18
           position:
                                     Pos(Num('2'), Num('6'))
19
           size:
                                     Num('4')
20
         },
21
       Symbol
22
23
           type qualifier:
                                     Empty()
24
                                     FunDecl(VoidType('void'), Name('main'), [])
           datatype:
25
           name:
                                     Name('main')
26
           value or address:
                                     Empty()
27
                                     Pos(Num('5'), Num('5'))
           position:
28
           size:
                                     Empty()
29
         },
30
       Symbol
31
         {
32
           type qualifier:
                                     Writeable()
33
                                     ArrayDecl([Num('2'), Num('1')], IntType('int'))
           datatype:
34
           name:
                                     Name('ar@main')
35
                                     Num('0')
           value or address:
                                     Pos(Num('6'), Num('6'))
36
           position:
37
                                     Num('2')
           size:
38
         }
39
```

Code 0.11: Symboltabelle für die Initialisierung eines Feldes.

Im PiocC-ANF Pass in Code 0.12 werden zuerst die Knoten für die Logischen Ausdrücke in den Blättern des Teilbaumes, dessen Wurzel der Feld-Initializer-Knoten Array([Array([BinOp(Num('3'), Add('+'), Num('1'))]), Array([Num('5')])]) ist nach dem Prinzip der Tiefensuche, von links-nach-rechts ausgewertet und durch Knoten erstetzt, welche das Ergebnis dieser Ausdrücke auf den Stack schreiben<sup>11</sup>.

Im finalen Schritt muss zwischen Globalen Statischen Daten bei der main-Funktion und Stackframe bei der Funktion fun unterschieden werden. Die auf den Stack ausgewerteten Logischen Ausdrücke werden mittels der Knoten Assign(Global(Num('0')), Stack(Num('2'))) bzw. Assign(Stackframe(Num('3')), Stack(Num('5'))), die in Tabelle ?? genauer beschrieben ist, versetzt in der selben Reihenfolge zu den Globalen Statischen Daten bzw. auf den Stackframe geschrieben.

Der Trick ist hier, dass egal wieviele Dimensionen und was für einen grundlegenden Datentyp<sup>12</sup> das Feld

<sup>&</sup>lt;sup>11</sup>Da der Zuweisungsoperator = rechtsassoziativ ist und auch rein logisch, weil man nichts zuweisen kann, was man noch nicht berechnet hat.

<sup>&</sup>lt;sup>12</sup>Z.B. ein Verbund, sodass es ein "Feld von Verbunden" ist.

hat, man letztendlich immer das gesamte Feld erwischt, wenn man einfach die Größe des Feldes viele Speicherzellen mit z.B. der Komposition Assign(Global(Num('0')), Stack(Num('2'))) verschiebt.

In die Knoten Global('0') und Stackframe('3') wurde hierbei die Startadresse des jeweiligen Feldes geschrieben, sodass man nach dem PicoC-ANF Pass nie mehr Variablen in der Symboltabelle nachsehen muss und gleich weiß, ob sie in Bezug zu den Globalen Statischen Daten oder dem Stackframe stehen.

```
File
 2
     Name './example_array_init.picoc_mon',
     Γ
 4
       Block
 5
         Name 'fun.1',
 6
         [
           // Assign(Name('ar'), Array([Array([Num('3'), Num('4')]), Array([Num('5'),
           → Num('6')])))
           Exp(Num('3'))
 9
           Exp(Num('4'))
10
           Exp(Num('5'))
11
           Exp(Num('6'))
           Assign(Stackframe(Num('3')), Stack(Num('4')))
13
           Return(Empty())
14
         ],
15
       Block
16
         Name 'main.0',
17
         Γ
18
           // Assign(Name('ar'), Array([Array([BinOp(Num('3'), Add('+'), Num('1'))]),

    Array([Num('5')]))))
19
           Exp(Num('3'))
20
           Exp(Num('1'))
21
           Exp(BinOp(Stack(Num('2')), Add('+'), Stack(Num('1'))))
22
           Exp(Num('5'))
23
           Assign(Global(Num('0')), Stack(Num('2')))
24
           Return(Empty())
25
         ]
     ]
26
```

Code 0.12: PicoC-ANF Pass für die Initialisierung eines Feldes.

Im RETI-Blocks Pass in Code 0.13 werden die Kompositionen Exp(exp) und Assign(Global(Num('0')), Stack(Num('2'))) bzw. Assign(Stackframe(Num('3')), Stack(Num('5'))) durch ihre entsprechenden RETI-Knoten ersetzt.

```
STOREIN SP ACC 1;
           # Exp(Num('4'))
12
13
           SUBI SP 1;
14
           LOADI ACC 4;
15
           STOREIN SP ACC 1;
16
           # Exp(Num('5'))
17
           SUBI SP 1;
18
           LOADI ACC 5;
19
           STOREIN SP ACC 1;
20
           # Exp(Num('6'))
21
           SUBI SP 1;
22
           LOADI ACC 6;
23
           STOREIN SP ACC 1;
24
           # Assign(Stackframe(Num('3')), Stack(Num('4')))
25
           LOADIN SP ACC 1;
26
           STOREIN BAF ACC -2;
           LOADIN SP ACC 2;
28
           STOREIN BAF ACC -3;
29
           LOADIN SP ACC 3;
30
           STOREIN BAF ACC -4;
31
           LOADIN SP ACC 4;
32
           STOREIN BAF ACC -5;
33
           ADDI SP 4;
34
           # Return(Empty())
35
           LOADIN BAF PC -1;
        ],
36
37
       Block
38
         Name 'main.0',
39
40
           # // Assign(Name('ar'), Array([Array([BinOp(Num('3'), Add('+'), Num('1'))]),

    Array([Num('5')]))))

           # Exp(Num('3'))
42
           SUBI SP 1;
43
           LOADI ACC 3;
44
           STOREIN SP ACC 1;
45
           # Exp(Num('1'))
46
           SUBI SP 1;
47
           LOADI ACC 1;
48
           STOREIN SP ACC 1;
49
           # Exp(BinOp(Stack(Num('2')), Add('+'), Stack(Num('1'))))
50
           LOADIN SP ACC 2;
51
           LOADIN SP IN2 1;
52
           ADD ACC IN2;
53
           STOREIN SP ACC 2;
54
           ADDI SP 1;
55
           # Exp(Num('5'))
56
           SUBI SP 1;
           LOADI ACC 5;
58
           STOREIN SP ACC 1;
59
           # Assign(Global(Num('0')), Stack(Num('2')))
60
           LOADIN SP ACC 1;
61
           STOREIN DS ACC 1;
62
           LOADIN SP ACC 2;
63
           STOREIN DS ACC 0;
           ADDI SP 2;
64
65
           # Return(Empty())
66
           LOADIN BAF PC -1;
```

```
67 ]
68 ]
```

Code 0.13: RETI-Blocks Pass für die Initialisierung eines Feldes.

#### 0.0.2.2 Zugriff auf einen Feldindex

Der Zugriff auf einen Feldindex (z.B. ar[0]) wird im Folgenden anhand des Beispiels in Code 0.14 erklärt.

```
void main() {
  int ar[1] = {42};
  ar[0];
}

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

Code 0.14: PicoC-Code für Zugriff auf einen Feldindex.

Der Zugriff auf einen Feldindex ar[0] wird im Abstrakten Syntaxbaum in Code 0.15 mithilfe des Knotens Subscr(Name('ar'), Num('0')) dargestellt.

```
File
    Name './example_array_access.ast',
    Γ
      FunDef
        VoidType 'void',
6
7
8
        Name 'main',
         [],
         Γ
           Assign(Alloc(Writeable(), ArrayDecl([Num('1')], IntType('int')), Name('ar')),
           → Array([Num('42')]))
10
          Exp(Subscr(Name('ar'), Num('0')))
11
        ],
12
      FunDef
13
        VoidType 'void',
14
        Name 'fun',
15
         [],
16
17
          Assign(Alloc(Writeable(), ArrayDecl([Num('3')], IntType('int')), Name('ar')),
           → Array([Num('1'), Num('2'), Num('3')]))
           Exp(Subscr(Name('ar'), BinOp(Num('1'), Add('+'), Num('1'))))
18
19
        ]
    ]
```

Code 0.15: Abstrakter Syntaxbaum für Zugriff auf einen Feldindex.

Im PicoC-ANF Pass in Code 0.16 wird vom Knoten Subscr(Name('ar'), Num('0')) zuerst im Anfangsteil?? die Adresse der Variable Name('ar') auf den Stack geschrieben. Bei den Globalen Statischen Daten der main-Funktion wird das durch die Komposition Ref(Global(Num('0'))) dargestellt und beim Stackframe der Funktionm fun wird das durch die Komposition Ref(Stackframe(Num('2'))) dargestellt.

In nächsten Schritt, dem Mittelteil ?? wird die Adresse ab der das Feldelement des Feldes auf das Zugegriffen werden soll anfängt berechnet. Dabei wurde im Anfangsteil bereits die Anfangsadresse des Feldes, in dem dieses Feldelement liegt auf den Stack gelegt. Da ein Index auf den Zugegriffen werden soll auch durch das Ergebnis eines komplexeren Ausdrucks, z.B. ar[1 + var] bestimmt sein kann, indem auch Variablen vorkommen können, kann dieser nicht während des Kompilierens berechnet werden, sondern muss zur Laufzeit berechnet werden.

Daher muss zuerst der Wert des Index, dessen Adresse berechnet werden soll bestimmt werden, z.B. im einfachen Fall durch Exp(Num('0')) und dann muss die Adresse des Index berechnet werden, was durch die Komposition Ref(Subscr(Stack(Num('2')), Stack(Num('1')))) dargestellt wird. Die Bedeutung der Komposition Ref(Subscr(Stack(Num('2')), Stack(Num('1')))) ist in Tabelle ?? dokumentiert.

Zur Adressberechnung ist es notwendig auf die Dimensionen (z.B. [Num('3')]) des Feldes, auf dessen Feldelement zugegriffen wird, zugreifen zu können. Daher ist der Felddatentyp (z.B. ArrayDecl([Num('3')], IntType('int'))) dem Knoten Ref(exp, datatype) als verstecktes Attribut datatype angehängt. Das versteckte Attribut wird während des Kompiliervorgangs im PiocC-ANF Pass dem Knoten Ref(exp, datatype) angehängt.

Je nachdem, ob mehrere Subscr(exp, exp) eine Komposition bilden (z.B. Subscr(Subscr(Name('var'), Num('1')), Num('1'))) ist es notwendig mehrere Adressberechnungsschritte für den Index Ref(Subscr(Stack(Num('2')), Stack(Num('1')))) einzuleiten und es muss auch möglich sein, z.B. einen Attributzugriff var.attr und eine Zugriff auf einen Arryindex var[1] miteinander zu kombinieren, was in Subkapitel ?? allgemein erklärt ist.

Im letzten Schritt, dem Schlussteil ?? wird der Inhalt des Index, dessen Adresse in den vorherigen Schritten berechnet wurde, nun auf den Stack geschrieben, wobei dieser die Adresse auf dem Stack ersetzt, die es zum Finden des Index brauchte. Dies wird durch den Knoten Exp(Stack(Num('1'))) dargestellt. Je nachdem, welchen Datentyp die Variable ar hat und auf welchen Unterdatentyp folglich im Kontext zuletzt zugegriffen wird, abhängig davon wird der Schlussteil Exp(Stack(Num('1'))) auf eine andere Weise verarbeitet (siehe Subkapitel ??). Der Unterdatentyp ist dabei ein verstecktes Attribut des Exp(Stack(Num('1')))-Knoten.

Der einzige Unterschied, je nachdem, ob der Zugriff auf einen Feldindex (z.B. ar[1]) in der main-Funktion oder der Funktion fum erfolgt, ist eigentlich nur beim Anfangsteil beim Schreiben der Adresse der Variable ar auf den Stack zu finden, bei dem unterschiedliche RETI-Befehle für eine Variable, die in den Globalen Statischen Daten liegt und eine Variable, die auf dem Stackframe liegt erzeugt werden müssen.

#### Anmerkung Q

Die Berechnung der Adresse, ab der ein Feldelement eines Feldes datatype  $ar[dim_1]...[dim_n]$  abgespeichert ist, kann mittels der Formel 0.0.1:

$$\mathtt{ref}(\mathtt{ar}[\mathtt{idx_1}]\dots[\mathtt{idx_n}]) = \mathtt{ref}(\mathtt{ar}) + \left(\sum_{\mathtt{i}=\mathtt{1}}^\mathtt{n} \left(\prod_{\mathtt{j}=\mathtt{i}+\mathtt{1}}^\mathtt{n} \mathtt{dim_j}\right) \cdot \mathtt{idx_i}\right) \cdot \mathtt{size}(\mathtt{datatype}) \tag{0.0.1}$$

aus der Betriebssysteme Vorlesung Scholl, "Betriebssysteme" berechnet werden $^{ab}$ .

Die Komposition Ref(Global(num)) bzw. Ref(Stackframe(num)) repräsentiert dabei den Summanden

```
1 File
    Name './example_array_access.picoc_mon',
 4
       Block
         Name 'main.1',
           // Assign(Name('ar'), Array([Num('42')]))
           Exp(Num('42'))
 9
           Assign(Global(Num('0')), Stack(Num('1')))
10
           // Exp(Subscr(Name('ar'), Num('0')))
11
           Ref(Global(Num('0')))
           Exp(Num('0'))
13
           Ref(Subscr(Stack(Num('2')), Stack(Num('1'))))
14
           Exp(Stack(Num('1')))
15
           Return(Empty())
16
         ],
17
       Block
18
         Name 'fun.0',
19
20
           // Assign(Name('ar'), Array([Num('1'), Num('2'), Num('3')]))
21
           Exp(Num('1'))
22
           Exp(Num('2'))
23
           Exp(Num('3'))
24
           Assign(Stackframe(Num('2')), Stack(Num('3')))
25
           // Exp(Subscr(Name('ar'), BinOp(Num('1'), Add('+'), Num('1'))))
26
           Ref(Stackframe(Num('2')))
27
           Exp(Num('1'))
28
           Exp(Num('1'))
29
           Exp(BinOp(Stack(Num('2')), Add('+'), Stack(Num('1'))))
30
           Ref(Subscr(Stack(Num('2')), Stack(Num('1'))))
           Exp(Stack(Num('1')))
32
           Return(Empty())
33
         ]
34
```

Code 0.16: PicoC-ANF Pass für Zugriff auf einen Feldindex.

Im RETI-Blocks Pass in Code 0.17 werden die Kompositionen Ref(Global(Num('0'))), Ref(Subscr(Stack(Num('2')) und Stack(Num('1')))) durch ihre entsprechenden RETI-Knoten ersetzt.

```
Name './example_array_access.reti_blocks',
       Block
         Name 'main.1',
           # // Assign(Name('ar'), Array([Num('42')]))
           # Exp(Num('42'))
           SUBI SP 1;
10
           LOADI ACC 42;
           STOREIN SP ACC 1;
11
12
           # Assign(Global(Num('0')), Stack(Num('1')))
13
           LOADIN SP ACC 1;
14
           STOREIN DS ACC 0;
           ADDI SP 1;
           # // Exp(Subscr(Name('ar'), Num('0')))
16
17
           # Ref(Global(Num('0')))
18
           SUBI SP 1;
19
           LOADI IN1 0;
20
           ADD IN1 DS;
21
           STOREIN SP IN1 1;
22
           # Exp(Num('0'))
23
           SUBI SP 1;
24
           LOADI ACC 0;
25
           STOREIN SP ACC 1;
26
           # Ref(Subscr(Stack(Num('2')), Stack(Num('1'))))
27
           LOADIN SP IN1 2;
28
           LOADIN SP IN2 1;
29
           MULTI IN2 1;
30
           ADD IN1 IN2;
31
           ADDI SP 1;
32
           STOREIN SP IN1 1;
33
           # Exp(Stack(Num('1')))
34
           LOADIN SP IN1 1;
35
           LOADIN IN1 ACC O;
36
           STOREIN SP ACC 1;
37
           # Return(Empty())
38
           LOADIN BAF PC -1;
39
         ],
40
       Block
41
         Name 'fun.0',
42
43
           # // Assign(Name('ar'), Array([Num('1'), Num('2'), Num('3')]))
44
           # Exp(Num('1'))
45
           SUBI SP 1;
46
           LOADI ACC 1;
47
           STOREIN SP ACC 1;
48
           # Exp(Num('2'))
49
           SUBI SP 1;
50
           LOADI ACC 2;
51
           STOREIN SP ACC 1;
52
           # Exp(Num('3'))
53
           SUBI SP 1;
54
           LOADI ACC 3;
           STOREIN SP ACC 1;
```

```
# Assign(Stackframe(Num('2')), Stack(Num('3')))
           LOADIN SP ACC 1;
57
58
           STOREIN BAF ACC -2;
59
           LOADIN SP ACC 2;
           STOREIN BAF ACC -3;
61
           LOADIN SP ACC 3;
62
           STOREIN BAF ACC -4;
63
           ADDI SP 3;
64
           # // Exp(Subscr(Name('ar'), BinOp(Num('1'), Add('+'), Num('1'))))
65
           # Ref(Stackframe(Num('2')))
66
           SUBI SP 1;
67
           MOVE BAF IN1;
68
           SUBI IN1 4;
69
           STOREIN SP IN1 1;
70
           # Exp(Num('1'))
71
           SUBI SP 1;
72
           LOADI ACC 1;
           STOREIN SP ACC 1;
           # Exp(Num('1'))
75
           SUBI SP 1;
           LOADI ACC 1;
76
77
           STOREIN SP ACC 1;
78
           # Exp(BinOp(Stack(Num('2')), Add('+'), Stack(Num('1'))))
79
           LOADIN SP ACC 2;
80
           LOADIN SP IN2 1;
           ADD ACC IN2;
81
82
           STOREIN SP ACC 2;
83
           ADDI SP 1;
84
           # Ref(Subscr(Stack(Num('2')), Stack(Num('1'))))
85
           LOADIN SP IN1 2;
           LOADIN SP IN2 1;
87
           MULTI IN2 1;
88
           ADD IN1 IN2;
89
           ADDI SP 1;
90
           STOREIN SP IN1 1;
91
           # Exp(Stack(Num('1')))
92
           LOADIN SP IN1 1;
93
           LOADIN IN1 ACC 0;
94
           STOREIN SP ACC 1;
95
           # Return(Empty())
96
           LOADIN BAF PC -1;
97
         ]
98
    ]
```

Code 0.17: RETI-Blocks Pass für Zugriff auf einen Feldindex.

#### 0.0.2.3 Zuweisung an Feldindex

Die **Zuweisung** eines Wertes an einen **Feldindex** (z.B. ar[2] = 42;) wird im Folgenden anhand des Beispiels in Code 0.18 erläutert.

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

4 }

Code 0.18: PicoC-Code für Zuweisung an Feldindex.

Im Abstrakten Syntaxbaum in Code 0.19 wird eine Zuweisung an einen Feldindex ar[2] = 42; durch die Komposition Assign(Subscr(Name('ar'), Num('2')), Num('42')) dargestellt.

```
File
Name './example_array_assignment.ast',

[
FunDef
VoidType 'void',
Name 'main',
[],
[],
[]
Exp(Alloc(Writeable(), ArrayDecl([Num('2')], IntType('int')), Name('ar')))
Assign(Subscr(Name('ar'), Num('1')), Num('42'))
]
]
```

Code 0.19: Abstrakter Syntaxbaum für Zuweisung an Feldindex.

Im PicoC-ANF Pass in Code 0.20 wird zuerst die rechte Seite des rechtsassoziativen Zuweisungsoperators =, bzw. des Knotens der diesen darstellt ausgewertet: Exp(Num('42')).

Danach ist das Vorgehen, bzw. sind die Kompostionen, die dieses darauffolgende Vorgehen darstellen: Ref(Global(Num('0'))), Exp(Num('2')) und Ref(Subscr(Stack(Num('2')), Stack(Num('1')))) identisch zum Anfangsteil und Mittelteil aus dem vorherigen Subkapitel 0.0.2.2. Es wird die Adresse des Index, dem das Ergebnis der Ausdrucks auf der rechten Seite des Zuweisungsoperators = zugewiesen wird berechet, wie in Subkapitel 0.0.2.2.

Zum Schluss stellt die Komposition Assign(Stack(Num('1')), Stack(Num('2')))<sup>13</sup> die Zuweisung = des Ergebnisses des Ausdrucks auf der rechten Seite der Zuweisung zum Feldindex, dessen Adresse im Schritt danach berechnet wurde dar.

<sup>&</sup>lt;sup>13</sup>Ist in Tabelle ?? genauer beschrieben ist

```
14 ]
15 ]
```

Code 0.20: PicoC-ANF Pass für Zuweisung an Feldindex.

Im RETI-Blocks Pass in Code 0.21 werden die Kompositionen Ref(Global(Num('0'))), Ref(Subscr(Stack(Num('2')), Stack(Num('1')))) und Assign(Stack(Num('1')), Stack(Num('2'))) durch ihre entsprechenden RETI-Knoten ersetzt.

```
1 File
     Name './example_array_assignment.reti_blocks',
     Ε
       Block
         Name 'main.0',
           # // Assign(Subscr(Name('ar'), Num('1')), Num('42'))
           # Exp(Num('42'))
 9
           SUBI SP 1;
10
           LOADI ACC 42;
           STOREIN SP ACC 1;
           # Ref(Global(Num('0')))
           SUBI SP 1;
13
14
           LOADI IN1 0;
           ADD IN1 DS;
16
           STOREIN SP IN1 1;
17
           # Exp(Num('1'))
18
           SUBI SP 1;
19
           LOADI ACC 1;
20
           STOREIN SP ACC 1;
           # Ref(Subscr(Stack(Num('2')), Stack(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(Stack(Num('1')), Stack(Num('2')))
29
           LOADIN SP IN1 1;
30
           LOADIN SP ACC 2;
31
           ADDI SP 2;
           STOREIN IN1 ACC 0;
33
           # Return(Empty())
34
           LOADIN BAF PC -1;
35
         ]
    ]
```

Code 0.21: RETI-Blocks Pass für Zuweisung an Feldindex.

### Literatur

### Vorlesungen

• Scholl, Christoph. "Betriebssysteme". Vorlesung. Vorlesung. Universität Freiburg, 2020. URL: https://abs.informatik.uni-freiburg.de/src/teach\_main.php?id=157 (besucht am 09.07.2022).