Albert Ludwigs Universität Freiburg

TECHNISCHE FAKULTÄT

PicoC-Compiler

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

BACHELORARBEIT

 $Abgabedatum: 28^{th}$ April 2022

 $\begin{array}{c} Author: \\ \text{J\"{u}rgen Mattheis} \end{array}$

Gutachter: Prof. Dr. Scholl

Betreung: M.Sc. Seufert

Eine Bachelorarbeit am Lehrstuhl für Betriebssysteme

ERKLÄRUNG
ERRLARONS
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

Abbilo	dungsverzeichnis	Ι
\mathbf{Codev}	verzeichnis	II
Tabell	lenverzeichnis	II
Defini	tionsverzeichnis	V
Gram	matikverzeichnis	V
	gebnisse und Ausblick	1
1.1	1.1.1 Kommandozeilenoptionen	1
		3 4
1.2	4	6
1.3	Erweiterungsideen	7
Litera	${ m tur}$	\mathbf{A}

Abbildungsverzeichnis		
1.1 Show-Mode in der Verwendung		

Codeverzeichnis 1.1 Shellaufruf und die Befehle compile und quit

Tabellenverzeichnis

Definitionsverzeichnis	

Grammatikverzeichnis	

1 Ergebnisse und Ausblick

Zum Schluss soll ein Überblick über das gegeben werden, was im Kapitel ?? implementiert wurde. In Unterkapitel 1.1 wird mithilfe kurzer Anleitungen ein grober Einblick in die wichtigsten Funktionalitäten des implementierten PicoC-Compilers und anderer mitimplementierter Tools gegeben. Im Unterkapitel 1.2 wird aufgezeigt, was zur Qualitätssicherung implementiert wurde, um zu gewährleisten, dass der PicoC-Compiler die Kompilierung der Programmiersprache L_{PicoC} in Syntax und Semantik identisch zur entsprechenden Untermenge der Programmiersprache L_C umsetzt. Als allerletztes wird im Unterkapitel 1.3 ein Ausblick gegeben, wie der PicoC-Compiler erweitert werden könnte.

1.1 Funktionsumfang

Bei der Implementierung des PicoC-Compilers wurden verschiedene Kommandozeilenoptionen und Modes implemenetiert. Diese werden in den folgenden Kapiteln 1.1.1, 1.1.2 und 1.1.3 mithilfe kurzer Anleitungen erklärt.

Die kurzen Anleitungen in dieser Schrifftlichen Ausarbeitung der Bachelorarbeit sollen nur zu einem schnellen, grundlegenden Verständnis der Verwendung des PicoC-Compilers und seiner Kommandozeilenoptionen und Befehle beihelfen, sowie zum Verständnis der weiteren implementierten Tools. Alle weiteren Kommandozeilenoptionen und Befehle sind für die Verwendung des PicoC-Compilers unwichtig und erweisen sich nur in speziellen Situationen als nütztlich, weshalb für diese auf die ausführlichere Dokumentation unter Link¹ verwiesen wird.

1.1.1 Kommandozeilenoptionen

Will man einfach nur ein Programm program.picoc kompilieren ist das mit dem PicoC-Compiler genauso unkompliziert wie mit dem GCC durch einfaches Angeben der Datei, die kompiliert werden soll:

> picoc_compiler program.picoc

. Als Ergebnis des Kompiliervorgangs wird eine Datei program.reti mit dem entsprechenden RETI-Code erstellt, wobei für die Benennung der Datei einfach nur der Basisname der Datei program an eine neue Dateiendung .reti angehängt wird².

Daneben gibt es allerdings auch die Möglichkeit Kommandozeilenoptionen <cli-options> in der Form

picoc_compiler <cli-options> program.picoc mitanzugeben, von denen die wichtigsten in Tabelle 1.1 erklärt sind. Alle weiteren Kommandozeilenoptionen können in der Dokumenation unter Link nachgelesen werden.

¹ https://github.com/matthejue/PicoC-Compiler/blob/new_architecture/doc/help-page.txt

²Beim GCC wird bei Nicht-Angabe eines Dateinamen mit der -o Option dagegen eine Datei mit der festen Namen a.out

Kommandozeilenoption	Beschreibung	Standardwert
-i,intermediate_stages	Gibt Zwischenschritte der Kompilierung in Form der verschiedenen Tokens, Ableitungsbäume, Abstrakten Syntaxbäume der verschiedenen Passes in Dateien mit entsprechenden Dateiendungen aber gleichem Basinamen aus. Im Shell-Mode erfolgt keine Ausgabe in Dateien, sondern nur im Terminal.	false, most_used: true
-p,print	Gibt alle Dateiausgaben auch im Terminal aus. Diese Option ist im Shell-Mode dauerhaft aktiviert.	false (true im Shell-Mode und für den most_used- Befehl)
-v,verbose	Fügt den verschiedenen Zwischenschritten der Kompilierung, unter anderem auch dem finalen RETI-Code Kommentare hinzu, welche ein Statement oder Befehl aus einem vorherigen Pass beinhalten, der durch die darunterliegenden Statements oder Befehle ersetzt wurde. Wenn dierun-Option aktivert ist, wird der Zustand der virtuellen RETI-CPU vor und nach jedem Befehl angezeigt.	false
-vv,double_verbose	Hat dieselben Effekte, wie die —verbose-Option, aber bewirkt zusätzlich weitere Effekte. PicoC-Knoten erhalten bei der Ausgabe in den Abstrakten Syntaxbäumen zustätzliche runde Klammern, sodass direkter abgelesen werden kann, wo ein Knoten anfängt und wo einer aufhört. In Fehlermeldungen werden mehr Tokens angezeigt, die an der Stelle der Fehlermeldung erwartet worden wären. Bei Aktivierung derintermediate_stages-Option werden in den dadurch ausgegebenen Abstrakten Syntaxbäumen ebenfalls versteckte Attribute, die Informationen zu Datentypen und für Fehlermeldungen beinhalten angezeigt.	false
-h,help	Zeigt die Dokumentation, welche ebenfalls unter Link gefunden werden kann im Terminal an. Mit dercolor-Option kann die Dokumentation mit farblicher Hervorhebung im Terminal angezeigt werden.	false
-R,run	Führt die RETI-Befehle, die das Ergebnis des Kompilierung sind mit einer virtuellen RETI-CPU aus. Wenn dieintermediate_stages-Option aktiviert ist, wird eine Datei basename>.reti_states erstellt, welche den Zustsand der RETI-CPU nach dem letzten ausgeführten RETI-Befehl enthält. Wenn dieverbose-oderdouble_verbose-Option aktiviert ist, wird der Zustand der RETI-CPU vor und nach jedem Befehl auch noch zusätlich in die Datei basename>.reti_states ausgegeben.	false, most_used: true
-B,process_begin	Setzt die relative Adresse, wo der Prozess bzw. das Codesegment für das ausgeführte Programm beginnt.	3
-D,datasegment_size	Setzt die Größe des Datensegments. Diese Option muss mit Vorsicht gesetzt werden, denn wenn der Wert zu niedrig gesetzt wird, dann können die Globalen Statischen Daten und der Stack miteinander kollidieren.	32

Alle kleingeschriebenen Kommandozeilenoptionen, wie -i, -p, -v usw. betreffen dabei den PicoC-Compiler und alle großgeschriebenen Kommandozeilenoptionen, wie -R, -B, -D usw. betreffen den RETI-Interpreter.

1.1.2 Shell-Mode

Will man z.B. eine Folge von Statements in der Programmiersprache L_{PicoC} schnell kompilieren ohne eine Datei erstellen zu müssen, so kann der PicoC-Compiler im sogenannten Shell-Mode aufgerufen werden. Hierzu wird der PicoC-Compiler ohne Argumente \rightarrow picoc_compiler aufgerufen, wie es in Code 1.1 zu sehen ist. Die angegebene Folge von Statements \leftarrow wird dabei automatisch in eine main-Funktion eingefügt: void main() \leftarrow seq-of-stmts \rightarrow .

Mit dem **> compile <cli-options> <filename>** Befehl (oder der Abkürzung cpl) kann PicoC-Code zu RETI-Code kompiliert werden. Die Kommandozeilenoptionen <cli-options> sind dieselben, wie wenn der Compiler direkt mit Kommandozeilenoptionen aufgerufen wird. Die wichtigsten dieser Kommandozeilenoptionen sind in Tabelle 1.1 angegeben.

Mit dem Befehl > quit kann der Shell-Mode wieder verlassen werden.

```
> picoc_compiler
PicoC Shell. Enter `help` (shortcut `?`) to see the manual.
PicoC> cpl "6 * 7;";
                   ----- RETI -----
SUBI SP 1;
LOADI ACC 6;
STOREIN SP ACC 1;
SUBI SP 1;
LOADI ACC 7;
STOREIN SP ACC 1;
LOADIN SP ACC 2;
LOADIN SP IN2 1;
MULT ACC IN2;
STOREIN SP ACC 2;
ADDI SP 1;
LOADIN BAF PC -1;
Compilation successfull
PicoC> quit
```

Code 1.1: Shellaufruf und die Befehle compile und quit

Wenn man möglichst alle nützlichen Kommandozeilenoptionen direkt aktiviert haben will, bei denen es keinen Grund gibt, sie nicht mitanzugeben, kann der Befehl > most_used <cli-options> <filename> (oder seine Abkürzung mu) genutzt werden, um diese Kommandozeilenoptionen mit dem compile-Befehl nicht jedes mal selbst Angeben zu müssen. In der Tabelle 1.1 sind in grau die Werte der einzelnen Kommandozeilenoptionen angegeben, die bei dem Befehl most_used gesetzt werden. In Code 1.2 ist der most_used-Befehl in seiner Verwendung zu sehen.

Dadurch, dass die --intermediate_stages- und die --run-Option beim most_used-Befehl aktiviert sind, werden die verschiedenen Zwischenstufen der Kompilierung, wie Tokens, Derivation Tree usw., sowie der Zustand der RETI-CPU nach der Ausführung des letzten Befehls angezeigt. Aus Platzgründen ist das meiste allerdings mit '...' ausgelassen.

```
PicoC> mu "int var = 42;";
         ----- Code -----
// stdin.picoc:
void main() {int var = 42;}
----- Tokens
  ----- Derivation Tree ------
  ----- Derivation Tree Simple -----
  ----- Abstract Syntax Tree ------
  ----- PicoC Shrink -----
   ----- PicoC Blocks -----
   ----- PicoC Mon -----
    ------ Symbol Table ------
   ------ RETI Blocks ------
----- RETI Patch -----
----- RETI -----
SUBI SP 1;
LOADI ACC 42;
STOREIN SP ACC 1;
LOADIN SP ACC 1;
STOREIN DS ACC 0;
ADDI SP 1;
LOADIN BAF PC -1;
         ----- RETI Run -----
Compilation successfull
```

Code 1.2: Shell-Mode und der Befehl most_used

Im Shell-Mode kann der Cursor mit den ← und → Pfeiltasten bewegt werden. In der Befehlshistorie kann sich mit den ↑ und ↓ Pfeiltasten rückwarts und vorwärts bewegt werden. Mit Tab kann ein Befehl automatisch vervollständigt werden.

Es gibt für den Shell-Mode noch weitere Befehle, wie color_toggle, history etc. und kleinere Funktionalitäten für die Shell, die sich in der ein oder anderen Situation als nützlich erweisen können. Für die Erklärung dieser wird allerdings auf die Dokumentation unter Link verwiesen, welche auch über den Befehl help angezeigt werden kann.

1.1.3 Show-Mode

Der Show-Mode ist ein Nebenprodukt der Implementierung des PicoC-Compilers. Dieser Mode wurde eigentlich nur implementiert, um beim Testen des PicoC-Compilers Bugs bei der Generierung des RETI-Code zu finden, indem im Terminal eine virtuelle RETI-CPU angezeigt wird, welches den kompletten

Zustand einer virtuell ausgeführten RETI mit allen Registern, SRAM, UART, EPROM und einigen weiteren Informationen anzeigt.

Allerdings bringt die Möglichkeit des Show-Mode, die RETI-Befehle des übersetzten Programmes in Ausführung zu sehen auch einen großen Lerneffekt mit sich, weshalb der Show-Mode noch weiterentwickelt wurde, sodass auch Studenten ihn auf unkomplizierte Weise nutzen können.

Der Show-Mode kann auf die einfachste Weise mittels der /Makefile des PicoC-Compilers mit dem Befehl make show FILEPATH=<path-to-file> <more-options> gestartet werden. Alle einstellbaren Optionen, die z.B. unter <more-options> noch für die Makefile gesetzt werden können sind in Tabelle 1.2 aufgelistet.

Kommandozeilenoption	Beschreibung	Standardwert
FILEPATH	Pfad zur Datei, die im Show-Mode angezeigt werden soll	Ø
TESTNAME	Name des Tests. Alles andere als der Basisname, wie die Dateiendung wird abgeschnitten	Ø
EXTENSION	Dateiendung, die an TESTNAME angehängt werden soll zu ./tests/TESTNAME.EXTENSION	reti_states
NUM_WINDOWS	Anzahl Fenster auf die ein Dateiinhalt verteilt werden soll	5
VERBOSE	Möglichkeit die Kommandozeilenoption -v oder -vv zu aktivieren für eine ausführlichere Ausgabe	Ø
DEBUG	Möglichkeit die Kommandozeilenoption -d zu aktivieren, um bei make test-show TESTNAME= <testname> den Debugger für den entsprechenden Test <testname> zu starten</testname></testname>	Ø

Tabelle 1.2: Makefileoptionen

Alternativ kann der Show-Mode mit dem Befehl make test-show TESTNAME=<testname> <more-options> auch für einen der geschriebenen Tests im Ordner /tests gestartet werden. Der Test wird bei diesem Befehl erst ausgeführt und dann der Show-Mode gestartet.

Der Show-Mode nutzt den Terminal Texteditor Neovim³ um einen Dateiinhalt über mehrere Fenster verteilt anzuzeigen, so wie es in Abbildung 1.1 zu sehen ist. Für den Show-Mode wird eine eigene Konfiguration für Neovim verwendet, welche in der Konfigurationsdatei /interpr_showcase.vim spezifiziert ist.

Gedacht ist der Show-Mode vor allem dafür etwas ähnliches wie ein RETI-Debugger zu sein und wird daher standardmäßig bei Nicht-Angabe einer EXTENSION auf die Datei program>.reti_states angewandt. Der Show-Mode kann aber auch dazu genutzt werden andere Dateien, welche verschiedene Zwischenschritte der Kompilierung darstellen anzuzeigen, indem EXTENSION auf eine andere Dateiendung gesetzt wird.

Im Show-Mode wird ein Trick angewandet, indem die verschiedenen Zustände der RETI-CPU nicht zur Laufzeit des Show-Mode berechnet werden, sondern schon berechnet wurden und nacheinander in die Datei \program>.reti_states ausgegeben wurden. Der Show-Mode macht nichts anderes, als immer an die Stelle zu springen, an welcher der nächste Zustand anfängt. Durch Drücken von Tab und \(\frac{1}{2}\)—Tab können auf diese Weise die verschiedenen Zuständen der RETI-CPU vor und nach der Ausführung eines Befehls angezeigt werden.

³Home - Neovim.

```
struction: ADDI SP 1;
                                                                                              020 LOADIN SP ACC 1;
021 STOREIN DS ACC 0;
                                                                                                                                                                                                                                                                                  LOADIN SP ACC 1;
                                                                                                                                                                                     059 LOADIN SP ACC 1;
                                                                                                                                                                               00059 LOADIN SP ACC 1;
00060 ADDI SP 1;
00061 CALL PRINT ACC;
00062 SUBI SP 1;
00063 LOADI ACC 0;
00064 STOREIN SP ACC 1;
       STMPLE:
N1 SIMPLE:
                                                                                                                                                                                                                                                                     00100 JUMP -32;
00101 SUBI SP 1;
00102 LOADIN DS ACC 0;
                                                                                             0024 LOADIN DS ACC 0:
                                                                                        00024 LOADIN DS ACC 0;
00025 STOREIN SP ACC 1;
00026 SUBI SP 1;
00027 LOADI ACC 4;
00028 STOREIN SP ACC 1;
00029 LOADIN SP ACC 2;
00030 LOADIN SP IN2 1;
                                                                                                                                                                               00065 LOADIN SP ACC 1;
00066 STOREIN DS ACC 0;
00067 ADDI SP 1;
00068 SUBI SP 1;
00068 LOADIN DS ACC 0;
00070 STOREIN SP ACC 1;
                             2147483686
                                                                                                                                                                                                                                                                     00103 STOREIN SP ACC 1:
     STMPLE:
                                                                                         00030 LOADIN SP IN.
00031 SUB ACC IN2;
00032 JUMP< 3;
00033 LOADI ACC 0;
                              2147483650
                                                                                                                                                                               00071 SUBI SP 1;
00072 LOADI ACC 2;
00073 STOREIN SP ACC 1;
                                                                                                                                                                                                                                                                     00109 STOREIN SP ACC 1:
                                                                                         00036 STOREIN SP ACC 2;
00037 ADDI SP 1;
00038 LOADIN SP ACC 1; <- PC
                                                                                                                                                                                                                                                                     00112 CALL PRINT ACC;
00113 LOADIN BAF PC -
00114 3 <- DS
     SIMPLE:
                                                                                                                                                                               00074 LOADIN SP ACC 2;
00075 LOADIN SP IN2 1;
                                                                                                                                                                               00076 SUB ACC IN2:
                                                                                                                                                                                                                                                                                                                                                            00000 LOADI DS -2097152: <- IN
  00001 2147483648
00002 0 <- BAF
                                                                                                                                                                               00077 JUMP< 3;
00078 LOADI ACC 0;
                                                                                                                                                                                                                                                                                                                                                           00001 MULTI DS 1024; <-
00002 MOVE DS SP;
                                                                                         00039 ADDI SP 1;
   00001 2147433643
00002 0 <- BAF
00003 CALL INPUT ACC; <- CS
                                                                                        00040 JUMP== 2;
00041 JUMP -32;
                                                                                                                                                                              00079 JUMP 2;
00079 JUMP 2;
00080 LOADI ACC 1;
00081 STOREIN SP ACC 2;
00082 ADDI SP 1;
00083 LOADIN SP ACC 1;
  98093 CALL INPUT ACC; <-
69004 SUBI SP, C1
69005 STOREIN SP ACC 1;
98006 LOADIN SP ACC 1;
98006 LOADIN SP ACC 0;
98008 ADDI SP 1;
98019 SUBI SP 1;
98010 LOADIN DS ACC 0;
98011 STOREIN SP ACC 1;
98012 SUBI SP 1;
98013 LOADI ACC 1;
98014 STOREIN SP ACC 1;
98015 LOADIN SP ACC 2;
                                                                                                                                                                                                                                                                                                                                                            00003 MOVE DS BAF:
                                                                                        00041 JUMP -32;
000042 SUBI SP 1;
00043 LOADIN DS ACC 0;
00044 STOREIN SP ACC 1;
00045 SUBI SP 1;
00046 LOADI ACC 2;
00047 STOREIN SP ACC 1;
                                                                                                                                                                                                                                                                                                                                                            00004 MOVE DS CS; <- IN2
00005 ADDI SP 145;
00006 ADDI BAF 2;
                                                                                                                                                                                                                                                                                                                                                            00007 ADDI CS 3;
00008 ADDI DS 114;
                                                                                                                                                                               00085 JUMP== 16;
00086 SUBI SP 1;
00087 LOADIN DS ACC 0;
00088 STOREIN SP ACC 1;
                                                                                         00048 LOADIN SP ACC 2;
00049 LOADIN SP IN2 1;
                                                                                                                                                                                                                                                                                                                                                         nstruction: LOADIN SP ACC 1:
                                                                                              050 SUB ACC IN2:
                                                                                              051 STOREIN SP ACC 2;
052 ADDI SP 1;
                                                                                                                                                                                                                                                                                                                                                            C_SIMPLE:
   00015 LOADIN SP ACC 2;
00016 LOADIN SP IN2 1;
00017 ADD ACC IN2;
                                                                                          00053 LOADIN SP ACC 1;
                                                                                                                                                                                     91 STOREIN SP ÁCC 1:
                                                                                                                                                                                           LOADIN SP ACC 2;
LOADIN SP IN2 1;
                                                                                                                                                                                                                                                                                                                                                         N1 SIMPLE:
                                                                                                                                                                                                                                                                                                                                                         N2 SIMPLE:
                  STOREIN SP ACC 2:
```

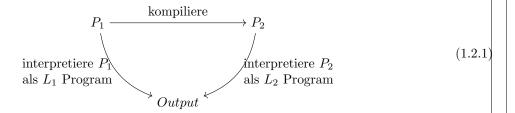
Abbildung 1.1: Show-Mode in der Verwendung

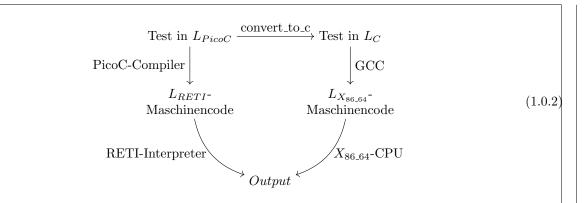
Zur besseren Orientierung wird für alle Register ebenfalls ein mit der Registerbezeichnung beschriffteter Zeiger <- REG an Adressen im EPROM, UART und SRAM angezeigt, je nachdem, ob der Wert im Register nach der Memory Map dem Adressbereich von EPROM, UART oder SRAM entspricht.

Durch Drücken von Esc oder q kann der Show-Mode wieder verlassen werden. Es gibt für den Show-Mode noch viele weitere Tastenkürzel, die sich in der ein oder anderen Situation als nützlich erweisen können. Für die Erklärung dieser wieder allerdings auf die Dokumentation unter Link verwiesen. Des Weiteren stehen durch die Nutzung des Terminal Texteditors Neovim auch alle Funktionalitäten dieses mächtigen Terminal Texteditors zur Verfügung, welche mittels der Eingabe von :help nachgelesen werden können oder mittels der Eingabe von :Tutor mithilfe einer kurzen Einführungsanleitung erlernt werden können.

1.2 Qualitätssicherung

Die Qualität des PicoC-Compilers wird





1.3 Erweiterungsideen

- •4 4	
Literatur	
Online	
• Home - Neovim. URL: http://neovim.io/ (besucht am 04.08.2022).	