

Objektorientierte Programmierung, Übungsaufgabe 2

## **C-Programm komprimieren**

Jan Tammen <foobar@fh-konstanz.de>

22. April 2005

---

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>3</b>
<b>2</b>	<b>Entwurf eines Automaten</b>	<b>3</b>
2.1	Automatengraph . . . . .	3
2.2	Automatentabelle . . . . .	4
2.3	Aktionen, Ausgaben . . . . .	5
2.4	Steuermatrix . . . . .	6

---

# 1 Einleitung

Ein (C-/C++-)Quelltextprogramm soll durch die Entfernung von Kommentaren und Weißzeichen komprimiert werden. Dabei müssen Strings, Zeichenkonstanten, Zahlen und Schlüsselwörter erhalten bleiben. Gleichzeitig werden alle erkannten Teile zeilenweise in einer neuen Datei abgelegt.

## 2 Entwurf eines Automaten

Das abstrahierte Problem lässt sich zunächst durch einen endlichen Automaten darstellen, welcher jeweils ein Zeichen des Quelltextes einliest und dann durch Übergang in Folgezustände und Durchführung von Aktionen den Quelltext komprimiert.

### 2.1 Automatengraph

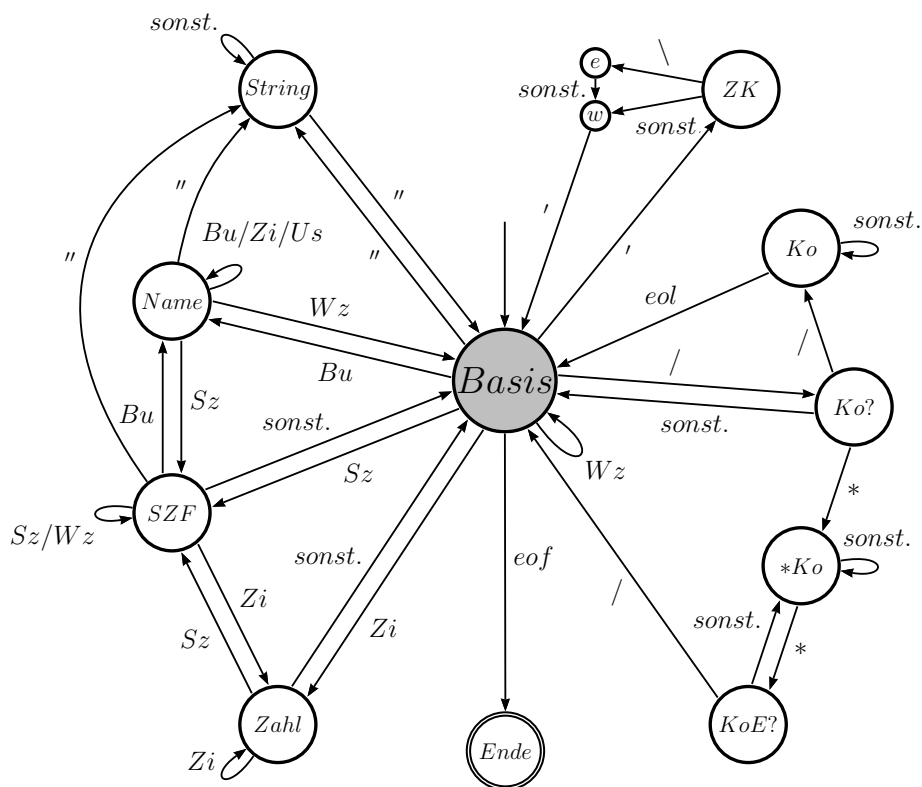


Abbildung 1: Automatengraph

## 2.2 Automatentabelle

Die Automatentabelle führt die Zustände sowie das Verhalten bei Eingabezeichen auf.

Im Zustand	bei Zeichen	Aktion	Zeichen lesen	Folgezustand
Basis	,	—	ja	ZK
	/	—	ja	Ko?
	”	—	ja	String
	Wz (Weißzeichen)	—	ja	Basis
	Zi (Ziffer)	Zi merken	ja	Zahl
	Sz (Sonderzeichen)	Sz merken	ja	SZF
	Bu (Buchstabe)	Bu merken	ja	Name
	eof (end of file)	—	nein	—
ZK	\	—	ja	ZKescape
	sonst.	Zeichen merken	ja	ZKwert
ZKescape	sonst.	Zeichen merken	ja	ZKwert
ZKwert	,	Ausgabe ZK	ja	Basis
Ko?	sonst.	—	ja	Basis
	/	—	ja	Ko
	*	—	ja	*Ko
Ko	sonst.	—	ja	Ko
	eol	—	ja	Basis
*Ko	sonst.	—	ja	*Ko
	*	—	ja	KoE?
KoE?	sonst.	—	ja	*Ko
	/	—	ja	Basis
Zahl	Zi	Zi dazumerken	ja	Zahl
	Sz	Zi ausgeben, Sz merken	ja	SZF
	sonst.	Zahl ausgeben	ja	Basis
SZF	Sz	Zeichen dazumerken	ja	SZF
	Wz	—	ja	SZF
	Bu	SZF ausgeben, Bu merken	ja	Name
	Zi	SZF ausgeben, Zi merken	ja	Zahl
	”	SZF ausgeben	ja	String
	sonst.	SZF ausgeben	ja	Basis
Name	Bu/Zi/Us	Zeichen dazumerken	ja	Name
	Sz	Name ausgeben, Sz merken	ja	SZF
	Wz	Name ausgeben	ja	Basis
	”	Name ausgeben	ja	String
String	”	String ausgeben	ja	Basis
	sonst.	Zeichen merken	ja	String

Tabelle 1: Automatentabelle

Um das komprimierte Quellprogramm in die Ausgabedatei zu schreiben, wurde der Automat nachfolgend um entsprechende **Aktionen** erweitert.

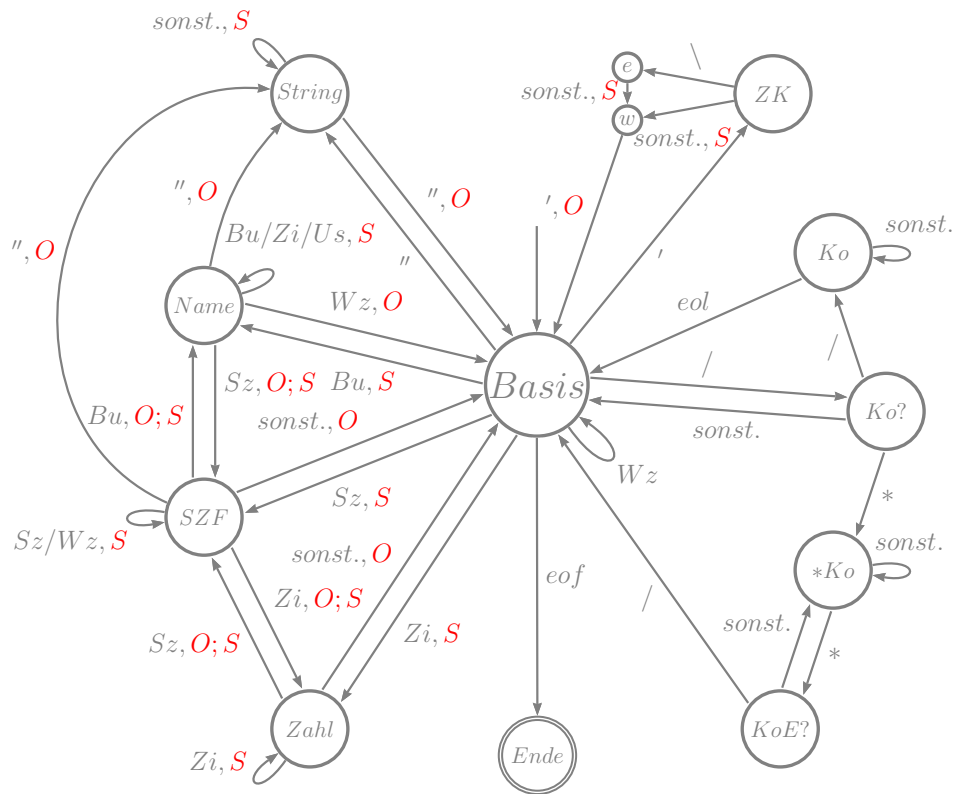


Abbildung 2: Erweiterter Automatengraph

Die benötigten Aktionen sind in nachstehender Tabelle benannt und kurz beschrieben.

Kurzname	Aktion	Beschreibung
STO (S)	Zeichen merken	Speichert aktuelles Zeichen (Ziffer, Sonderzeichen, etc.) zwischen.
ADD (A)	Zeichen dazumerken	Fügt Zeichen an den Zwischenspeicher an.
OUT (O)	Ausgabe $XY$	Gibt gespeicherte Zeichenkette + Newline (auf den Dateistrom) aus.

Tabelle 2: Aktionen

Es werden also lediglich *drei Grundoperationen* benötigt, wobei die Aktionen (S) und (A) programmiertechnisch später vermutlich noch weiter zusammengefasst werden können.

---

## 2.4 Steuermatrix

Für die Umsetzung im Programm wird die Automatentabelle in eine Steuermatrix transformiert. Die Spalten beinhalten die Zeichen bzw. –gruppen, die Zeilen die Zustände. Die Zeilen sind in zwei Bereiche unterteilt: oben ist der Folgezustand eingetragen, unten die Aktion.

Es werden folgende Eingabezeichen zu Gruppen zusammengefasst:

- Buchstaben: a-zA-Z
- Ziffern: 0-9
- Sonderzeichen: + - \* ; , [ ] ( ) { } \
- sonst.: Bezeichnet die „restlichen“, nicht näher zu spezifizierenden Eingaben

Von Zustand	Bei Zeichen / -gruppe										
	Bu	Zi	Sz	Wz	'	”	/	*	\	sonst.	eol
<b>Basis</b>	Name STO	Zahl STO	SZF STO	Basis —	ZK —	String —	Ko? —	SZF STO	SZF STO	— —	— —
<b>ZK</b>	— —	— —	— —	— —	— —	— —	— —	— —	ZKe —	ZKw STO	— —
<b>ZKe</b>	— —	— —	— —	— —	— —	— —	— —	— —	— —	ZKw STO	— —
<b>ZKw</b>	— —	— —	— —	— —	Basis OUT	— —	— —	— —	— —	— —	— —
<b>Ko?</b>	— —	— —	— —	— —	— —	— —	Ko —	*Ko —	— —	Basis —	— —
<b>Ko</b>	— —	— —	— —	— —	— —	— —	— —	— —	— —	Ko —	Basis —
<b>*Ko</b>	— —	— —	— —	— —	— —	— —	— —	KoE? —	— —	*Ko —	— —
<b>KoE?</b>	— —	— —	— —	— —	— —	— —	Basis —	— —	— —	*Ko —	— —
<b>Zahl</b>	— —	Zahl ADD	SZF O; S	— —	— —	— —	— —	— —	— —	Basis OUT	— —
<b>SZF</b>	Name O; S	Zahl O; S	SZF ADD	SZF —	— —	String OUT	— —	— —	— —	Basis OUT	— —
<b>Name</b>	Name ADD	Name ADD	SZF O; S	Basis OUT	— —	String OUT	— —	— —	— —	— —	— —
<b>String</b>	— —	— —	— —	— —	— —	Basis OUT	— —	— —	— —	String ADD	— —

Tabelle 3: Steuermatrix