Im Inneren der CPU

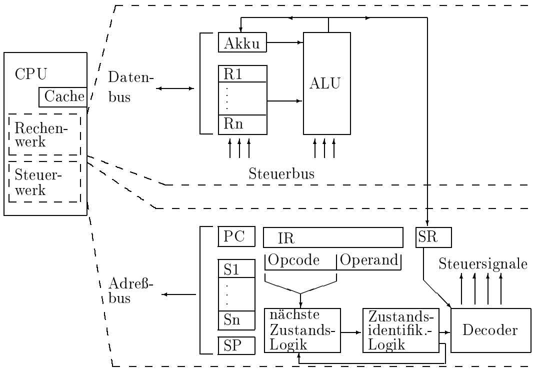

Abbildung 1.5: CPU

Programmausführung:

```
PC := StartingAddress;
RunBit := 1;
while RunBit = 1 do begin
  IR := memory(PC);   PC := PC + Befehlslänge(IR);
                                 (häufig=1, fest)
  InstrType := DetermineInstrType(IR) ;
  DataAddress := LocateData(IR,InstrType,PC);
  if DataAddress <> nil then
    data := memory(DataAddress);
  else
    data := non_valid;
  Execute(IR,InstrType,data,memory,registers,PC,RunBit);
end;
```

Ein von-Neumann-Rechner ist durch folgende Merkmale gekennzeichnet (zu 6.-8. vgl. auch Kapitel 1.2):

Prinzipien:

1. **Computerbestandteile:** CPU (= Rechenwerk und Steuerwerk), Speicher, Ein- und Ausgabewerk.
2. **Universalität:** Spezialisierung nur durch von außen eingegebenes Programm.
3. **Programme als Daten:** Eingabedaten, Zwischen- und Endergebnisse sowie Programme werden im selben Speicher abgelegt.
4. **Linearer Speicher:** Der Speicher ist in gleichgroße Zellen, die fortlaufend numeriert sind, eingeteilt.
5. **Sequentielle Ausführung:** Aufeinanderfolgende Befehle eines Programms werden in aufeinanderfolgenden Speicherzellen abgelegt. Beim Programmablauf wird im allgemeinen der Befehlszähler fortlaufend inkrementiert.
6. **Sprünge:** Sprungbefehle ermöglichen ein Durchbrechen der linearen Ausführungsreihenfolge.
7. **Zusätzliche Primitiva:** Arithmetische Operationen, logische und Vergleichsoperationen, Graphik-/Multimediabefehle; indizierte, relative, ... Adressierung.
8. **Binäre Codierung:** Daten (Befehle, Adressen, Operanden) werden binär codiert. Geeignete Schaltwerke (Decodierer) sorgen für die richtige Entschlüsselung.

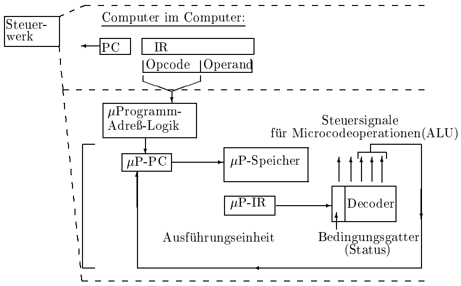

Abbildung 1.6: CISC = Microprogrammierte CPU

Das Konzept des *Zusammenbaus* komplexer Aufgaben aus einer Sequenz primitiver Aktionen war so erfolgreich, daß in Zeiten, als die Datentransferzeit zwischen Speicher und CPU im Vergleich zur Ausführungszeit eines Primitivbefehls sehr groß war (vor $\sim$ 1970), das Rechenwerk wiederum als (Mikro-)Computer aufgebaut wurde: Die Software reduziert den Schaltungsaufwand (=Anzahl der Gatter, Anzahl der Leitungen, ...) und „interpretiert“ den Maschinencode der CPU (Computer im Computer). Heute wird für leistungsfähige PC's und Workstations das Rechenwerk wieder direkt in Hardware aufgebaut:

RISC (= reduced instruction set computer)-Designprinzip :

- Analysiere Anwendungen, um die Schlüsseloperationen zu finden.
- Reduziere unter allen Umständen die „Datenpfadzykluszeit“ (Register $\to$ ALU $\to$ Register) für diese Operationen, also: *keine* Microcode-Interpretation (Maschinencode=Microcode).
- Jedes „neue“ Feature ist nur zuzulassen, wenn es häufig benutzt wird und die Mehrzahl der existierenden anderen Feature nicht verlangsamt.

(Wiederhole diesen Design-Zyklus für die anderen CPU-Ressourcen: Cache, Speichermanagement, Gleitkomma-Koprozessoren, ...)

Tabelle 1.2: *Merkmale der 1. bis 3. Computergeneration*

	1. Generation ab 1951	**2. Generation** ab 1960	**3. Generation** ab 1965
Basiselement	Elektronenröhre	Flächentransistor	integrierter Schaltkreis
Arbeitsspeicher	Magnettrommel-, Umlaufspeicher	Ferritkernspeicher	Mikroferritkernspeicher
Externe Speicher	keine	Magnetbandspeicher, -trommelspeicher	Magnetband-, Wechselplattenspeicher
Ein- und Ausgabegeräte	Lochkarten- , Lochstreifengeräte, elektrische Schreibmaschine	wie 1. Generation sowie Walzendrucker	wie 2. Generation sowie Bildschirmgeräte, Datenfernübertragungsgeräte
Programmierung	Maschinencode, Assemblersprache	Assemblersprache, erste problemorientierte Sprachen, einfache Betriebssysteme	Assemblersprache, universelle Sprachen, komplexe Betriebssysteme
Einsatz als	wissenschaftlich-technische Rechner	wissenschaftlich-technische, kommerzielle und Prozeßrechner	universelle und Prozeßrechner, Einbaurechner

Quelle: D.Werner (Hrsg.): Taschenbuch der Informatik; Fachbuchverlag Leipzig, 1995

Merkmale der 4. Generation (heute):

- VLSI-Schaltkreise
- in wenigen Jahren sich verdoppelnde Rechengeschwindigkeiten, Speichergrößen, Übertragungsgeschwindigkeiten ($\to$ in 4 Jahren ist ein Computersystem „total“ veraltet)
- Software-Entwicklungskosten höher als Hardwarekosten; wegen der schnellen Hardware-Innovationszyklen sind meist „nur“ noch portable Software-Produkte schnell genug verfügbar (UNIX, TEX, ...)
- GUI's = *graphical user interfaces*
- Computer werden immer mehr zu IT-Geräten (Datennetzwerke)
- Laserdrucker, fotorealistischer Druck, Scanner mit 24Bit Farbtiefe

5. Generation: Für nach 2000 angestrebte Merkmale:

- natürlichsprachiger Umgang mit Computern
  - Handschrift

– Sprache

- Multimedia, Telekonferenzen, Unterhaltung, Teleshoping, T-online, WWW, Heimarbeitsplätze, Fernkurse
- Methoden der Künstlichen Intelligenz (KI)
  - Zeichen-, Bild-, Sprach-, Schrift-, Bewegungserkennung
  - simultane automatische Sprachübersetzung
  - Wissensverarbeitung (automatische Krankheitsdiagnose, …)
  - Nutzeridentifizierung durch Iris-Mustererkennung ...

1.2 Datenformate, Maschinenbefehle und Adressierungsarten

1.2.1 Schrift / Zeichen

1.2.1.1 Zeichensätze

Der ASCII-Code (*American Standard Code for Information Interchange*) ist auch heute noch Grundlage vieler Zeichencodes:

00	NUL	01	SOH	02	STX	03	ETX	04	EOT	05	ENQ	06	ACK	07	BEL
08	BS	09	HT	0A	NL	0B	VT	0C	NP	0D	CR	0E	SO	0F	SI
10	DLE	11	DC1	12	DC2	13	DC3	14	DC4	15	NAK	16	SYN	17	ETB
18	CAN	19	EM	1A	SUB	1B	ESC	1C	FS	1D	GS	1E	RS	1F	US
20	SP	21	!	22	"	23	#	24	$	25	%	26	&	27	´
28	(	29	)	2A	*	2B	+	2C	,	2D	-	2E	.	2F	/
30	0	31	1	32	2	33	3	34	4	35	5	36	6	37	7
38	8	39	9	3A	:	3B	;	3C	<	3D	=	3E	>	3F	?
40	@	41	A	42	B	43	C	44	D	45	E	46	F	47	G
48	H	49	I	4A	J	4B	K	4C	L	4D	M	4E	N	4F	O
50	P	51	Q	52	R	53	S	54	T	55	U	56	V	57	W
58	X	59	Y	5A	Z	5B	[	5C	\	5D	]	5E	^	5F	_
60	`	61	a	62	b	63	c	64	d	65	e	66	f	67	g
68	h	69	i	6A	j	6B	k	6C	l	6D	m	6E	n	6F	o
70	p	71	q	72	r	73	s	74	t	75	u	76	v	77	w
78	x	79	y	7A	z	7B	{	7C	$\vert$	7D	}	7E	~	7F	DEL

Tabelle 1.3: ASCII-Code

ISO Nr.	Zeichensatz	Dezimalform												
		ID	35	36	64	91	92	93	94	96	123	124	125	126
6	ANSI ASCII	0U	#	$	@	[	\	]	ˆ	`	{	$\vert$	}	˜
11	Schweden: Namen	0S	#		É	Ä	Ö	Å	Ü	`	ä	ö	å	ü
10	Schweden	3S	#		@	Ä	Ö	Å	ˆ	`	ä	ö	å	
17	Spanien	2S	£	$	§	¡	Ñ	¿	ˆ	`	°	ñ	ç	˜
69	Frankreich	1F	£	$	à	°	ç	§	ˆ	μ	é	ù	è	
21	Deutschland	1G	#	$	§	Ä	Ö	Ü	ˆ	`	ä	ö	ü	ß
4	Großbritannien	1E	£	$	@	[	\	]	ˆ	`	{	$\vert$	}	
16	Portugal	4S	#	$	§	Ã	Ç	Õ	ˆ	`	ã	ç	õ	°
60	Norwegen 1	0D	#	$	@	Æ	Ø	Å	ˆ	`	æ	ø	å	
61	Norwegen 2	1D	§	$	@	Æ	Ø	Å	ˆ	`	æ	ø	å	$\vert$
2	IRV		#		@	[	\	]	ˆ	`	{	$\vert$	}	
15	Italien	0I	£	$	§	°	ç	é	ˆ	`	à	ò	è	ì

Tabelle 1.4: ISO-Austauschtabelle

Da sieben Bit, also 128 Zeichen nicht für landesspezifische Sonderzeichen ausreichten, entstanden die landesspezifischen Varianten durch Zeichenersetzung (vgl. Tabelle 1.4). Alternativ wurden acht Bit (256 Zeichen) für landesspezifische Sonderzeichen, mathematische Symbole, graphische Symbole zum Tabellendruck bzw. für Sonderzwecke (Spiele, ...) besetzt, etwa im Industriestandard PC-8 Zeichensatz (Tabelle 1.5)[1]. In Windows 3.x wurden jedoch leider andere Codierungen genutzt (Tabelle 1.6). Der inzwischen verabschiedete Standard der *International Standardization Organisation* (ISO), der ISO-8859 Latin 1(ECMA-94 Latin 1) Zeichensatz (Tabelle 1.7) setzte sich insbesondere bei Workstations und neueren Hard- und Softwareprodukten durch. Neben der Latin-1 Version existieren auch noch einige andere nationale Sonderformen des ISO 8859 Codes (vgl. Tabelle 1.8).

[1] In dieser – wie auch in einigen der folgenden Tabellen – sind leere Codestellen entweder unbesetzt oder wegen Problemen beim Satz der entsprechenden Zeichen in diesem Skript freigelassen worden.

0	16	32	0 48	@ 64	P 80	‘ 96	p 112	Ç 128	É 144	á 160	176	192	208	α 224	≡ 240
1	17	! 33	1 49	A 65	Q 81	a 97	q 113	ü 129	æ 145	í 161	177	193	209	ß 225	± 241
2	↕ 18	” 34	2 50	B 66	R 81	b 98	r 114	é 130	Æ 146	ó 162	178	⊤ 194	210	Γ 226	≥ 242
3	!! 19	# 35	3 51	C 67	S 82	c 99	s 115	â 131	ô 147	ú 163	$\vert$ 179	⊢ 195	211	π 227	≤ 243
4	¶ 20	$ 36	4 52	D 68	T 83	d 100	t 116	ä 132	ö 148	ñ 164	⊣ 180	196	212	∑ 228	244
♣ 5	§ 21	% 37	5 53	E 69	U 84	e 101	u 117	à 133	ò 149	Ñ 165	181	197	213	σ 229	245
♠ 6	22	& 38	6 54	F 70	V 85	f 102	v 118	å 134	û 150	^{a} 166	182	198	214	μ 230	÷ 246
7	23	’ 39	7 55	G 71	W 86	g 103	w 119	ç 135	ù 151	^{o} 167	183	199	215	τ 231	≈ 247
8	↑ 24	(40	8 56	H 72	X 87	h 104	x 120	ê 136	ÿ 152	¿ 168	184	200	216	Φ 232	° 248
○ 9	↓ 25	) 41	9 57	I 73	Y 88	i 105	y 121	ë 137	Ö 153	169	185	201	217	Θ 233	· 249
10	→ 26	* 42	: 58	J 74	Z 89	j 106	z 122	è 138	Ü 154	¬ 170	$\Vert$ 186	202	218	Ω 234	· 250
11	← 27	+ 43	; 59	K 75	[90	k 107	{ 123	ï 139	ç 155	171	187	203	219	δ 235	√ 251
12	28	, 44	< 60	L 76	\ 91	l 108	$\vert$ 124	î 140	£ 156	172	188	204	220	∞ 236	^{n} 252
13	↔ 29	- 45	= 61	M 77	] 92	m 109	} 125	ì 141	157	¡ 173	189	205	221	ϕ 237	^{2} 253
14	30	. 46	> 62	N 78	ˆ 93	n 110	˜ 126	Ä 142	Pt 158	≪ 174	190	206	222	ϵ 238	254
15	31	/ 47	? 63	O 79	_ 94	o 111	127	Å 143	ƒ 159	≫ 175	191	207	223	∩ 239	255

Tabelle 1.5: PC-8 Zeichensatz

NUL 0	16	32	0 48	@ 64	P 80	` 96	p 112	128	144	160	° 176	À 192	208	à 224	240
1	17	! 33	1 49	A 65	Q 81	a 97	q 113	129	‘ 145	¡ 161	± 177	Á 193	Ñ 209	á 225	ñ 241
2	18	” 34	2 50	B 66	R 81	b 98	r 114	130	’ 146	ç 162	^{2} 178	Â 194	Ò 210	â 226	ò 242
3	19	# 35	3 51	C 67	S 82	c 99	s 115	131	147	£ 163	^{3} 179	Ã 195	Ó 211	ã 227	ó 243
4	20	$ 36	4 52	D 68	T 83	d 100	t 116	132	148	164	´ 180	Ä 196	Ô 212	ä 228	ô 244
5	21	% 37	5 53	E 69	U 84	e 101	u 117	133	149	165	μ 181	Å 197	Õ 213	å 229	õ 245
6	22	& 38	6 54	F 70	V 85	f 102	v 118	134	150	$\vert$ 166	¶ 182	Æ 198	Ö 214	æ 230	ö 246
BEL 7	23	′ 39	7 55	G 71	W 86	g 103	w 119	135	151	§ 167	· 183	Ç 199	× 215	ç 231	÷ 247
BS 8	24	(40	8 56	H 72	X 87	h 104	x 120	136	152	¨ 168	184	È 200	Ø 216	è 232	ø 248
HT 9	25	) 41	9 57	I 73	Y 88	i 105	y 121	137	153	© 169	^{1} 185	É 201	Ù 217	é 233	ù 249
LF 10	26	* 42	: 58	J 74	Z 89	j 106	z 122	138	154	^{a} 170	^{o} 186	Ê 202	Ú 218	ê 234	ú 250
VT 11	ESC 27	+ 43	; 59	K 75	[90	k 107	{ 123	139	155	≪ 171	≫ 187	Ë 203	Û 219	ë 235	û 251
FF 12	28	, 44	< 60	L 76	\ 91	l 108	$\vert$ 124	140	156	¬ 172	$\frac{1}{4}$ 188	Ì 204	Ü 220	ì 236	ü 252
CR 13	29	- 45	= 61	M 77	] 92	m 109	} 125	141	157	- 173	$\frac{1}{2}$ 189	Í 205	Ý 221	í 237	ý 253
SO 14	30	. 46	> 62	N 78	ˆ 93	n 110	˜ 126	142	158	174	$\frac{3}{4}$ 190	Î 206	222	î 238	254
SI 15	31	/ 47	? 63	O 79	_ 94	o 111	127	143	159	¯ 175	¿ 191	Ï 207	ß 223	ï 239	ÿ 255

Tabelle 1.6: Zeichensatz für Windows 3.x

0	16	32	0 48	@ 64	P 80	‘ 96	p 112	128	144	160	° 176	À 192	208	à 224	240
1	17	! 33	1 49	A 65	Q 81	a 97	q 113	129	145	¡ 161	± 177	Á 193	Ñ 209	á 225	ñ 241
2	18	” 34	2 50	B 66	R 81	b 98	r 114	130	146	162	^{2} 178	Â 194	Ò 210	â 226	ò 242
3	19	# 35	3 51	C 67	S 82	c 99	s 115	131	147	£ 163	^{3} 179	Ã 195	Ó 211	ã 227	ó 243
4	20	$ 36	4 52	D 68	T 83	d 100	t 116	132	148	164	´ 180	Ä 196	Ô 212	ä 228	ô 244
5	21	% 37	5 53	E 69	U 84	e 101	u 117	133	149	165	μ 181	Å 197	Õ 213	å 229	õ 245
6	22	& 38	6 54	F 70	V 85	f 102	v 118	134	150	| 166	¶ 182	Æ 198	Ö 214	æ 230	ö 246
7	23	’ 39	7 55	G 71	W 86	g 103	w 119	135	151	§ 167	· 183	Ç 199	× 215	ç 231	÷ 247
8	24	(40	8 56	H 72	X 87	h 104	x 120	136	152	¨ 168	184	È 200	Ø 216	è 232	ø 248
9	25	) 41	9 57	I 73	Y 88	i 105	y 121	137	153	© 169	^{1} 185	É 201	Ù 217	é 233	ù 249
10	26	* 42	: 58	J 74	Z 89	j 106	z 122	138	154	ª 170	º 186	Ê 202	Ú 218	ê 234	ú 250
11	27	+ 43	; 59	K 75	[90	k 107	{ 123	139	155	« 171	» 187	Ë 203	Û 219	ë 235	û 251
12	28	, 44	< 60	L 76	\ 91	l 108	| 124	140	156	¬ 172	$\frac{1}{4}$ 188	Ì 204	Ü 220	ì 236	ü 252
13	29	- 45	= 61	M 77	] 92	m 109	} 125	141	157	- 173	$\frac{1}{2}$ 189	Í 205	Ý 221	í 237	ý 253
14	30	. 46	> 62	N 78	^ 93	n 110	~ 126	142	158	174	$\frac{3}{4}$ 190	Î 206	222	î 238	254
15	31	/ 47	? 63	O 79	_ 94	o 111	127	143	159	¯ 175	¿ 191	Ï 207	ß 223	ï 239	ÿ 255

Tabelle 1.7: ISO-8859 Latin 1(ECMA-94 Latin 1) Zeichensatz

Code Set	Name	Coverage	Approved
8859-1	Latin-1	western Europe	15 February 1987
8859-2	Latin-2	eastern Europe	15 February 1987
8859-3	Latin-3	Maltese, Catalan, Galician, Esperanto	15 April 1988
8859-4	Latin-4	Baltic and Nordic region	15 April 1988
8859-5	Cyrillic	Slavic countries	1 December 1988
8859-6	Arabic	Arab countries	15 August 1987
8859-7	Greek	Greece	15 November 1987
8859-8	Hebrew	Israel	1 June 1988
8859-9	Latin-5	8859-1 minus Iceland plus Turkey	15 May 1989

Tabelle 1.8: nationale ISO8859-Varianten

Da die Tastaturen der Bildschirmarbeitsplätze nicht für alle möglichen Zeichen eigene Tasten besitzen (Abbildung 1.7), wird es ermöglicht, durch sogenannte Compose-Tastensequenzen die fehlenden Zeichen zu erzeugen (Tabelle 1.9). Auf

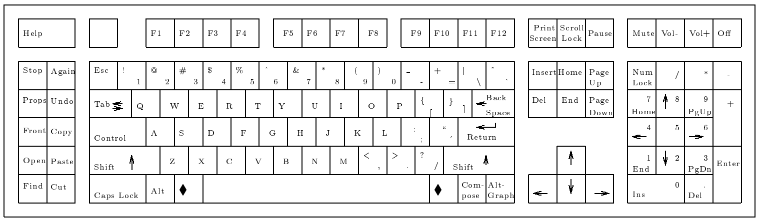

Abbildung 1.7: SunKeyBoard Type 5

PC's existiert daneben noch die Möglichkeit durch gleichzeitiges Drücken der Modifikationstasten `AltGraph`, `Shift`, `AltGraph+Shift` eine Vierfachbelegung einiger Tasten zu realisieren, etwa `AltGraph+Q` für @: Q @ Damit immer noch nicht erreichbare Zeichen können durch ihre Dezimalcodes unter Benutzung der Alt-Taste erzeugt werden: <Alt><2><3><6> (die Alt-Taste muß dabei während der Ziffernengabe niedergedrückt bleiben und die Ziffern müssen im Zifferneingabeblock bei eingeschaltetem Num-Modus betätigt werden).

Zeichen	**Compose-Tastenfolge**
Á	`A '`
á	`a '`
À	`A `` `
à	`a `` `
Â	`A ^`
â	`a ^`
Ã	`A ~`
ã	`a ~`
Ä	`A "`
ä	`a "`
Å	`A *`
å	`a *`
Æ	`A E`
æ	`a e`
Ç	`C ,`
ç	`c ,`
	`d -`
	`D -`
È	`E `` `
è	`e `` `

⋮

Zeichen	**Compose-Tastenfolge**
Í	`I '`
í	`i '`
Î	`I ^`
î	`i ^`
Ï	`I "`
ï	`i "`
Ñ	`N ~`
ñ	`n ~`
Ò	`O `` `
ò	`o `` `
Ó	`O '`
ó	`o '`
Ô	`O ^`
ô	`o ^`
Õ	`O ~`
õ	`o ~`
Ö	`O "`
ö	`o "`
Ø	`O /`
ø	`o /`

Tabelle 1.9: Eingabe von Sonderzeichen / Tastaturen: Compose-Tastensequenzen

1.2.1.2 Zeichencodierung durch Zeichenfolgen

In Textverarbeitungssystemen werden Sonderzeichen häufig als internes Speicherformat in Zeichenfolgen codiert, z.B. in SGML, Tabelle 1.10.

á	`á`	Small a with acute accent
Á	`Á`	Capital A with acute accent
ă	`ă`	Small a with breve accent
Ă	`Ă`	Capital A with breve accent
â	`â`	Small a with circumflex accent
Â	`Â`	Capital A with circumflex accent
æ	`æ`	Small ae diphtong (ligature)
Æ	`Æ`	Capital AE diphtong (ligature)
à	`à`	Small a with grave accent
À	`À`	Capital A with grave accent
ā	`ā`	Small a with macron accent
Ā	`Ā`	Capital A with macron accent
å	`å`	Small a with ring accent
Å	`Å`	Capital A with ring accent
ã	`ã`	Small a with tilde accent
Ã	`Ã`	Capital A with tilde accent
ä	`ä`	Small a with umlaut (diaeresis) accent
Ä	`Ä`	Capital A with umlaut (diaeresis) accent
č	`č`	Small c with caron accent
Č	`Č`	Capital C with caron accent
	⋮	
\	`\`	Backslash (reverse solidus)
√	`√`	Radical sign
∐	`∐`	Coproduct (amalgamation) operator
∏	`∏`	Product operator
∫	`∫`	Integral operator
∮	`&coint;`	Integral operator with central o
∑	`∑`	Summation operator
∇	`∇`	Hamilton operator
⊔	`⊔`	Square union sign
⊓	`⊓`	Square intersection sign
⊑	`⊑`	Square subset of, or equal to, sign
⊒	`⊒`	Square superset of, or equal to, sign
§	`§`	Section sign
†	`&dag;`	Dagger
‡	`&Dag;`	Double dagger
¶	`¶`	Paragraph sign (pilcrow)

Tabelle 1.10: SGML = *standard generalized Markup Language*

HTML = *hypertext markup language* als ein spezieller SGML-Typ sieht etwa wie in Abbildung 1.8 gezeigt aus: