

Computer können Daten nur in **binärer Darstellung** verarbeiten (1 - Strom fließt, Spannung vorhanden / 0 - kein Strom bzw. keine Spannung). Daher werden alle im Computer zu speichernden Informationen in das Binärsystem **codiert** (z.B. 11001101 10001011).

Eine Binärstelle (also 0 oder 1) nennt man dabei ein **Bit**, 8 Bits werden stets zu einem **Byte** zusammengefasst.

Grundlage für die Binärcodierung (und andere Codierungen) sind Zahlen-Positionssysteme. In einem **Positionssystem** hat jede Stelle einer Zahl einen bestimmten Grundwert, der sich als Potenz der Basis b des Zahlensystems ergibt. Dieser Grundwert (Positionswert) wird mit der entsprechenden Ziffer multipliziert. Die möglichen Ziffern der Zahl kommen aus dem Bereich von 0 bis $b-1$.

Allgemeine Darstellung (Basis b , Ziffern von 0 bis $b-1$):

$$z_{n-1} z_{n-2} \dots z_1 z_0 = z_{n-1} \cdot b^{n-1} + z_{n-2} \cdot b^{n-2} + \dots + z_1 \cdot b^1 + z_0 \cdot b^0$$

Beispiel Dezimalsystem (Basis $b=10$, Ziffern von 0 bis 9):

$$26078_{(10)} = 2 \cdot 10^4 + 6 \cdot 10^3 + 0 \cdot 10^2 + 7 \cdot 10^1 + 8 \cdot 10^0$$

Beispiel Binärsystem (Basis $b=2$, Ziffern 0 und 1):

$$101110_{(2)} = 1 \cdot 2^5 + 0 \cdot 2^4 + 1 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0 = 32 + 0 + 8 + 4 + 2 + 0 = 46_{(10)}$$

Beispiel Hexadezimalsystem (Basis $b=16$, Ziffern 0 bis 9, A bis F):

$$3BF7_{(16)} = 3 \cdot 16^3 + 11 \cdot 16^2 + 15 \cdot 16^1 + 7 \cdot 16^0 = 12288 + 2816 + 240 + 7 = 15351_{(10)}$$

Die Hexadezimaldarstellung wird oft als „Abkürzung“ für die Binärdarstellung verwendet, da immer 4 Bits zu einer Hexadezimalziffer zusammengefasst werden können.

$$3BF7_{(16)} = 0011\ 1011\ 1111\ 0111_{(2)} = 15351_{(10)}$$

In Computern werden also Informationen in Zahlen überführt (Zeichencodierungen, Farbwerte, Samplewerte usw.) und diese dann mit Hilfe des Binärformats als Daten gespeichert und verarbeitet.

Der **American Standard Code for Information Interchange** (ASCII) ist eine 7-Bit-Zeichencodierung und bildet die Grundlage für spätere mehrbittige Zeichensätze und -codierungen.

ASCII wurde im Jahr 1967 erstmals als Standard veröffentlicht und im Jahr 1986 zuletzt aktualisiert. Die Zeichenkodierung definiert 128 Zeichen, bestehend aus 33 nicht druckbaren sowie 95 druckbaren Zeichen.

Die Zeichen umfassen das lateinische Alphabet in Groß- und Kleinschreibung, die zehn arabischen Ziffern sowie einige Satz- und Steuerzeichen. Der Zeichenvorrat entspricht weitgehend dem einer Tastatur oder Schreibmaschine für die englische Sprache.

ASCII-Codetabelle										
+	0	1	2	3	4	5	6	7	8	9
30				!	"	#	\$	%	&	'
40	()	*	+	,	-	.	/	0	1
50	2	3	4	5	6	7	8	9	:	;
60	<	=	>	?	@	A	B	C	D	E
70	F	G	H	I	J	K	L	M	N	O
80	P	Q	R	S	T	U	V	W	X	Y
90	Z	[\]	^	_	`	a	b	c
100	d	e	f	g	h	i	j	k	l	m
110	n	o	p	q	r	s	t	u	v	w
120	x	y	z	{		}	~			

In der obigen Tabelle sind die darstellbaren Zeichen der ASCII-Codetabelle abgebildet. Diese Zeichen haben die Werte zwischen 32 und 126. Wert 32 ist das Leerzeichen. Den Wert eines Zeichens, auch Codenummer genannt, ermittelt man, indem man die Werte aus Spalten- und Zeilenüberschrift des Zeichens addiert. Ein großes K hat in dieser Tabelle beispielsweise den Zeilenwert 70 und den Spaltenwert 5. Daraus ergibt sich der Wert 75 für dieses Zeichen.

Erweiterungen des ASCII-Codes basieren auf einer 8-bit-Darstellung und enthalten somit 256 Zeichen (z.B. je nach Erweiterung verschiedene länderspezifische Zeichen), wobei die ersten 128-Zeichen dem ASCII-Code entsprechen.

Alle darstellbaren Zeichen (aller Kulturen und Sprachen) werden im sogenannten **Unicode** (Zeichensatz mit 17 Ebenen zu je 16 bit, bisher 6 Ebenen genutzt, die ersten 128 Zeichen entsprechen dem ASCII-Code) erfasst, der heute von den meisten Computersystemen verwendet wird. Mit den UTF-Codierungen (z.B. UTF-16 bei Windows und OS X) wird der Unicode praktisch implementiert.