

Internetanwendungstechnik (Übung)

Datenrepräsentation

S. Bissell, G. Mühl

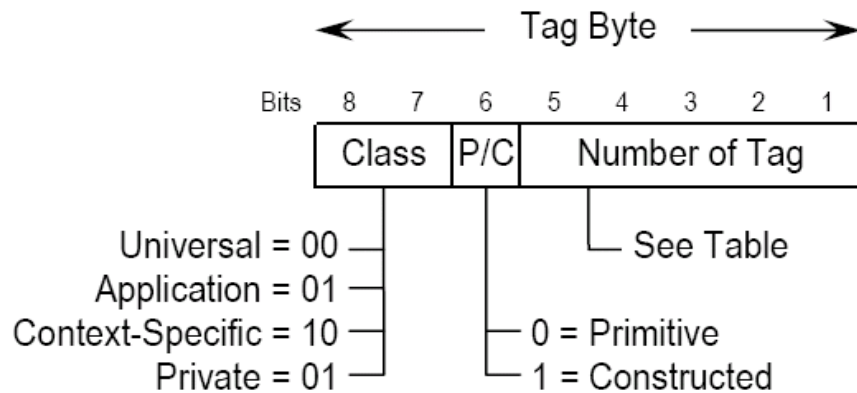
Technische Universität Berlin

Fakultät IV – Elektrotechnik und Informatik

Kommunikations- und Betriebssysteme (KBS)

Einsteinufer 17, Sekr. EN6, 10587 Berlin

ASN.1 Kodierung des Tag-Bytes



0	Reserved for BER
1	BOOLEAN
2	INTEGER
3	BIT STRING
4	OCTET STRING
5	NULL
6	OBJECT IDENTIFIER
7	ObjectDescriptor
8	INSTANCE OF, EXTERNAL
9	REAL

10	ENUMERATED
11	EMBEDDED PDV
12	UTF8String
13	RELATIVE-OID
14-15	Reserved for Future Editions
16	SEQUENCE, SEQUENCE OF
17	SET, SET OF
18	NumericString
19	PrintableString
20	TeletexString, T61String
21	VideotexString
22	IA5String
23	UTCTime
24	GeneralizedTime
25	GraphicString
26	VisibleString, ISO64String
27	GeneralString
28	UniversalString
29	CHARACTER STRING
30	BMPString
31-...	Reserved for Addenda

Tabelle ausschließlich für Universal Typen!

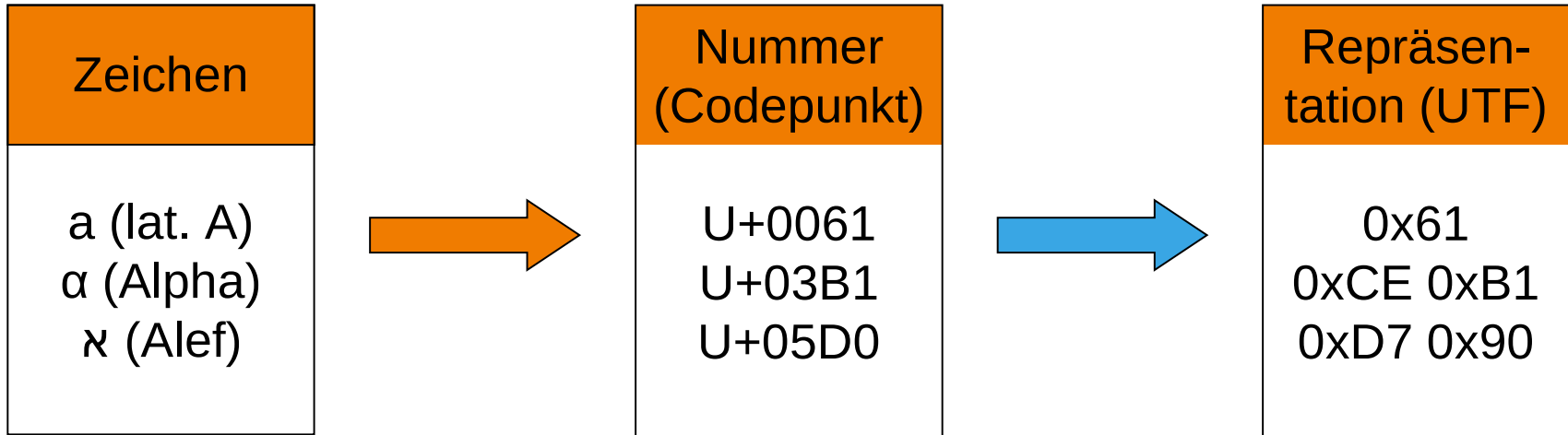
Base64 Kodierungsschema

Text	M	a	n	
ASCII	0x4D (77)	0x61 (97)	0x6E (110)	
Bit Pattern	0 1 0 0 1 1 0 1	0 1 1 0 0 0 0 1	0 1 1 0 1 1 1 0	
Index	0x13 (19)	0x15 (22)	0x05 (5)	0x2E (46)
Base64	T	W	F	u

Base64 Tabelle

Wert	Zeichen	Wert	Zeichen	Wert	Zeichen	Wert	Zeichen
0x00 (00)	A	0x10 (16)	Q	0x20 (32)	g	0x30 (48)	w
0x01 (01)	B	0x11 (17)	R	0x21 (33)	h	0x31 (49)	x
0x02 (02)	C	0x12 (18)	S	0x22 (34)	i	0x32 (50)	y
0x03 (03)	D	0x13 (19)	T	0x23 (35)	j	0x33 (51)	z
0x04 (04)	E	0x14 (20)	U	0x24 (36)	k	0x34 (52)	0
0x05 (05)	F	0x15 (21)	V	0x25 (37)	l	0x35 (53)	1
0x06 (06)	G	0x16 (22)	W	0x26 (38)	m	0x36 (54)	2
0x07 (07)	H	0x17 (23)	X	0x27 (39)	n	0x37 (55)	3
0x08 (08)	I	0x18 (24)	Y	0x28 (40)	o	0x38 (56)	4
0x09 (09)	J	0x19 (25)	Z	0x29 (41)	p	0x39 (57)	5
0x0A (10)	K	0x1A (26)	a	0x2A (42)	q	0x3A (58)	6
0x0B (11)	L	0x1B (27)	b	0x2B (43)	r	0x3B (59)	7
0x0C (12)	M	0x1C (28)	c	0x2C (44)	s	0x3C (60)	8
0x0D (13)	N	0x1D (29)	d	0x2D (45)	t	0x3D (61)	9
0x0E (14)	O	0x1E (30)	e	0x2E (46)	u	0x3E (62)	+
0x0F (15)	P	0x1F (31)	f	0x2F (47)	v	0x3F (63)	/

Unicode



Unicode gibt jedem Zeichen eine Nummer.

Unicode Transformation Format (UTF) bildet die Nummern auf Bytesequenzen ab.

Beispiele für Alphabete im Unicode

- > Latein/ASCII (Basic, 0000 – 007F)
 - > Beispiele: ' , a , k , l , t , v
- > Latein (Supplement, 0080 – 00FF)
 - > Beispiele: ð , ä , ö , ü
- > Griechisch (0370 – 03FF)
 - > Beispiele: Δ , í , α , β , ε , κ , λ , ν , ω
- > International Phonetic Alphabet (IPA Extensions, 0250 – 02AF)
 - > Beispiele: ɔ , ε

Was heißt nun Δεν καταλαβαίνω?

UTF-8 Kodierung

Unicode Bereich	UTF-8 Kodierung
0000 0000 – 0000 007F	0xxxxxxx
0000 0080 – 0000 07FF	110xxxxx 10xxxxxx
0000 0800 – 0000 FFFF	1110xxxx 10xxxxxx 10xxxxxx
0001 0000 – 0010 FFFF	11110xxx 10xxxxxx 10xxxxxx 10xxxxxx

> Startbyte

- > 0xxxxxxx (äquivalent zu ASCII)
- > 1...10x...x (Anzahl der führenden 1 = Anzahl der Bytes gesamt)

> Folgebyte(s)

- > 10xxxxxx (soviel wie notwendig)

Fragen?

