



Unicode

A decorative graphic on the left side of the slide, consisting of two overlapping semi-circles. The top one is a dark teal color, and the bottom one is a lighter, semi-transparent teal color.

Gliederung

- Was ist der Unicode?
- Warum gibt es den Unicode?
- Wie funktioniert er?
- Wo ist mein Schriftzeichen?
- Kritische Stimmen



Was ist der Unicode?

- ein Datensatz von Schriftzeichen wie viele andere auch
- herausgegeben von Unicode Consortium (1991 gegründet)
- auch bekannt als ISO 10646
→ internationaler Industriestandard
- Ziel: alle Schriftzeichen der Welt zu vereinen

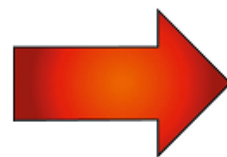
Warum gibt es den Unicode?

- Computer arbeiten intern mit „1“ und „0“ (Binärsystem)

| Dezimal | Binär | Dezimal | Binär |
|---------|-------|---------|-------|
| 0 | 0000 | 8 | 1000 |
| 1 | 0001 | 9 | 1001 |
| 2 | 0010 | 10 | 1010 |
| 3 | 0011 | 11 | 1011 |
| 4 | 0100 | 12 | 1100 |
| 5 | 0101 | 13 | 1101 |
| 6 | 0110 | 14 | 1110 |
| 7 | 0111 | 15 | 1111 |

- um Schriftzeichen verarbeiten zu können, müssen jedem einzelnen konkrete Binärzahlen zugeordnet werden können

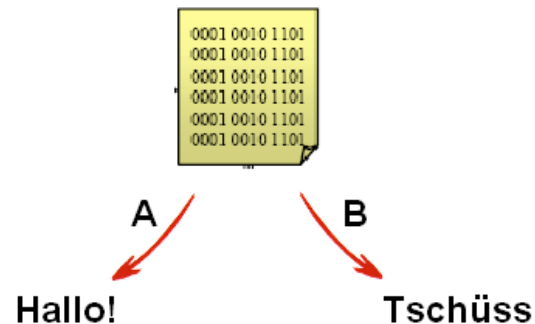
\$ = **11001**



Unicode

Warum gibt es den Unicode?

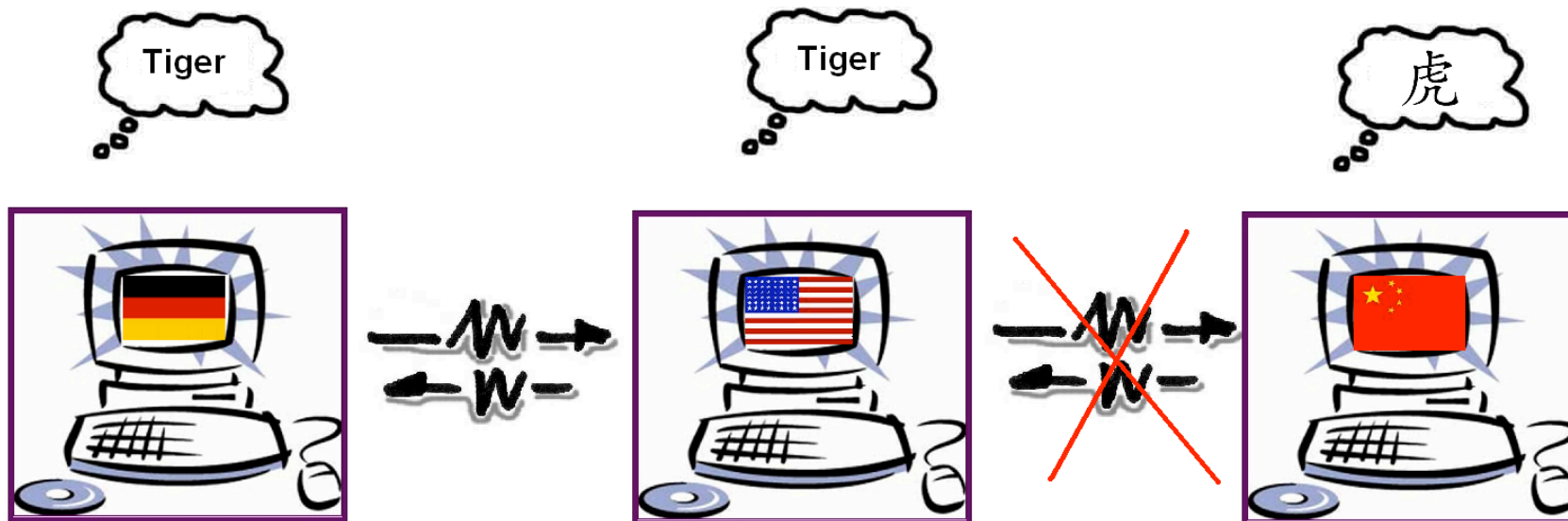
- Im Laufe der Geschichte haben Programmierer aller Welt zunächst ihre eigenen Datensätze entwickelt, die häufig auf 128 (7 Bit) oder 256 (8 Bit) Zeichen beschränkt waren und somit neben den nötigen Steuer- und Sonderzeichen nur eine Zeichensprache enthielten.
- → Inkompatibilität beim Austausch von Texten



Warum gibt es den Unicode?



- 1985: Entstehung Latin-1 (ISO 8859-1) behebt Problem für Sprachen mit Lateinischem Zeichensatz, aber viele Zeichen bleiben inkompatibel





Warum gibt es den Unicode?



- 1991 wurde dann die erste Version von Unicode veröffentlicht
- angelegt zunächst für maximal 65.536 Zeichen, später 1.114.112
- jedes Jahr werden nach strenger Prüfung neue Zeichen unterschiedlichster Sprachen hinzugefügt
- Mittlerweile Version 5.1.0 mit über 100.000 codierten Zeichen, darunter tibetische, phönizische und viele weitere Sprachzeichen
- Anmerkung: auch das Eurozeichen € musste 1998 extra codiert werden



Wie funktioniert der Unicode?

- In einem Datensatz wird ein Schriftzeichen mit einer für den Computer speicherbaren und eindeutigen Binärzahl verknüpft.
- jedes Schriftzeichen bekommt eigene Zeichennummer

\$ = 11001



Wie funktioniert der Unicode?

- Um den Unmengen an Schriftzeichen (momentan ca. 100.000) auf elegante Weise Herr zu werden, werden zur Speicherplatzbenennung **Hexadezimalzahlen** verwendet (statt Binärzahlen).

Exkurs: Hexadezimalsystem

- Zahlensystem zur Basis 16 (2^4)
- Gezählt wird mit den Ziffern:
0,1,2,3,4,5,6,7,8,9,A,B,C,D,E,F
- in der Datenverarbeitung häufig, da es komfortable Verwaltung von Binärzahlen ermöglicht
- es können 4 Bit mit nur einem Zeichen dargestellt werden

| binär | hexad. | dezimal |
|---------|--------|---------|
| 0 0 0 1 | 1 | 1 |
| 0 0 1 0 | 2 | 2 |
| ... | ... | ... |
| 1 1 1 1 | F | 15 |

Exkurs: Hexadezimalsystem

| | | | | | | | | | | | | | | | |
|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
| 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 1A | 1B | 1C | 1D | 1E | 1F |
| 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 2A | 2B | 2C | 2D | 2E | 2F |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| F0 | F1 | F2 | F3 | F4 | F5 | F6 | F7 | F8 | F9 | FA | FB | FC | FD | FE | FF |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| FF0 | FF1 | FF2 | FF3 | FF4 | FF5 | FF6 | FF7 | FF8 | FF9 | FFA | FFB | FFC | FFD | FFE | FFF |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| FFF0 | FFF1 | FFF2 | FFF3 | FFF4 | FFF5 | FFF6 | FFF7 | FFF8 | FFF9 | FFFA | FFFB | FFFC | FFFD | FFFE | FFFF |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |

- Aussprache immer nach Zeichen getrennt:
15 = Eins, Fünf
F7 = eff, sieben

Exkurs: Hexadezimalsystem



| Dezimal | Binär | Hexadezimal |
|---------|------------------|-------------|
| 0 | 0b | 0h |
| 1 | 1b | 1h |
| 2 | 10b | 2h |
| 3 | 11b | 3h |
| 4 | 100b | 4h |
| 5 | 101b | 5h |
| 6 | 110b | 6h |
| 7 | 111b | 7h |
| 8 | 1000b | 8h |
| 9 | 1001b | 9h |
| 10 | 1010b | Ah |
| 11 | 1011b | Bh |
| 12 | 1100b | Ch |
| 128 | 10000000b | 80h |
| 129 | 10000001b | 81h |
| 256 | 100000000b | 100h |
| 1024 | 10000000000b | 400h |
| 4096 | 1000000000000b | 1000h |
| 65535 | 111111111111111b | FFFFh |

- jeder Eintrag innerhalb eines 16 Bit Speichers ($2^{16} = 65.536$ Einträge - inklusive „0“) kann mit einer vierstelligen Hexadezimalzahl eindeutig aufgerufen werden, wobei $0 = 0$ und $65535 = \text{FFFF}$
- Hexadezimalzahlen werden üblicherweise entweder mit dem Präfix `0x` oder dem Suffix `h` versehen, also `0x1B3F` oder `1B3Fh`

Exkurs: Hexadezimalsystem

- um eine Hexadezimalzahl in eine Dezimalzahl umzurechnen, unterteilt man sie in Potenzen
 - Beispiel: $0x1B3F$
 - = $1 * 16^3$ ($1 * 4096 = 4096$)
 - + $B * 16^2$ ($11 * 256 = 2816$)
 - + $3 * 16^1$ ($3 * 16 = 48$)
 - + $F * 16^0$ ($15 * 1 = 15$)
- $0x1B3F = 6975$



Wie funktioniert der Unicode?

- 65536 Speicherplätze für Schriftzeichen
- jeder mit 4-stelliger Hexadezimalzahl eindeutig ansteuerbar
- zur besseren Übersichtlichkeit:
Unterteilung in 256 Blöcke mit je 256 Einträgen, z.B. Block **00** für alle Einträge von **0001** bis **00FF**
- in der Regel ganze Blöcke pro Schriftsystem

Wie funktioniert der Unicode?



| | | | | | | | | | | | | | | | |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 00 | 01 | 02 | 03 | 04 | 05 | 06 | 07 | 08 | 09 | 0A | 0B | 0C | 0D | 0E | 0F |
| 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 1A | 1B | 1C | 1D | 1E | 1F |
| 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 2A | 2B | 2C | 2D | 2E | 2F |
| 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 3A | 3B | 3C | 3D | 3E | 3F |
| 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 4A | 4B | 4C | 4D | 4E | 4F |
| 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 5A | 5B | 5C | 5D | 5E | 5F |
| 60 | 61 | 62 | 63 | 64 | 65 | 66 | 67 | 68 | 69 | 6A | 6B | 6C | 6D | 6E | 6F |
| 70 | 71 | 72 | 73 | 74 | 75 | 76 | 77 | 78 | 79 | 7A | 7B | 7C | 7D | 7E | 7F |
| 80 | 81 | 82 | 83 | 84 | 85 | 86 | 87 | 88 | 89 | 8A | 8B | 8C | 8D | 8E | 8F |
| 90 | 91 | 92 | 93 | 94 | 95 | 96 | 97 | 98 | 99 | 9A | 9B | 9C | 9D | 9E | 9F |
| A0 | A1 | A2 | A3 | A4 | A5 | A6 | A7 | A8 | A9 | AA | AB | AC | AD | AE | AF |
| B0 | B1 | B2 | B3 | B4 | B5 | B6 | B7 | B8 | B9 | BA | BB | BC | BD | BE | BF |
| C0 | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 | CA | CB | CC | CD | CE | CF |
| D0 | D1 | D2 | D3 | D4 | D5 | D6 | D7 | D8 | D9 | DA | DB | DC | DD | DE | DF |
| E0 | E1 | E2 | E3 | E4 | E5 | E6 | E7 | E8 | E9 | EA | EB | EC | ED | EE | EF |
| F0 | F1 | F2 | F3 | F4 | F5 | F6 | F7 | F8 | F9 | FA | FB | FC | FD | FE | FF |

- Lateinische Schriften und Symbole
- Lautschriften
- Andere europäische Schriften
- Nahost- und Südwestasiatische Schriften
- Afrikanische Schriften
- Südasiatische Schriften
- Südostasiatische Schriften
- Ostasiatische Schriften
- CJK-Ideogramme
- Kanadische Silben
- Symbole
- Diakritika
- UTF-16-Surrogates und privater Nutzungsbereich
- Verschiedene Zeichen
- Nicht belegte Codebereiche



Wie funktioniert der Unicode?

- 65.536 Zeichen reichen nicht
→ Speicherbereich wird versiebzehnfacht
- Codierung von 17 Planes á 65.536
- bisheriger Unicode bleibt als Plane 0 oder Basic Multilingual Plane (aktuell gebräuchliche Schriftsysteme, Satzzeichen etc.) erhalten (vorherige Folie)
- jetzt: max. 1.114.112 Zeichen

Wie funktioniert der Unicode?



- Unicode speichert Schriftzeichen, aber keine Glyphen
- Glyphe ist das Schriftbild/die Grafik, welche wir sehen/schreiben
- viele unterschiedliche Glyphen für identisches Schriftzeichen
- Glyphen sind in Schriftart-Datensätzen gespeichert

Aa
Œa
Aa
Aa
Aa

Das Prädikat „Unicode-konform“ einer Schriftart bedeutet, dass sie die Schriftzeicheneinträge des Unicode verwendet, aber nicht, dass sie alle Unicode-Schriftzeichen als Glyphe grafisch darstellen kann.



Wie funktioniert der Unicode?

- Umgang mit Unicode gestaltet sich zunächst einfach, da in modernen Betriebssystemen/Editoren integriert
- Wir arbeiten mit ihm, ohne es zu merken.
- Unter Windows häufig Unicode-Direkteingabe möglich mit Alt + *****, wobei ***** Unicode-Eintrag als Dezimalzahl (umrechnen) meint
- Verständnishilfe: Windows-Zeichentabelle
(Start→Programme→Zubehör→Systemprogramme)



Wo ist mein Schriftzeichen?

- Suche nach einem Zeichen
 - a) Windows-Zeichentabelle
 - b) Unicode-Datenbank
- Suche in Unicode-Datenbank auf unterschiedliche Weise
 - a) über Unicode Character Name Index
 - b) über Unicode NamesList
 - c) über Unicode Character Code Charts By Script

Wo ist mein Schriftzeichen?



Index

| | | | | | | | | |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| A | B | C | D | E | F | G | H | I |
| J | K | L | M | N | O | P | Q | R |
| S | T | U | V | W | X | Y | Z | |

Related Links

[About the Character Names Index](#)
[Code Charts](#)
[About the Unicode Standard](#)
[Versions of the Unicode Standard](#)
[Latest Version of the Unicode Standard](#)
[The Unicode Standard, Version 5.0](#)

Unicode Character Name Index

Click on an index letter on the left or on a character range below.

[\[A - D\]](#) [\[E - N\]](#) [\[O - Z\]](#)

[A]

| | |
|--|----------------------|
| A WITH ACUTE, LATIN CAPITAL LETTER | 00C1 |
| A WITH ACUTE, LATIN SMALL LETTER | 00E1 |
| A WITH BREVE, LATIN SMALL LETTER | 0103 |
| A WITH CARON, LATIN SMALL LETTER | 01CE |
| A WITH CIRCUMFLEX, LATIN CAPITAL LETTER | 00C2 |
| A WITH CIRCUMFLEX, LATIN SMALL LETTER | 00E2 |
| A WITH DIAERESIS, LATIN CAPITAL LETTER | 00C4 |
| A WITH DIAERESIS, LATIN SMALL LETTER | 00E4 |
| A WITH DOT ABOVE, LATIN SMALL LETTER | 0227 |
| A WITH DOT BELOW, LATIN SMALL LETTER | 1EA1 |
| A WITH DOUBLE GRAVE, LATIN SMALL LETTER | 0201 |
| A WITH GRAVE, LATIN CAPITAL LETTER | 00C0 |
| A WITH GRAVE, LATIN SMALL LETTER | 00E0 |
| A WITH HOOK ABOVE, LATIN SMALL LETTER | 1EA3 |
| A WITH INVERTED BREVE, LATIN SMALL LETTER | 0203 |
| A WITH MACRON, LATIN SMALL LETTER | 0101 |
| A WITH OGONEK, LATIN SMALL LETTER | 0105 |
| A WITH RIGHT HALF RING, LATIN SMALL LETTER | 1E9A |
| A WITH RING ABOVE, LATIN CAPITAL LETTER | 00C5 |
| A WITH RING ABOVE, LATIN SMALL LETTER | 00E5 |
| A WITH RING BELOW, LATIN SMALL LETTER | 1E01 |
| A WITH STROKE, LATIN CAPITAL LETTER | 023A |

Wo ist mein Schriftzeichen?



```
@@@ The Unicode Standard 5.1
@@@+ U51M080326.lst
Final Unicode 5.1 names list.
This file is semi-automatically derived from UnicodeData.txt and
a set of manually created annotations using a script to select
or suppress information from the data file. The rules used
for this process are aimed at readability for the human reader,
at the expense of some details; therefore, this file should not
be parsed for machine-readable information.

@+ Copyright (c) 1991-2008 Unicode, Inc.
For terms of use, see http://www.unicode.org/terms\_of\_use.html
@@ 0000 CO Controls and Basic Latin (Basic Latin) 007F
@@+
@ CO controls
@+ Alias names are those for ISO/IEC 6429:1992. Commonly used alternative aliases are also shown.
0000 <control>
= NULL
0001 <control>
= START OF HEADING
0002 <control>
= START OF TEXT
0003 <control>
= END OF TEXT
0004 <control>
= END OF TRANSMISSION
0005 <control>
= ENQUIRY
0006 <control>
= ACKNOWLEDGE
0007 <control>
= BELL
0008 <control>
= BACKSPACE
0009 <control>
= CHARACTER TABULATION
= horizontal tabulation (HT), tab
000A <control>
= LINE FEED (LF)
= new line (NL), end of line (EOL)
000B <control>
= LINE TABULATION
= vertical tabulation (VT)
```

Wo ist mein Schriftzeichen?



The Unicode Character Code Charts By Script

updated for Unicode 5.1

[SYMBOLS AND PUNCTUATION](#) | [NAME INDEX](#) | [HELP AND LINKS](#)

| European Alphabets | African Scripts | Indic Scripts | East Asian Scripts |
|-------------------------------------|-------------------------------|-------------------------|------------------------------------|
| (see also Comb. Marks) | Ethiopic | Bengali | Han Ideographs |
| Armenian | Ethiopic | Devanagari | Unified CJK Ideographs (5MB) |
| Armenian | Ethiopic Supplement | Gujarati | CJK Ideographs Ext. A (2MB) |
| <i>Armenian Ligatures</i> | Ethiopic Extended | Gurmukhi | CJK Ideographs Ext. B (13MB) |
| Coptic | Other African scripts | Kannada | |
| Coptic | N'Ko | Lepcha | Compatibility Ideographs (.5MB) |
| <i>Coptic in Greek block</i> | Osmanya | Limbu | ... Supplement (.5MB) |
| Cyrillic | Tifinagh | Malayalam | Kanbun |
| Cyrillic | Vai | Oi Chiki | (see also Unihan Database) |
| Cyrillic Supplement | Middle Eastern Scripts | Oriya | Radicals and Strokes |
| Cyrillic Extended A | Arabic | Saurashtra | CJK Radicals |
| Cyrillic Extended B | Arabic | Sinhala | KangXi Radicals |
| Georgian | Arabic Supplement | Syloti Nagri | CJK Strokes |
| Georgian | Arabic Present. Forms A | Tamil | Ideographic Description |
| Georgian Supplement | Arabic Present. Forms B | Telugu | Chinese-specific |
| Greek | Hebrew | South East Asian | Bopomofo |
| Greek | Hebrew | Balinese | Bopomofo Extended |
| Greek Extended | <i>Hebrew Present. Forms</i> | Buginese | Japanese-specific |
| (see also Ancient Greek) | Syriac | Cham | Hiragana |
| Latin | Syriac | Kayah Li | Katakana |
| Basic Latin | Thaana | Khmer | Katakana Phonetic Ext. |
| Latin-1 | Thaana | Khmer Symbols | <i>Halfwidth Katakana</i> |
| Latin Extended A | American scripts | Lao | Korean-specific |
| Latin Extended B | Canadian Syllabics | Myanmar | Hangul Syllables (4MB) |
| Latin Extended C | Cherokee | New Tai Lue | Hangul Jamo |
| Latin Extended D | Deseret | Rejang | Hangul Compatibility Jamo |
| Latin Extended Additional | Philippine Scripts | Sundanese | Halfwidth Jamo |
| Latin Ligatures | Buhid | Tai Le | Yi |
| Fullwidth Latin Letters | Hanunoo | Thai | Yi (.6MB) |
| Small Forms | Tagalog | Other Scripts | Yi Radicals |
| (see also Phonetic Symbols) | Tagbanwa | Shavian | |

Wo ist mein Schriftzeichen?



0000

C0 Controls and Basic Latin

007F

| | 000 | 001 | 002 | 003 | 004 | 005 | 006 | 007 |
|---|-------------|-------------|------------|-----------|-----------|-----------|-----------|-----------|
| 0 | NUL 0000 | DLE 0010 | SP 0020 | 0 0030 | @ 0040 | P 0050 | ` 0060 | p 0070 |
| 1 | SOH 0001 | DC1 0011 | ! 0021 | 1 0031 | A 0041 | Q 0051 | a 0061 | q 0071 |
| 2 | STX 0002 | DC2 0012 | " 0022 | 2 0032 | B 0042 | R 0052 | b 0062 | r 0072 |
| 3 | ETX 0003 | DC3 0013 | # 0023 | 3 0033 | C 0043 | S 0053 | c 0063 | s 0073 |
| 4 | EOT 0004 | DC4 0014 | \$ 0024 | 4 0034 | D 0044 | T 0054 | d 0064 | t 0074 |
| 5 | ENQ 0005 | NAK 0015 | % 0025 | 5 0035 | E 0045 | U 0055 | e 0065 | u 0075 |
| 6 | ACK 0006 | SYN 0016 | & 0026 | 6 0036 | F 0046 | V 0056 | f 0066 | v 0076 |
| 7 | BEL 0007 | ETB 0017 | ' 0027 | 7 0037 | G 0047 | W 0057 | g 0067 | w 0077 |

Wo ist mein Schriftzeichen?



1C50

Ol Chiki

1C7F

| | 1C5 | 1C6 | 1C7 |
|---|-----------|-----------|-----------|
| 0 | 0 1C50 | b 1C60 | ⓪ 1C70 |
| 1 | 𑄠 1C51 | 𑄡 1C61 | 𑄢 1C71 |
| 2 | 𑄣 1C52 | 𑄤 1C62 | 𑄥 1C72 |
| 3 | 𑄧 1C53 | 𑄨 1C63 | 𑄩 1C73 |
| 4 | 𑄫 1C54 | 𑄬 1C64 | 𑄭 1C74 |
| 5 | 𑄯 1C55 | 𑄰 1C65 | 𑄱 1C75 |
| 6 | 𑄴 1C56 | 𑄵 1C66 | 𑄶 1C76 |
| 7 | 𑄹 1C57 | 𑄺 1C67 | 𑄻 1C77 |

Digits

| | | |
|------|---|----------------------|
| 1C50 | 0 | OL CHIKI DIGIT ZERO |
| 1C51 | 𑄠 | OL CHIKI DIGIT ONE |
| 1C52 | 𑄣 | OL CHIKI DIGIT TWO |
| 1C53 | 𑄧 | OL CHIKI DIGIT THREE |
| 1C54 | 𑄫 | OL CHIKI DIGIT FOUR |
| 1C55 | 𑄯 | OL CHIKI DIGIT FIVE |
| 1C56 | 𑄴 | OL CHIKI DIGIT SIX |
| 1C57 | 𑄹 | OL CHIKI DIGIT SEVEN |
| 1C58 | 𑄴 | OL CHIKI DIGIT EIGHT |
| 1C59 | 𑄫 | OL CHIKI DIGIT NINE |

Letters

| | | |
|------|---|---------------------|
| 1C5A | 𑄲 | OL CHIKI LETTER LA |
| 1C5B | 𑄱 | OL CHIKI LETTER AT |
| 1C5C | 𑄫 | OL CHIKI LETTER AG |
| 1C5D | 𑄣 | OL CHIKI LETTER ANG |
| 1C5E | 𑄡 | OL CHIKI LETTER AL |
| 1C5F | 𑄢 | OL CHIKI LETTER LAA |
| 1C60 | b | OL CHIKI LETTER AAK |
| 1C61 | 𑄡 | OL CHIKI LETTER AAJ |
| 1C62 | 𑄤 | OL CHIKI LETTER AAM |
| 1C63 | 𑄨 | OL CHIKI LETTER AAW |
| 1C64 | 𑄬 | OL CHIKI LETTER LI |
| 1C65 | 𑄰 | OL CHIKI LETTER IS |
| 1C66 | 𑄴 | OL CHIKI LETTER IH |
| 1C67 | 𑄶 | OL CHIKI LETTER INY |
| 1C68 | 𑄺 | OL CHIKI LETTER IR |
| 1C69 | 𑄻 | OL CHIKI LETTER LU |
| 1C6A | 𑄻 | OL CHIKI LETTER UC |
| 1C6B | 𑄻 | OL CHIKI LETTER UD |
| 1C6C | 𑄻 | OL CHIKI LETTER UNN |



Wo ist mein Schriftzeichen?



- manche Zeichen werden auch als Kombination mehrerer anderer Zeichen dargestellt
→ erschwerte Suche

| Character | Code Points | Linguistic Usage |
|-----------|----------------|-----------------------------|
| ch | 0063 0068 | Slovak, traditional Spanish |
| ᵗ | 0074 02B0 | |
| ꝛ | 0078 0323 | Native American languages |
| Ą | 019B 0313 | |
| ą | 00E1 0328 | Lithuanian |
| ı | 0069 0307 0301 | |
| ᵛ | 30C8 309A | Ainu in kana transcription |

Wo ist mein Schriftzeichen?

- Beispiel: indische Halbzeichen
- Darstellung erfolgt durch Kombination von normalem Zeichen mit dem indischen Halant (diakritisches Zeichen)

| Representative shapes in code chart | | Display appearance |
|-------------------------------------|---|--------------------|
| क | ् | क |

Weitergehende Darstellungen setzen gezieltes Wissen der einzelnen Zeichensprachen voraus.

Kritische Stimmen

- Zwar Industriestandard, aber nicht weltweit beliebt
- Besonders: Han-Vereinheitlichung
→ chinesische, japanische und koreanische
Schriftzeichen als eine Zeichenfamilie
- zukünftige Codierung Ägyptischer Hieroglyphen auf
Forschungen eines einzelnen Forschers basierend
- teilweise sehr lange Diskussionen über
Sprachzugehörigkeiten
→ Problem fernab der Informatik

| | | | |
|-----|-----|-----|-----|
| CHS | CHT | JAP | KOR |
| 次 | 次 | 次 | 次 |



Ende



конец

末端

!!!

نهاية

Exkurs: Binärsystem

- Zahlensystem, basierend auf 2 Ziffern: „0“ und „1“
- gezählt wird: 0, 1, 10, 11, 100, 101, 110, 111, 1000, 1001...
- in Informatik: einstellige Zahl = 1 Bit, zweistellige Zahl = 2 Bit ...
- 1 Bit hat $2 = 2^1$ Einträge,
2 Bit haben $4 = 2^2$,
3 Bit haben $8 = 2^3$
usw.

