WIKIPEDIA

Optical Character Recognition (Unicode block)

Optical Character Recognition is a Unicode block containing signal characters for OCR and MICR standards.

Contents

Optical Character Recognition	
Range	U+2440..U+245F (32 code points)
Plane	BMP
Scripts	Common
Symbol sets	OCR controls
Assigned	11 code points
Unused	21 reserved code points
Source standards	ISO 2033
Unicode version history	
1.0.0 (1991) 11 (+11)	
Chart	
Code chart (https://www.unicode.org/charts/PDF/U2440.pdf)	
Note: [1][2]	

Block

Optical Character Recognition[1][2] Official Unicode Consortium code chart (https://www.unicode.org/charts/PDF/U2440.pdf) (PDF)																
	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
U+244x	⑀	⑁	⑂	⑃	⑄	⑅	⑆	⑇	⑈	⑉	⑊					
U+245x																

Notes

1.^ As of Unicode version 14.0
2.^ Grey areas indicate non-assigned code points

Subheadings

The Optical Character Recognition block has three informal subheadings (groupings) within its character collection: OCR-A, MICR, and OCR.[3]

OCR-A

The OCR-A subheading contains six characters taken from the OCR-A font described in the ISO 1073-1:1976 standard: U+2440 ⑀ OCR HOOK, U+2441 ⑁ OCR CHAIR, U+2442 ⑂ OCR FORK, U+2443 ⑃ OCR INVERTED FORK, U+2444 ⑄ OCR BELT BUCKLE, and U+2445 ⑅ OCR BOW TIE. The OCR bow tie is given the informative alias "unique asterisk".

MICR

The MICR subheading contains four punctuation characters for bank cheque identifiers, taken from the magnetic ink character recognition E-13B font (codified in the ISO 1004:1995 standard): U+2446 ⑆ OCR BRANCH BANK IDENTIFICATION, U+2447 ⑇ OCR AMOUNT OF CHECK, U+2448 ⑈ OCR DASH, and U+2449 ⑉ OCR CUSTOMER ACCOUNT NUMBER.

The latter two characters are misnamed: their names were inadvertently switched when they were named in the 1993 (first) edition of ISO/IEC 10646,[4] a mistake which had been present since Unicode 1.0.0.[5] Although their formal names remain unchanged due to the Unicode stability policy, they both have corrected normative aliases: U+2448 ⑈ is MICR ON US SYMBOL, and U+2449 ⑉ is MICR DASH SYMBOL[6] (the standard notes that "the Unicode character names include several misnomers").

These symbols had previously been encoded by the ISO-IR-98 encoding defined by ISO 2033:1983, in which they were simply named SYMBOL ONE through SYMBOL FOUR.[7] All four characters have informative aliases in the Unicode charts: "transit", "amount", "on us", and "dash" respectively.

OCR

The OCR subheading consists of a single character: U+244A ⑊ OCR DOUBLE BACKSLASH.

History

The following Unicode-related documents record the purpose and process of defining specific characters in the Optical Character Recognition block:

Version	Final code points[a]	Count	L2 ID	WG2 ID	Document
1.0.0	U+2440..244A	11			(to be determined)
			L2/10-416R (https://www.unicode.org/L2/L2010/10416.htm)		Moore, Lisa (2010-11-09), "Consensus 125-C39", *UTC #125 / L2 #222 Minutes*, "Create two formal aliases, U+2448 MICR ON US SYMBOL and U+2449 MICR DASH SYMBOL for Unicode 6.1."
				N4103 (https://www.unicode.org/wg2/docs/n4103.pdf)	"T.3. Optical Character Recognition", *Unconfirmed minutes of WG 2 meeting 58*, 2012-01-03

a. Proposed code points and characters names may differ from final code points and names

References

1. "Unicode character database" (https://www.unicode.org). *The Unicode Standard*. Retrieved 2016-07-09.
2. "Enumerated Versions of The Unicode Standard" (https://www.unicode.org/versions/enumeratedversions.html). *The Unicode Standard*. Retrieved 2016-07-09.

3. "Unicode Code Charts: Optical Character Recognition" (https://www.unicode.org/charts/PDF/U2440.pdf) (PDF). *The Unicode Standard, Version 6.3*. Retrieved 27 February 2014.

4. ISO/IEC JTC 1/SC 2/WG 2 (2012-01-03). "T.3. Optical Character Recognition". *Unconfirmed minutes of WG 2 meeting 58* (https://www.unicode.org/wg2/docs/n4103.pdf) (PDF). p. 29. SC2 N4188 / WG2 N4103. "These Magnetic Ink Character Recognition (MICR) symbols are used by banks on checks. The names of these characters were inadvertently mixed up in the 1993 edition of ISO/IEC 10646."

5. "3.8: Block-by-Block Charts" (https://www.unicode.org/versions/Unicode1.0.0/CodeCharts2.pdf) (PDF). *The Unicode Standard*. version 1.0. Unicode Consortium.

6. Freytag, Asmus; McGowan, Rick; Whistler, Ken (2017-04-10). *Known Anomalies in Unicode Character Names* (https://www.unicode.org/notes/tn27/tn27-4.html) (4 ed.). Unicode Consortium. Unicode Technical Note #27.

7. ISO/TC97/SC2 (1985-08-01). *ISO-IR-98: E13B Graphic Character Set* (https://www.itscj-ipsj.jp/ir/098.pdf) (PDF). ITSCJ/IPSJ.

Retrieved from "https://en.wikipedia.org/w/index.php?title=Optical_Character_Recognition_(Unicode_block)&oldid=1076374760"

This page was last edited on 10 March 2022, at 20:15 (UTC).