

BBN-ANG-183 Typography

Digital fonts

Zoltán G. Kiss & Péter Szigetvári

Dept of English Linguistics, Eötvös Loránd University

topic outline

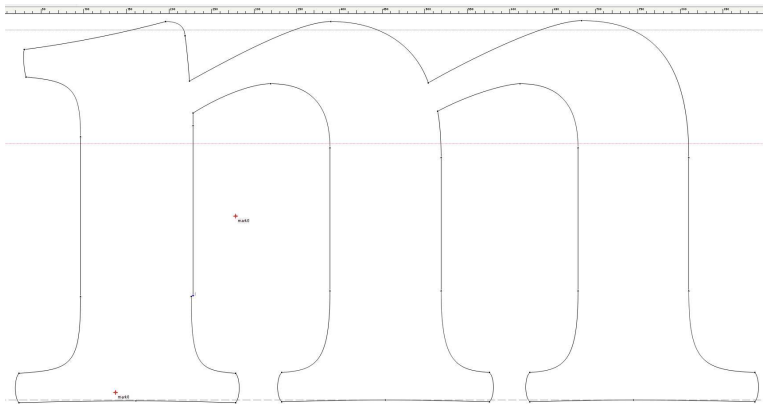
- ▶ digital font formats, digital font technology (for print and screen)
- ▶ what's inside a digital font?
- ▶ how does the computer handle fonts?
- ▶ codepages, UNICODE
- ▶ smart font options, web font formats

what's in a digital font?

fonts are computer programs that usually contain:

- ▶ **painting/drawing instructions** (e.g., line/curve outlines defined in PostScript; hinting)
- ▶ **metrical data** (x-height, set-width, etc.) in ems
- ▶ **kerning** information in ems
- ▶ **ligature information** (“lig pairs”)
- ▶ **character set** and **encoding** (mapping between character/glyph to a code number)

outlines of the glyph “m” defined by the PostScript language



what's in a digital font?

fonts are computer programs that usually contain:

- ▶ **painting/drawing instructions** (e.g., line/curve outlines defined in PostScript; hinting)
- ▶ **metrical data** (x-height, set-width, etc.) in ems
- ▶ **kerning** information in ems
- ▶ **ligature information** (“lig pairs”)
- ▶ **character set** and **encoding** (mapping between character/glyph to a code number)

kerning

V + A → VA

what's in a digital font?

fonts are computer programs that usually contain:

- ▶ **painting/drawing instructions** (e.g., line/curve outlines defined in PostScript; hinting)
- ▶ **metrical data** (x-height, set-width, etc.) in ems
- ▶ **kerning** information in ems
- ▶ **ligature information** (“lig pairs”)
- ▶ **character set** and **encoding** (mapping between character/glyph to a code number)

ligature

f+i → fi

what's in a digital font?

fonts are computer programs that usually contain:

- ▶ **painting/drawing instructions** (e.g., line/curve outlines defined in PostScript; hinting)
- ▶ **metrical data** (x-height, set-width, etc.) in ems
- ▶ **kerning** information in ems
- ▶ **ligature information** (“lig pairs”)
- ▶ **character set** and **encoding** (mapping between character/glyph to a code number)

two meanings of font “format”

format, meaning 1

the **platform** for which the font was designed (Windows PC vs. Mac)

format, meaning 2

the **way** that the typographic information itself is **described** and how that information is **organized** (bitmap vs. outline; Type 1 vs. TrueType, etc.)

basic font formats (w.r.t. glyph description)

- ▶ **bitmap(ped)** fonts (a.k.a. screen/printer fonts)
- ▶ **outline** fonts (a.k.a. postscript, vector, scalable fonts)

outline fonts

characteristics

- ▶ collections of PostScript-based curve outlines of glyphs
- ▶ every glyph is drawn with the help of straight line and curve segments

outline fonts

advantages

- ▶ outlines can be mathematically scaled to any size without distorting the shapes/proportions of the glyphs
- ▶ excellent/clean pdf output
- ▶ device & resolution independence
- ▶ *de facto* standard of publishing & printing industry
- ▶ most digital fonts available are outline

outline fonts

disadvantages

- ▶ at low resolution shapes are often distorted (“hinting”)
- ▶ font shapes depend on the rendering engine technology, on-screen display is often different based on the operating system (Windows vs. Mac, iOS vs. Android)

outline font formats

- ▶ (METAFONT)
- ▶ **PostScript fonts: Type 1** (Type 2, Type 3)
- ▶ **TrueType** fonts
- ▶ (MultipleMaster (MM) fonts)
- ▶ **OpenType** fonts: PS and TT “flavours”

PostScript (Type 1) fonts

features

- ▶ developed by Adobe Systems, San Jose CA
- ▶ use the PostScript language, glyphs: defined with lines & Bézier curves
- ▶ two files for each font:
 - ▶ **outline** data (*.pfb/*.pfa)
 - ▶ **metrics** data (*.pfm/*.afm), they must be in the same directory
- ▶ standard for graphic design, desktop publishing industry
- ▶ 8-bit font format (max. $2^8 = 256$ characters)

TrueType (TT) fonts

features

- ▶ developed by Apple and Microsoft as a reaction against Adobe's high-cost licensing of Type 1 fonts
- ▶ use modified Bézier curves
- ▶ better algorithm for hinting/non-printed displays
- ▶ TT fonts come in a single (binary) file (*.ttf): both outlines and metrics
- ▶ also 8-bit font format (max. $2^8 = 256$ characters)

OpenType (OT) fonts

features

- ▶ jointly developed by Adobe and Microsoft since 1996
- ▶ has taken over as the standard font format, constantly developed
- ▶ two subformats of *glyph outlines*:
 - ▶ **PS-flavoured** OT (*.otf) (for print, also called Compact Font Format/CFF)
 - ▶ **TT-flavoured** OT (*.ttf) (for screens)
- ▶ come in a single file containing both outline and metrics data (in binary form), and many more data, too

why OpenType?

advantages

- ▶ **platform independent & cross-codepage compatible**
- ▶ supports **Unicode**
- ▶ includes **smartfont** options within same file
⇒ more sophisticated typography
- ▶ better **protection for font data**

digital fonts: important terms

- ▶ character
- ▶ glyph
- ▶ character set/repertoire
- ▶ character encoding
- ▶ font encoding

character – glyph

glyph

the actual image/drawing of a symbol used in a writing system or in a notational system (e.g., music, maths, phonetics, etc.)

character

- ▶ a simple **description of class of glyphs**, an abstraction
- ▶ the description is usually illustrated with an **illustrative glyph**

character – glyph

character: ‘the capital Latin letter double-you, W’

glyphs:



character set, code, & font encoding

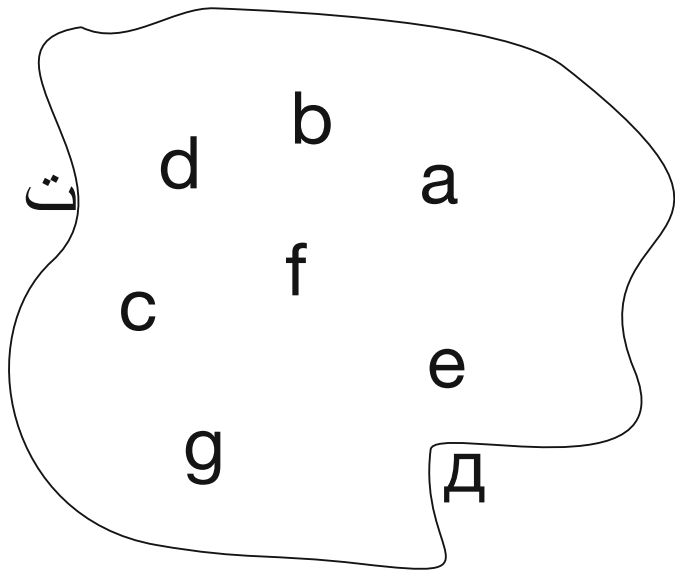
character set/repertoire

- ▶ a collection of distinct characters
- ▶ e.g.: the characters that can be used to represent the letters of the English alphabet

unorganized characters

ا b a
d
c f
g e
Д

characters collected into a character set ('Latin')



character encoding

character encoding

- ▶ computers work with numbers, so characters need to be identified by a number
- ▶ a unique **numerical code** is assigned to each character in the set
- ▶ an “organized” numbered set of characters that the computer can consistently identify
- ▶ called **character codepage/table**

character encoding/a codepage

97 → a

98 → b

99 → c

100 → d

101 → e

102 → f

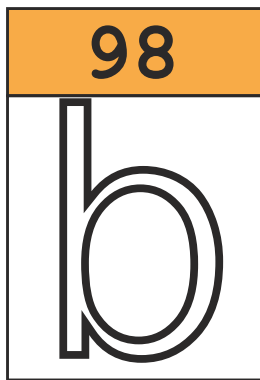
103 → g

character repertoire, code, & font encoding

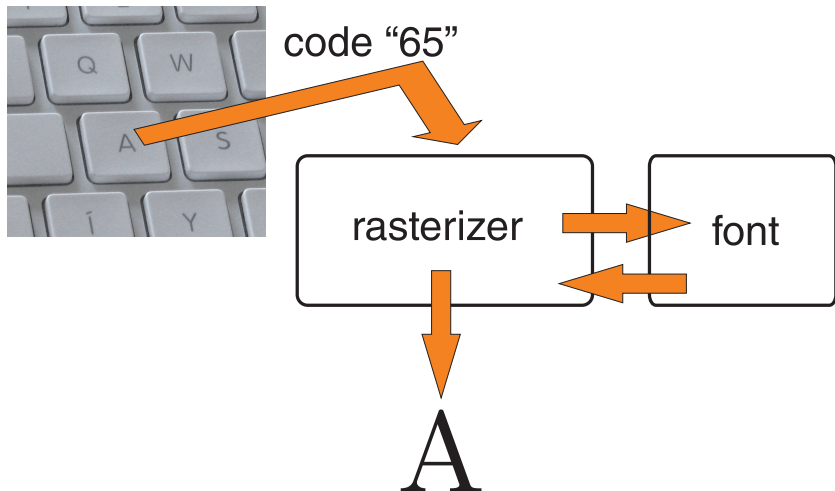
font encoding

- ▶ a character encoding **within a digital font**
- ▶ a mapping between character **code** and the **glyph**
- ▶ drawing instructions of glyphs organized into a numbered set

font encoding (= encoding inside the font file)



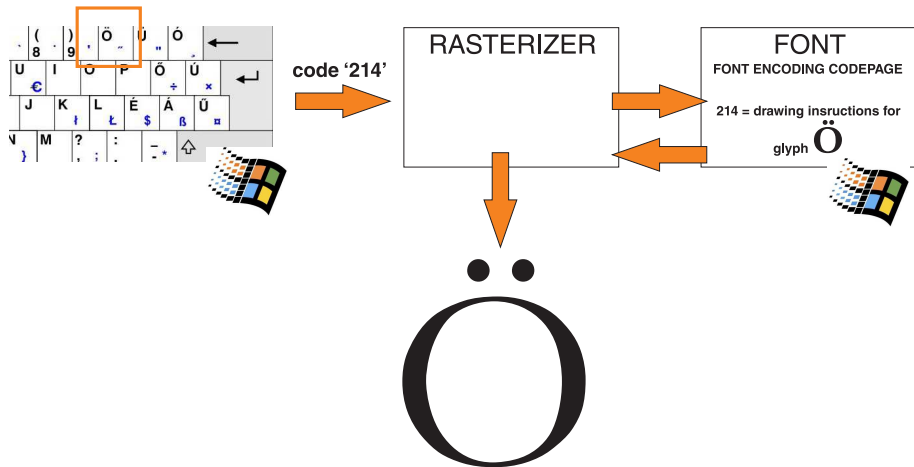
from code 65 to glyph A



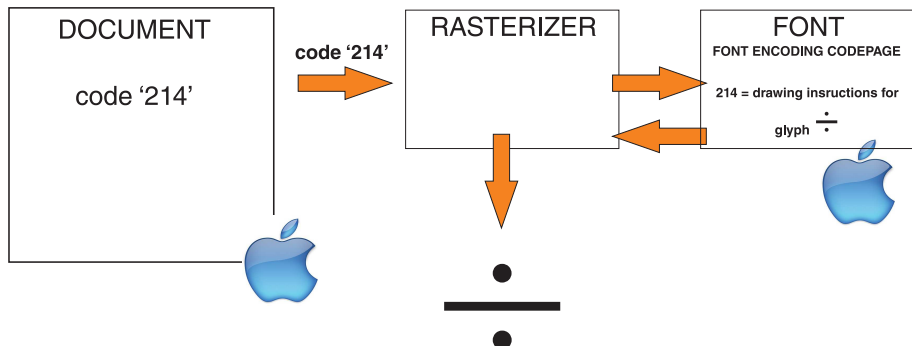
problems with older digital fonts

- ▶ characters are displayed unexpectedly, or not displayed at all: you get an empty rectangle
- ▶ reason: font-encoding differences between operating systems
- ▶ limited number of characters in a font ($2^8 = 256$)

typing Ö on Hungarian Windows (CP1250/Latin 2 enc.)



code 214 on a Mac (CP1275/Mac Roman encoding)



CP1250/Latin 2 Windows encoding

Codepage 1250 - Latin 2 Windows

	-0	-1	-2	-3	-4	-5	-6	-7	-8	-9	-A	-B	-C	-D	-E	-F	
0-		0001	0002	0003	0004	0005	0006	0007	0008	0009	000A	000B	000C	000D	000E	000F	
1-	0010	0011	0012	0013	0014	0015	0016	0017	0018	0019	001A	001B	001C	001D	001E	001F	
2-	0020	!	"	#	\$	%	&	'	()	*	+	,	-	.	/	
3-	0030	0	1	2	3	4	5	6	7	8	9	:	<	=	>	?	
4-	0040	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5-	0050	P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
6-	0060	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
7-	0070	p	q	r	s	t	u	v	w	x	y	z	{		}	~	?
8-	20AC	€	,	”	…	†	‡		%	Š	<	Š	Ť	Ž	Ž		
9-	0090	‘	’	“	”	•	—		™	š	>	š	ı	ž	ž		
A-	00A0	˘	Ł	ł	Ą	ą	§	”	©	§	«	¬	®	Ž			
B-	00B0	±	ı	ı	ı	ı	ı	ı	ı	ı	ı	ı	ı	ı	ı	ı	
C-	0154	Ř	Á	Â	Ã	Ä	Å	Ĺ	Ć	Ç	Č	É	Ê	Ë	Í	Î	
D-	0110	Ď	Ň	Ń	Ó	Ô	Õ	Ö	×	Ř	Š	Ť	Ú	Û	Ý	Ţ	
E-	0155	ř	á	â	ã	ä	å	ł	ç	ç	č	é	ê	ë	í	î	
F-	0111	ď	ň	ń	ó	ô	õ	ö	÷	ř	š	ť	ú	û	ý	ţ	

CP1275/Mac Roman encoding

Codepage 1275 - Apple Latin 1

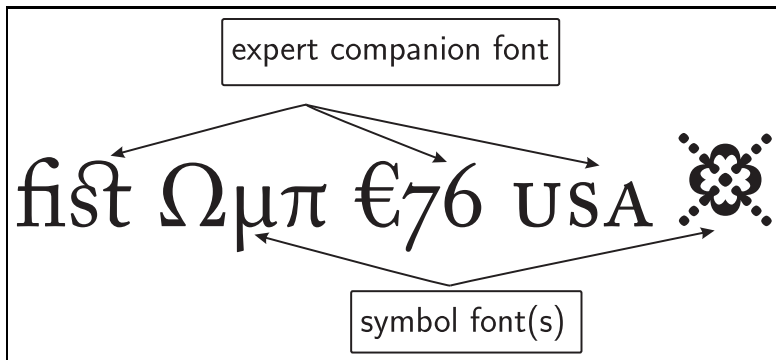
	-0	-1	-2	-3	-4	-5	-6	-7	-8	-9	-A	-B	-C	-D	-E	-F		
0-		0001	0002	0003	0004	0005	0006	0007	0008	0009	000A	000B	000C	000D	000E	000F		
1-	0010	⌘	√	◆	IBM	F8FF	0015	0016	0017	0018	0019	001A	001B	001C	001D	001E	001F	
2-	0020	!	"	#	\$	%	&	'	()	*	+	,	-	.	/		
3-	0030	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?	
4-	0040	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	
5-	0050	P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_	
6-	0060	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	
7-	0070	p	q	r	s	t	u	v	w	x	y	z	{		}	~	007F	
8-	0080	À	Á	Â	Ã	Ä	Å	Ö	Û	á	à	â	ã	ä	å	ç	é	è
9-	0090	ê	ë	ì	í	î	ï	ñ	ó	ò	ô	õ	ö	ù	ú	û	ü	
A-	00A0	ÿ	°	€	£	§	•	¶	ß	®	©	™	'	''	≠	Æ	Ø	
B-	00B0	∞	±	≤	≥	¥	µ	ð	Σ	Π	π	∫	∂	∂	∂	∂	∂	
C-	00C0	ˆ	˜	√	f	≈	Δ	«	»	...	À	Ã	Ö	Œ	œ			
D-	00D0	-	—	“	”	‘	’	◊	ÿ	ÿ	/	□	<	>	fi	fl		
E-	00E0	‡	·	ˆ	˜	‰	À	Á	È	É	Í	Ì	Î	Ì	Ó	Ô		
F-	00F0	IBM	Ò	Ú	Û	Ü	Ù	ˆ	˜	ˆ	˜	ˆ	˜	ˆ	˜	ˆ	˜	

problems with older digital fonts: limited number of available character slots



For example, you want to include all these in a document:

- ▶ ligatures
- ▶ Greek text, using same font
- ▶ special symbols (e.g., phonetics)
- ▶ ornaments
- ▶ true small caps
- ▶ old style figures, etc.



problems with the expert/symbol set solution

- ▶ extra cost (sold separately)
 - ▶ major problems with searching
 - ▶ major problems with hyphenation
 - ▶ change of font is necessary to include expert glyphs
-
- ▶ **never use a separate “expert”/“symbol” font** to insert special characters
 - ▶ use **one** font which is **Unicode-encoded** (= OpenType), and insert the special symbol from that font

Unicode (ISO 10646)

basic features

- ▶ an industry standard designed to assign unique character codes (codepoints) to practically **all characters used by humanity**
- ▶ the Unicode codepoint is a hexadecimal (hex) number associated with a character
- ▶ coderange is between $0 \dots 2^{16}$ (decimal $0 \dots 65,536$), but this can be multiplied 2×2^{16} , 3×2^{16} , etc.

some Unicode codepoints

a	→	97	0x0061	Я	→	1103	0x044F
á	→	225	0x00E1	Ń	→	1488	0x05D0
ą	→	261	0x0105	☺	→	9787	0x263B
α	→	945	0x03B1	練	→	32244	0x7DF4

Unicode subranges/planes (examples)

(<http://unicode.org/>)

1. Basic Multilingual Plane

- ▶ Basic Latin (English)
- ▶ Latin Extended-A (Central Europe, Baltic, Esperanto. . .)
- ▶ Latin Extended-B
- ▶ International Phonetic Alphabet
- ▶ diacritical marks
- ▶ Greek, Coptic
- ▶ Cyrillic, Armenian
- ▶ Semitic, Arabic
- ▶ various other scripts (Chinese, Japanese. . .)
- ▶ punctuation, digits, various symbols

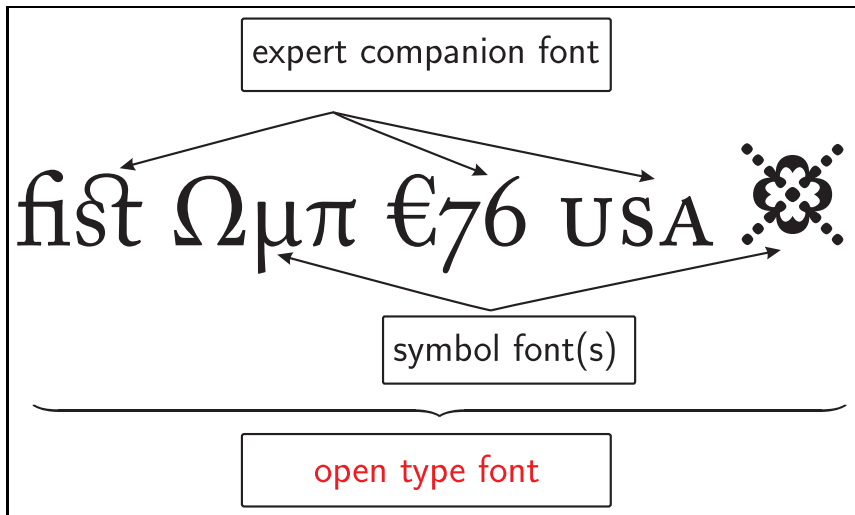
Unicode subranges/planes (examples)

2. Supplementary Multilingual Plane

- ▶ historic & unusual scripts: e.g. cuneiform, Phoenician, etc.
- ▶ musical notations
- ▶ math symbols

Scripts proposed for addition

- ▶ Old Hungarian runic script
- ▶ Tolkien's *The Lord of the Rings* script



OpenType and Unicode

- ▶ the OT technology makes it possible to accommodate **all the Unicode character set**
- ▶ instead of one font for each language group (Western Roman, CE, Baltic, etc.), OT character sets can include all of these code pages in **one single font**
- ▶ OT makes **advanced typographic and language dependent features** available in one font
- ▶ OT font files are **cross-OS**
- ▶ OT fonts can be used on older systems, too (but only the first 256 characters will be recognized)
- ▶ more & more applications take advantage of OT features

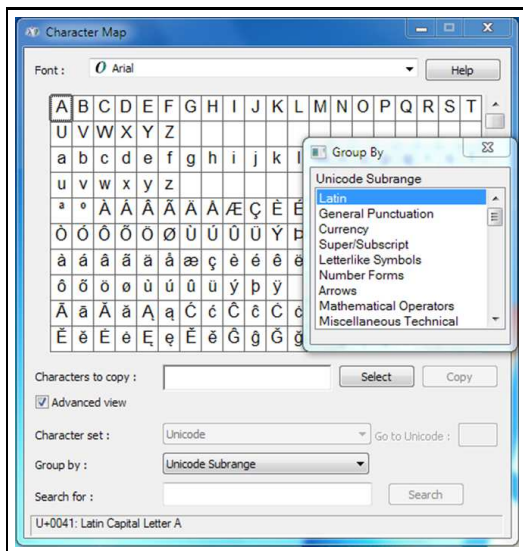
always, always use Unicode

- ▶ Unicode is now used by all modern operating systems as the default character codepage (well, from Windows 11...)
- ▶ Unicode-encoded fonts are also part of all modern OSs
- ▶ **there is no reason why you should not use Unicode!**
- ▶ so: set your text encoding to Unicode (typically: **utf-8**) & use OT fonts

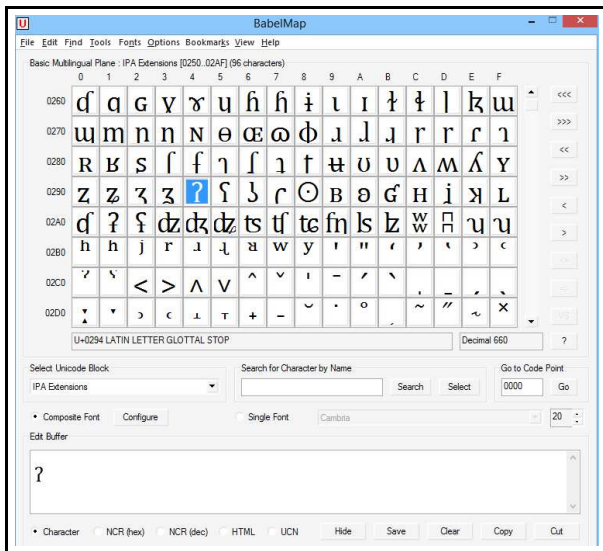
how can all these Unicode subranges be accessed?

- ▶ character map applications
- ▶ use typesetting applications that can access OpenType smartfont tags

Unicode ranges in an OT font



helpful application: BabelMap



OpenType smartfont tags/options

properties

- ▶ they **automate the process of accessing special characters**, features in a Unicode-encoded OT font
- ▶ 4-letter identifiers that can be switched on for a document, or part of it, without changing to another font
- ▶ two main types:
 1. **script/language**-related tags
 2. **layout feature**-related tags
- ▶ official list of registered tags:
<https://docs.microsoft.com/en-us/typography/opentype/spec/ttoreg>

layout tag example: ligatures

- ▶ liga is a smartfont layout tag
- ▶ when it is switched on – and when a font supports it –, it will **automatically** use ligatures instead of sequences of characters
- ▶ e.g., you type `f i` but you will automatically get `f i`

OpenType tags

main **script/language** tags (currently: around 40 defined)

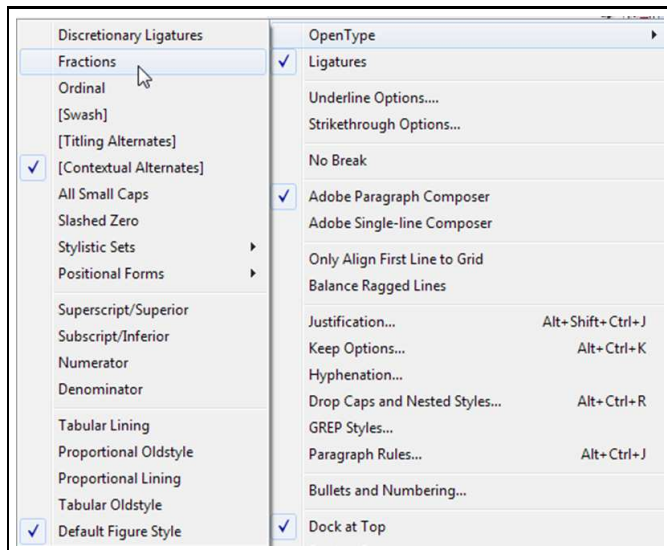
- ▶ activate **language-specific typographic options**
- ▶ latn: Latin alphabet
- ▶ cyrl: Cyrillic
- ▶ arab: Arabic
- ▶ grek: Greek
- ▶ ...

OpenType

main **layout** tags (currently: around 134 defined)

- ▶ all caps, small caps, petite caps (`case`, `smcp`, `c2sc`, `c2pc`)
- ▶ capital spacing (`cpssp`)
- ▶ standard ligatures (`liga`) or special discretionary ligatures (`dlig`)
- ▶ historical forms (`hist`)
- ▶ fractions (`frac`)
- ▶ ordinals, superscript, subscript (`ordn`, `sup`, `sub`)
- ▶ figure types: lining – oldstyle; tabular – proportional (`lnum`, `onum`, `tnum`, `pnum`)
- ▶ swash, stylistic alternates, sets (`swsh`, `salt`, `ssxx`)
- ▶ access all alternates (`aalt`)
- ▶ ...

OT layout tags in Adobe InDesign



OpenType layout features

demo

smcp: small capitals

Minion Pro (Adobe):

usa USA *usa*
USA USA *USA*

c2sc: small capitals from capitals

Minion Pro (Adobe):

unicef UNICEF
unicef UNICEF

pcap: petite caps (from lowercase)

Mrs Eaves (Emigre):

UNICEF unicef
all caps

UNICEF UNICEF
small caps petite caps

acex ACEx

uni: unicase

Filosofia (Emigre):

Filosofia Unicase
FILOSOFIA UNICASE

case: case-sensitive forms

Mrs Eaves (Emigre):

3 ip-sum:

3 IP-SUM:

3 IP-SUM:

cpsp: capital spacing

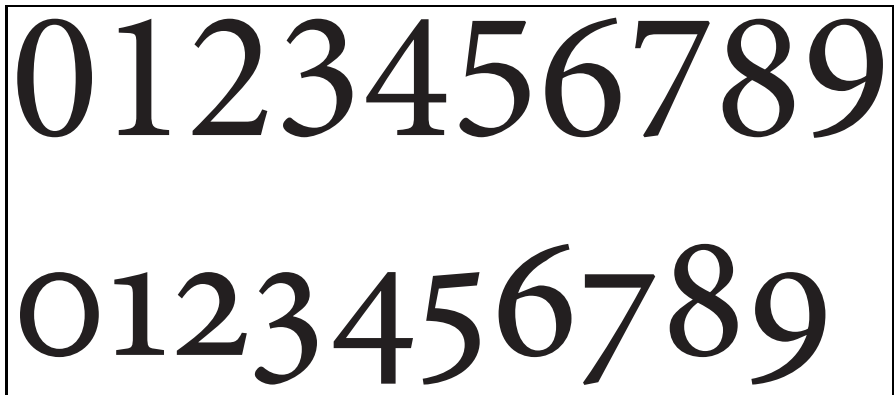
Adobe Garamond Pro (Adobe):



ABCDEFGHIJKLM
ABCDEFGHIJKLM

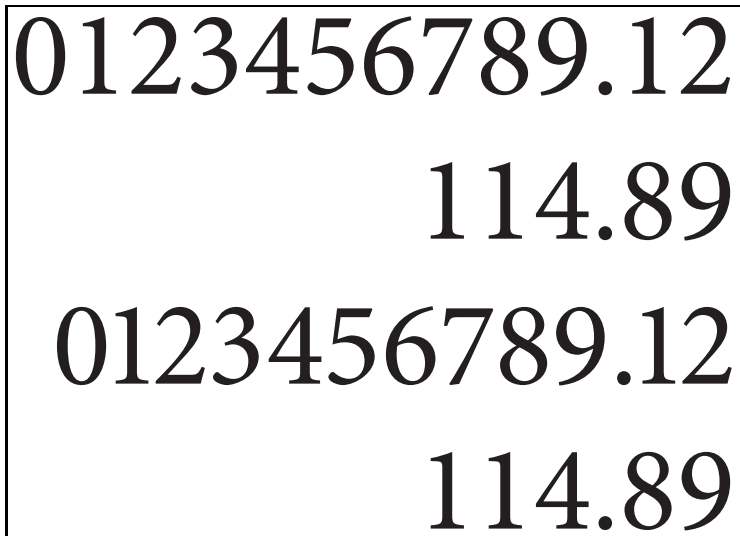
lnum vs. onum: lining vs. oldstyle numbers

Minion Pro (Adobe):



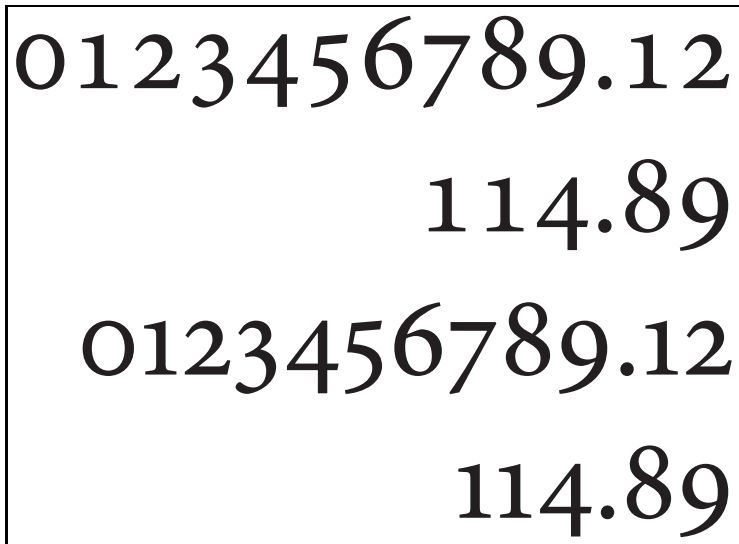
lnum + tnum vs. pnum: tabular vs. proportional numbers

Minion Pro (Adobe):



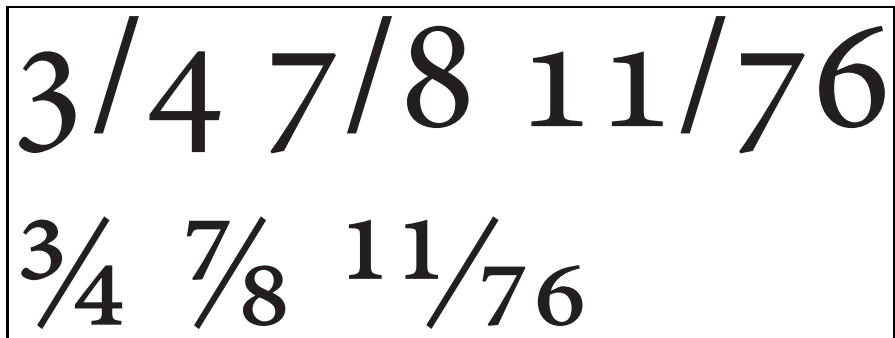
onum + tnum vs. pnum: tabular vs. proportional numbers

Minion Pro (Adobe):



frac: fractions

Minion Pro (Adobe):



3/4 7/8 11/76

3/4 7/8 11/76

sups: superscript

Minion Pro (Adobe):

0	1	2	3	4	5	6	7	8	9	a	e	st
0	1	2	3	4	5	6	7	8	9	a	e	st
footnote ⁶ 1 st 1 ^{er} 2 ^e												
footnote ⁶ 1 st 1 ^{er} 2 ^e												

liga: standard ligatures

Adobe Caslon Pro (Adobe):

fine fine soufflé fjord
fine fine soufflé fjord

dlig: discretionary ligatures

Mrs Eaves (Emigre):



strict spice *strict spice*
strict spice *strict spice*

dlig: discretionary ligatures

Mrs Eaves (Emigre):

hogy happy lucky waggon

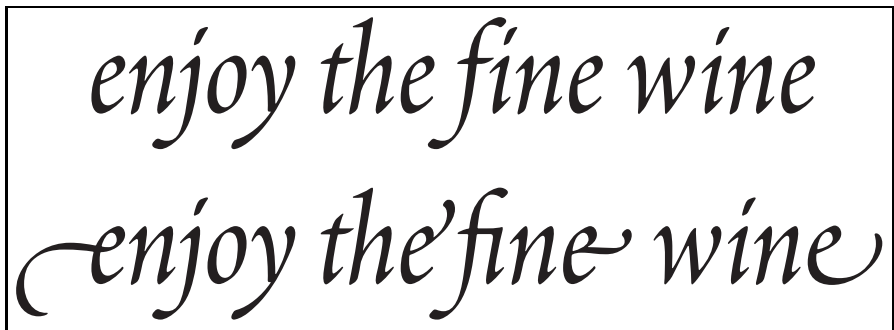
hogy happy lucky waggon

hogy happy lucky waggon TheTHE

hogy happy lucky waggon TheTE

`init`, `medi`, `fin`: initial, medial, final forms

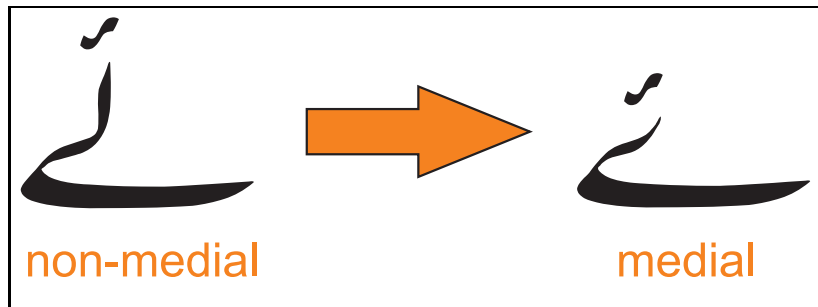
Poetica (Adobe):



enjoy the fine wine
enjoy the'fine wine

init, medi: initial vs. medial forms in Arabic

Nafees Nasta'leeq (the character 'yeh barree with hamza' above):



some drawbacks of OT fonts

- ▶ difficult to manage (lot of clicking), need for clever typesetting applications
- ▶ most are very expensive, and rarely sold separately (“all-or-nothing”)
- ▶ not all OT fonts have the same character coverage, naming does not always help
- ▶ **Std** = basic Unicode coverage, few smartfont options
Pro = medium Unicode coverage, more smartfont options
Premier Pro = wide range of Unicode coverage, lot of smartfont options
- ▶ a list of the most comprehensive free Unicode fonts:
en.wikipedia.org/wiki/List_of_typefaces#Unicode_fonts

Adobe Garamond Premier Pro: Unicode coverage

<https://www.fonts.com/font/adobe/garamond-premier/roman>

Garamond Premier Pro Roman

Designers: Robert Slimbach, Claude Garamond / Design foundry: Unknown / Source Foundry: Adobe / Classifications: Serif, Old Style Serif

Buy from €35 View Family

Select Font Format Add to Wish List

Try It Waterfall Character Map OpenType Technical Details Browser Previews

Desktop Web Font

View available characters! Select the font and unicode character range that you wish to display from the dropdown lists below. The character map feature will automatically display all available characters in the grid!

Select unicode range Num Chars.

0400..04FF Cyrillic 108

È	É	Ђ	ƒ	€	Š
È	É	Ђ	ƒ	€	Š
0400	0401	0402	0403	0404	0405
ı	İ	Ј	Љ	Њ	Ћ
ı	İ	Ј	Љ	Њ	Ћ
0406	0407	0408	0409	040A	040B
Ɔ	Ɔ	Ɔ	Ɔ	Ɔ	Ɔ
Ɔ	Ɔ	Ɔ	Ɔ	Ɔ	Ɔ
040C	040D	040E	040F	0410	0411
В	Г	Д	Е	Ж	З
В	Г	Д	Е	Ж	З

Adobe Garamond Premier Pro: smartfont options

<https://www.fonts.com/font/adobe/garamond-premier/roman>

Garamond Premier Pro Roman

[mbach, Claude Garamond](#) / Design foundry: [Unknown](#) / Source Foundry: [Adobe](#) / Classifications: [Serif, Old Style Serif](#)

Buy from €35
View Family

T Select Font Format ★ Add to Wish List

Try it
Waterfall
Character Map
OpenType
Technical Details
Browser Previews

aA
1/3
7/8
ff
123
2/
619
1st
123
H:0
H²
123
456
e
Bb
SM

d
f
H₂
ct
SS
01
Ø
SS
02
SS
03
SS
04
A·B
T_O
SIZE

Small Capitals From Capitals

Tag: `c2sc`

Function: Small Capitals From Capitals

web fonts

- ▶ fonts that are **fine-tuned for onscreen displays** at all kinds of resolutions, and are typically **hosted on servers** and loaded from them in web documents
- ▶ other than these, they are OT fonts, with the same capabilities (Unicode, smart font options, etc.)
- ▶ **Web Open Font Format (woff, woff2)**: this format is what is now fully ratified by the World Wide Web Consortium
- ▶ others: TT-flavoured OpenType (ttf), Scalable Vector Graphics font (svg) – mostly on MacOS and iOS Safari

font site tips

- ▶ good fonts come from a modern, professional, well-designed website, which never have ads
- ▶ good fonts clearly list the name of the designer, include information about the design process of the font, come with a clear license
- ▶ good fonts have features like language support, full glyph sets, ligatures, kerning pairs, OpenType smartfont features
- ▶ affordable font outlets: www.futurefonts.xyz; Font of the month club (djr.com/font-of-the-month-club); Adobe fonts (fonts.adobe.com/fonts)
- ▶ good free fonts: www.fontshare.com; velvetyne.fr; fonts.google.com; www.fontsquirrel.com
- ▶ large font corporate sites (expensive): www.myfonts.com; monotype.com; www.linotype.com; fonts.com