

Strings & Representation

Learning Goals

1. Understanding Representational Harms



History of Symbol Representation

**"Typewheel" for the telegraph,
covers 56 characters**

1846

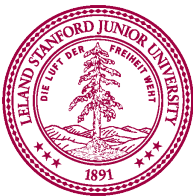


History of Symbol Representation

4-digit character-to-code mapping system
for the telegraph

1881

〇〇〇〇	〇一八	〇一六	〇一四	〇一二	〇一〇	〇〇八	〇〇六	〇〇四	〇〇二	〇〇〇
	傍	𠂔	伏	𠂔	仕	丁	尸	丁	丁	丁
一	〇一八二	〇一六二	〇一四二	〇一二二	〇一〇二	〇〇八二	〇〇六二	〇〇四二	〇〇二二	〇〇〇二
	佳	位	伐	伏	仁	也	丁	丁	丁	丁
二	〇一八三	〇一六三	〇一四三	〇一二三	〇一〇三	〇〇八三	〇〇六三	〇〇四三	〇〇二三	〇〇〇三
	餅	低	休	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔	𠂔
三	〇一八四	〇一六四	〇一四四	〇一二四	〇一〇四	〇〇八四	〇〇六四	〇〇四四	〇〇二四	〇〇〇四
	信	住	伙	仔	仆	交	事	乏	丰	丈
四	〇一八五	〇一六五	〇一四五	〇一二五	〇一〇五	〇〇八五	〇〇六五	〇〇四五	〇〇二五	〇〇〇五
	𠂔	佐	伯	伶	仇	亥	𠂔	乖	串	三
五	〇一八六	〇一六六	〇一四六	〇一二六	〇一〇六	〇〇八六	〇〇六六	〇〇四六	〇〇二六	〇〇〇六
	使	佑	估	仲	今	亦	乘	𠂔	上	𠂔
六	〇一八七	〇一六七	〇一四七	〇一二七	〇一〇七	〇〇八七	〇〇六七	〇〇四七	〇〇二七	〇〇〇七
	侃	估	佚	伙	介	享	𠂔	𠂔	下	𠂔
七	〇一八八	〇一六八	〇一四八	〇一二八	〇一〇八	〇〇八八	〇〇六八	〇〇四八	〇〇二八	〇〇〇八
	來	何	你	𠂔	仍	𠂔	于	𠂔	不	𠂔
八	〇一八九	〇一六九	〇一四九	〇一二九	〇一〇九	〇〇八九	〇〇六九	〇〇四九	〇〇二九	〇〇〇九
	侈	伐	𠂔	𠂔	仔	亨	云	九	𠂔	𠂔
九	〇一九〇	〇一七〇	〇一五〇	〇一三〇	〇一一〇	〇〇九〇	〇〇七〇	〇〇五〇	〇〇三〇	〇〇一〇
	例	余	伴	伴	仕	京	互	乙	凡	丑
〇	〇一九一	〇一七一	〇一五一	〇一三一	〇一一一	〇〇九一	〇〇七一	〇〇五二	〇〇三二	〇〇一二
	侍	余	伶	攸	他	亭	五	九	丹	且
一	〇一九二	〇一七二	〇一五二	〇一三二	〇一一二	〇〇九二	〇〇七二	〇〇五三	〇〇三三	〇〇一三
	侏	佛	仲	价	仗	亮	井	乞	主	丕
二	〇一九三	〇一七三	〇一五三	〇一三三	〇一一三	〇〇九三	〇〇七三	〇〇五三	〇〇三三	〇〇一三
	𠂔	作	伺	任	付	𠂔	亘	也	𠂔	世
三	〇一九四	〇一七四	〇一五四	〇一三四	〇一二四	〇〇九四	〇〇七四	〇〇五四	〇〇三四	〇〇一四
	𠂔	𠂔	伴	仿	仙	𠂔	互	𠂔	𠂔	丘
四	〇一九五	〇一七五	〇一五五	〇一三五	〇一二五	〇〇九五	〇〇七五	〇〇五五	〇〇三五	〇〇一五
	𠂔	𠂔	似	企	全	𠂔	况	乳	𠂔	丙
五	〇一九六	〇一七六	〇一五六	〇一三六	〇一二六	〇〇九六	〇〇七六	〇〇五六	〇〇三六	〇〇一六
	𠂔	佩	伽	𠂔	𠂔	些	乾	父	丞	𠂔
六	〇一九七	〇一七七	〇一五七	〇一三七	〇一二七	〇〇九七	〇〇七七	〇〇五七	〇〇三七	〇〇一七
	𠂔	𠂔	𠂔	伊	𠂔	亞	亂	乃	丟	𠂔
七	〇一九八	〇一七八	〇一五八	〇一三八	〇一二八	〇〇九八	〇〇七八	〇〇五八	〇〇三八	〇〇一八
	供	𠂔	但	𠂔	代	𠂔	𠂔	久	並	𠂔
八	〇一九九	〇一七九	〇一五九	〇一三九	〇一二九	〇〇九九	〇〇七九	〇〇五九	〇〇三九	〇〇一九
	依	伴	𠂔	伍	令	𠂔	𠂔	𠂔	𠂔	𠂔
九	〇二〇〇	〇一八〇	〇一六〇	〇一四〇	〇一二〇	〇〇一〇	〇〇九〇	〇〇七〇	〇〇五〇	〇〇三〇
	𠂔	𠂔	佈	𠂔	以	人	𠂔	𠂔	之	𠂔



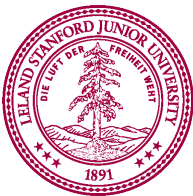
ASCII

"American Standard Code for Information Interchange" covers 256 symbols

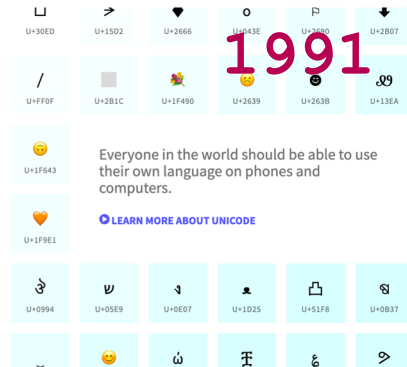
USASCII code chart

Bits					Column				Row			
b ₇	b ₆	b ₅	b ₄	b ₃	b ₂	b ₁	b ₀		0	1	2	3
0	0	0	0	0	0	0	0	0	NUL	DLE	SP	@
0	0	0	0	1	0	0	0	1	SOH	DC1	!	A
0	0	0	1	0	0	0	0	2	STX	DC2	"	B
0	0	0	1	1	0	0	0	3	ETX	DC3	#	C
0	0	1	0	0	0	0	0	4	EOT	DC4	\$	D
0	0	1	0	1	0	0	0	5	ENQ	NAK	%	E
0	0	1	1	0	0	0	0	6	ACK	SYN	&	F
0	0	1	1	1	0	0	0	7	BEL	ETB	'	G
0	1	0	0	0	0	0	0	8	BS	CAN	(H
0	1	0	0	1	0	0	0	9	HT	EM)	I
0	1	0	1	0	0	0	0	10	LF	SUB	*	J
0	1	0	1	1	0	0	0	11	VT	ESC	+	K
0	1	1	0	0	0	0	0	12	FF	FS	,	L
0	1	1	0	1	0	0	0	13	CR	GS	-	M
0	1	1	1	0	0	0	0	14	SO	RS	.	N
0	1	1	1	1	0	0	0	15	SI	US	/	O

1963

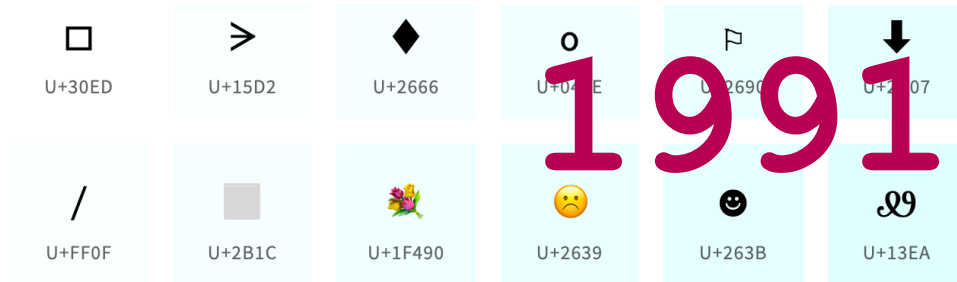


The image shows a detailed view of a mechanical watch movement, possibly a pocket watch, housed in a dark wooden case. The movement is complex, with numerous gears, levers, and a prominent balance wheel at the top. A small white label is placed near the bottom left of the movement, providing technical specifications. The overall appearance is that of a well-preserved historical timepiece.

[illegible]

Unicode

Covers over 1 million symbols,
Full coverage of 90 languages,
basic coverage for 200 languages



Adopt a Character



Emoji



Basic Info



News

Events

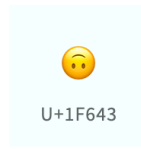
Connect



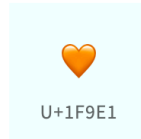
Membership



Press



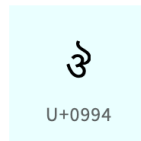
U+1F643



U+1F9E1

Everyone in the world should be able to use
their own language on phones and
computers.

[▶ LEARN MORE ABOUT UNICODE](#)



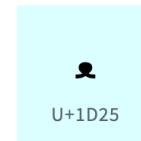
U+0994



U+05E9



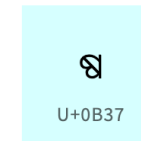
U+0E07



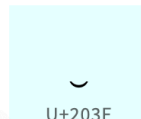
U+1D25



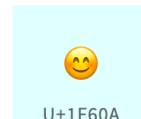
U+51F8



U+0B37



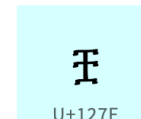
U+203F



U+1F60A



U+03CE



U+127F



U+060F



U+1557



Representing a Burrito in Unicode/UTF-8

Emoji added in 2007

Burrito (U+1F32F) added in 2015

Apple



Google



Microsoft



Unicode and Representation



U+30ED



U+15D2



U+2666



U+043E



U+2690



U+2B07



U+FF0F



U+2B1C



U+1F490



U+2639



U+263B



U+13EA



U+1F643



U+1F9E1

Everyone in the world should be able to use their own language on phones and computers.

[▶ LEARN MORE ABOUT UNICODE](#)



U+0994



U+05E9



U+0E07



U+1D25



U+51F8



U+0B37



U+203F



U+1F60A



U+03CE



U+127F



U+060F



U+1557



Adopt a Character

Emoji

Basic Info

News

Events

Connect

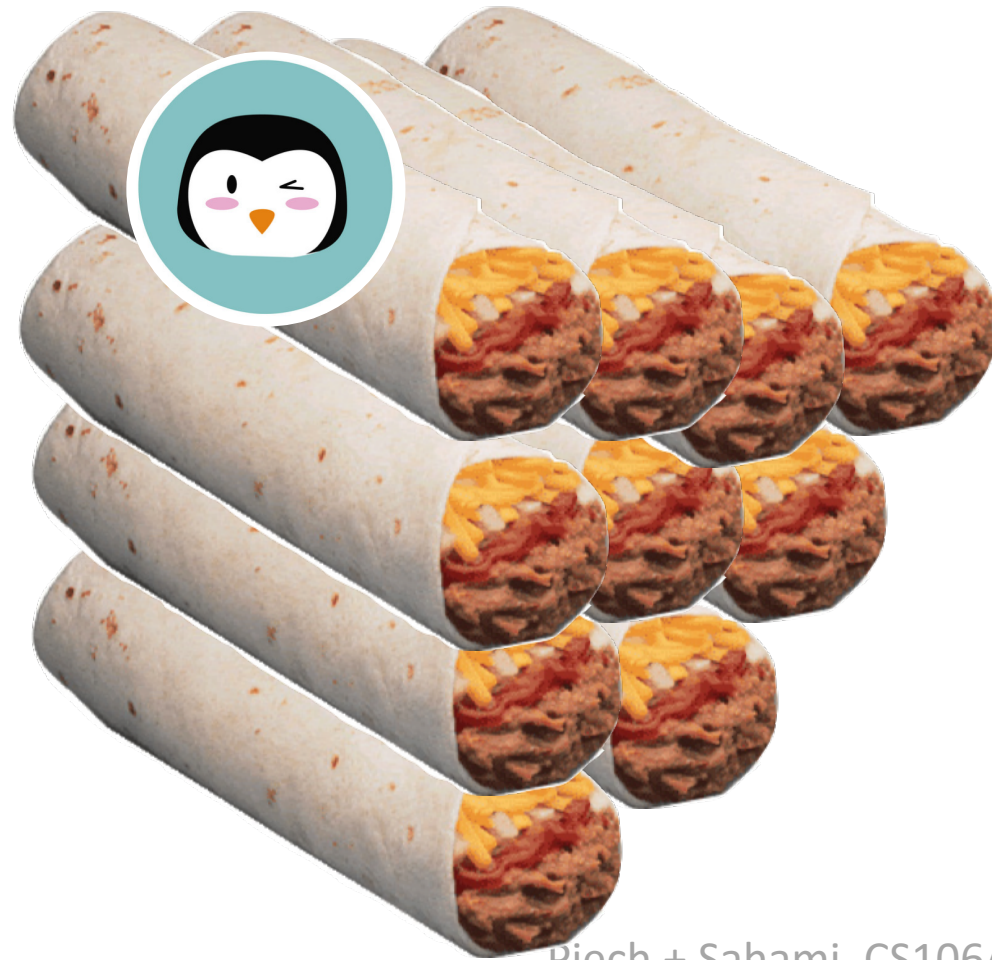
Membership

Press

What kind of harm is a
lack of representation?

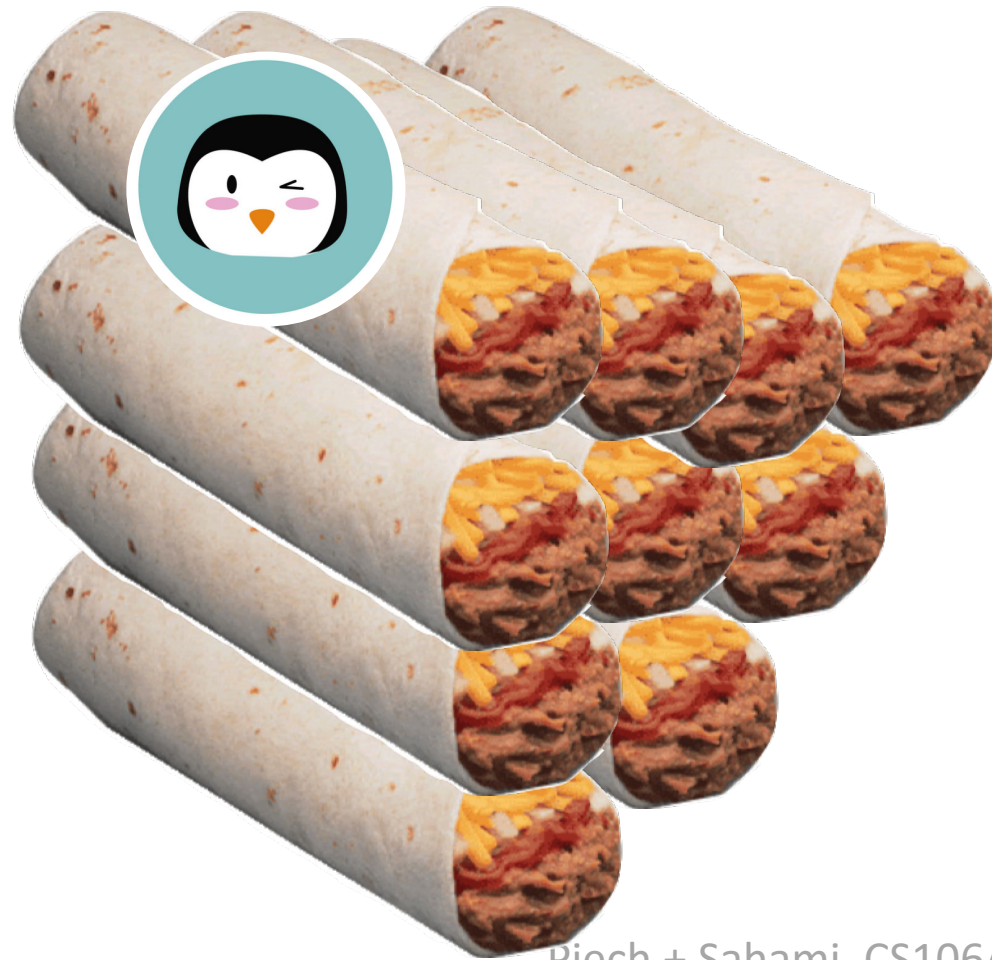
Distributive Harms

How should things or outcomes be distributed?



Distributive Harms

Equality of Opportunity: everyone has same opportunity to pursue good thing (burritos)



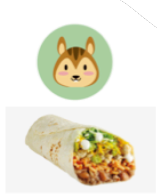
Distributive Harms

Equality of Outcome: everyone gets same good things (burritos) and same responsibilities



Distributive Harms

How should things or outcomes be distributed?



Equality of opportunity:
Everyone has same opportunity to
pursue good things (burritos)



Equality of outcome:
everyone has same good things
(burritos) and same responsibilities



Equality of welfare:
Everyone has well-being/happiness
(but only some people need burritos
for happiness)

Representational Harms

Am I represented in this system?

Can I express myself in it?



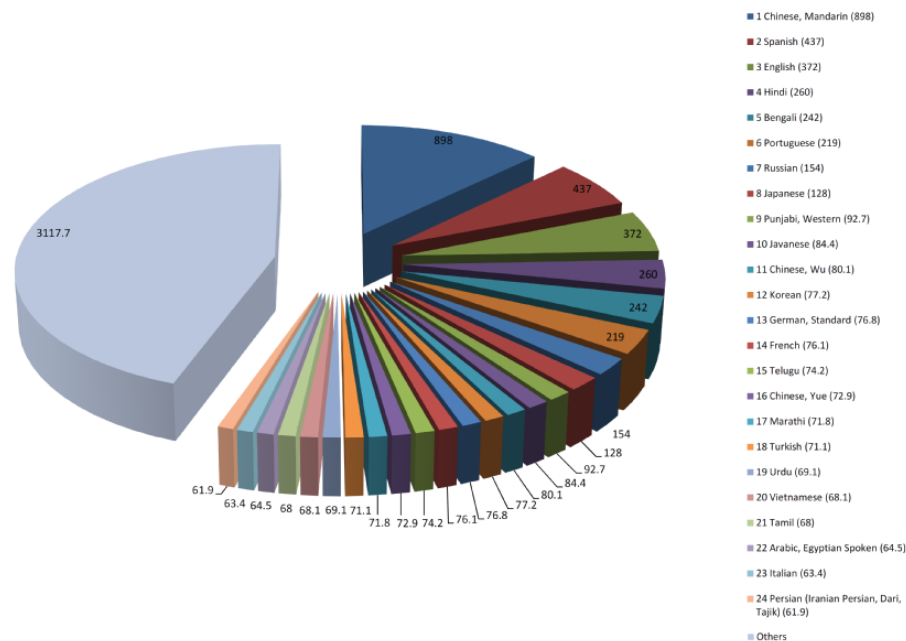
Does this system recognize me, my culture, and my self-expression?



Language Representation in Unicode

There are at least 7,100 languages spoken today. Of those, there are at least 24 with 50 million speakers.

Are the languages you speak representable in Unicode?



Languages with at least 50 million first-language speakers, millions

Total: 7000 million (2011)

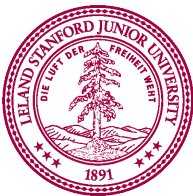


Representation beyond Unicode



What else is needed for linguistic representation besides Unicode?

- Keyboard support
 - Texting keyboards!
- Rendering support
 - Skin-tone emojis



Emoji Representation in Unicode

2015: "Family" aka "family_man_woman_boy"

Codepoints

👨 U+1F468

U+200D

👩 U+1F469

U+200D

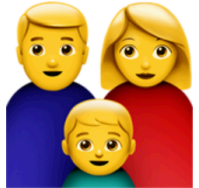
👦 U+1F466

Shortcodes

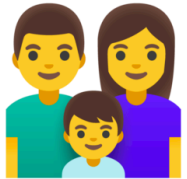
:family: (Slack)

:family_man_woman_boy: (Github, Emojipedia)

Apple



Google



Microsoft



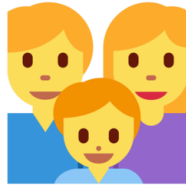
Samsung



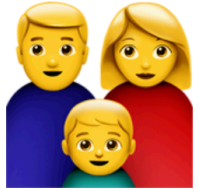
Emoji Representation in Unicode

2015: "Family" aka "family_man_woman_boy"

Twitter



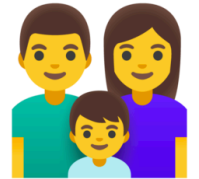
Apple



Facebook



Google



JoyPixels



Microsoft



OpenMoji



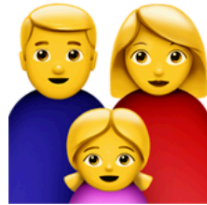
Samsung



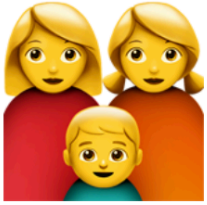
Emoji Representation in Unicode

`"family_man_woman_girl"` `"family_woman_woman_boy"`

Apple



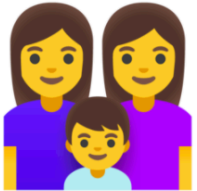
Apple



Google



Google



Microsoft



Microsoft



Samsung



Samsung



Scientific Representation

Representation of people is different from representation of natural phenomena.

This is a special challenge for social science.

