How do computers compute?

Representing data

- Representing all data with 1s and 0s, on/off, true/false
- Numbers -> binary (positional like decimal)
- Letters/symbols -> ASCII, Unicode, UTF8

Positional Notation

- Binary numbers, like decimal numbers, use *place notation*

$$1101 = 1\times1000 + 1\times100 + 0\times10 + 1\times1$$

$$= 1\times10^3 + 1\times10^2 + 0\times10^1 + 1\times10^0$$

except that the base is 2 not 10

$$1101 = 1\times8 + 1\times4 + 0\times2 + 1\times1$$

$$= 1\times2^3 + 1\times2^2 + 0\times2^1 + 1\times2^0$$

Base or radix

1101 in binary is 13 in decimal

Binary to Decimal

- What is the decimal equivalent of the binary number 10011?
- A. 11
- B. 12
- C. 13
- D. 18
- E. 19

Decimal to Binary

- What is the binary representation of the decimal value 21?
- A. 10010
- B. 10100
- C. 10101
- D. 10110
- E. 10111

Binary combinations, True/False possibilities

- One bit
- 0
- 1

- Two Bits
- 00
- 01
- 10
- 11

- Three bits
- 000
- 001
- 010
- 011
- 100
- 101
- 110
- 111

- Four bits
- 0000
- 0001
- 0010
- 0011
- 0100
- 0101
- 0110
- 0111
- 1000
- 1001
- 1010
- 1011
- 1100
- 1101
- 1110
- 1111

Binary Addition

0	0	1	1
+0	+1	+0	+1

Binary Addition

```
  001101
+ 010111
```

Binary Addition

$$\begin{array}{r} 011101 \\ +010011 \\ \hline \end{array}$$

A. 101010
B. 110101
C. 110000
D. 111000
E. 101101

Bytes

- A byte is eight bits treated as a unit
  - Adopted by IBM in 1960s
  - A standard measure ever since
  - Bytes encode the Latin alphabet using ASCII -- the American Standard Code for Information Interchange

0101 0101
0101 0111

ASCII

ASCII	0000 0001	0000 0010	0000 0011	0000 0100	0000 0101	0000 0110	0000 0111	0000 1000	0000 1001	0000 1010	0000 1011	0000 1100	0000 1101	0000 1110	0000 1111
0000	NU	SH	SX	EX	ET	EQ	AK	BL	BS	HT	LF	YT	FF	CR	S0	SI
0001	DL	D1	D2	D3	D4	NK	SY	EΣ	CN	EM	SB	EC	FS	GS	RS	US
0010		!	"	#	$	%	&	'	(	)	*	+	,	-	.	/
0011	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
0100	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
0101	P	Q	R	S	T	U	V	W	X	Y	Z	[	\	]	^	_
0110	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
0111	p	q	r	s	t	u	v	w	x	y	z	{	|	}	~	DT
1000	80	81	82	83	IN	NL	SS	ES	HS	HJ	YS	PD	PV	RI	S2	S3
1001	DC	P1	PZ	SE	CC	MM	SP	EP	Q8	QQ	QA	CS	ST	OS	PM	AP
1010	A0	¡	¢	£	♀	¥	¦	§	¨	©	♂	«	¬	-	®	¯
1011	°	±	²	³	´	µ	¶	·	¸	¹	º	»	¼	½	¾	¿
1100	À	Á	Â	Ã	Ä	Å	Æ	Ç	È	É	Ê	Ë	Ì	Í	Î	Ï
1101	Ð	Ñ	Ò	Ó	Ô	Õ	Ö	×	Ø	Ù	Ú	Û	Ü	Ý	Þ	β
1111	ð	ñ	ò	ó	ô	õ	ö	÷	ø	ù	ú	û	ü	ý	Þ	ÿ

0100 0011
0101 0011
0101 0000

0100 0111|0110 1111|0010 0000|0101 0011|0110 1100|0111 0101|0110 0111|0111 0011

ASCII

What is the first letter of the message at the bottom?

A. G

B. t

ASCII	0000 0000	0000 0001	0000 0010	0000 0011	0000 0100	0000 0101	0000 0110	0000 0111	0000 1000	0000 1001	0000 1010	0000 1011	0000 1100	0000 1101	0000 1110	0000 1111
0000	NU	SH	SX	EX	ET	EQ	AK	BL	BS	HT	LF	YT	FF	CR	S0	SI
0001	DL	D1	D2	D3	D4	NK	SY	EΣ	CN	EM	SB	EC	FS	GS	RS	US
0010		!	"	#	$	%	&	'	(	)	*	+	,	-	.	/
0011	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
0100	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
0101	P	Q	R	S	T	U	V	W	X	Y	Z	[	\	]	^	_
0110	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
0111	p	q	r	s	t	u	v	w	x	y	z	{	|	}	~	DT
1000	80	81	82	83	IN	NL	SS	ES	HS	HJ	YS	PD	PV	RI	S2	S3
1001	DC	P1	PZ	SE	CC	MM	SP	EP	Q8	QQ	QA	CS	ST	OS	PM	AP
1010	A0	¡	¢	£	♀	¥	¦	§	¨	©	♂	«	¬	-	®	¯
1011	°	±	²	³	´	µ	¶	·	¸	¹	º	»	¼	½	¾	¿
1100	À	Á	Â	Ã	Ä	Å	Æ	Ç	È	É	Ê	Ë	Ì	Í	Î	Ï
1101	Ð	Ñ	Ò	Ó	Ô	Õ	Ö	×	Ø	Ù	Ú	Û	Ü	Ý	Þ	β
1111	ð	ñ	ò	ó	ô	õ	ö	÷	ø	ù	ú	û	ü	ý	þ	ÿ

0100 0111|0110 1111|0010 0000|0101 0011|0110 1100|0111 0101|0110 0111|0111 0011

UTF-8

Uniform Transformation Format for bytes

(UTF-8) is universal … all characters have a place: 1-4 Bytes

(Unicode has space for over 1 million symbols)

لماذا لا يتكلمون اللّغة العربية فحسب؟
Защо те просто не могат да говорят **български**?
Per què no poden simplement parlar en **català**?
他們爲什麼不說中文（台灣）？
Proč prostě nemluví **česky**?
Hvorfor kan de ikke bare tale **dansk**?
Warum sprechen sie nicht einfach **Deutsch**?
Μα γιατί δεν μπορούν να μιλήσουν **Ελληνικά**;
Why can't they just speak **English**?
¿Por qué no pueden simplemente hablar en **castellano**?
Miksi he eivät yksinkertaisesti puhu **suomea**?
Pourquoi, tout simplement, ne parlent-ils pas **français** ?
למה הם פשוט לא מדברים **עברית**?
Miért nem beszélnek egyszerűen **magyarul**?
Af hverju geta þeir ekki bara talað **íslensku**?
Perché non possono semplicemente parlare **italiano**?
なぜ、みんな**日本語**を話してくれないのか？
세계의 모든 사람들이 **한국어**를 이해한다면 얼마나 좋을까?
Waarom spreken ze niet gewoon **Nederlands**?
Hvorfor kan de ikke bare snakke **norsk**?
Dlaczego oni po prostu nie mówią po **polsku**?
Porque é que eles não falam em **Português (do Brasil)**?
Oare ăştia de ce nu vorbesc **româneşte**?
Почему же они не говорят **по-русски**?
Zašto jednostavno ne govore **hrvatski**?
Pse nuk duan të flasin vetëm **shqip**?
Varför pratar dom inte bara **svenska**?
ทำไมเขาถึงไม่พูดภาษาไทย
Neden **Türkçe** konuşamıyorlar?

UTF-8

Bits of code point	First code point	Last code point	Bytes in sequence	Byte 1	Byte 2	Byte 3	Byte 4
7	U+0000	U+007F	1	`0xxxxxxx`			
11	U+0080	U+07FF	2	`110xxxxx`	`10xxxxxx`		
16	U+0800	U+FFFF	3	`1110xxxx`	`10xxxxxx`	`10xxxxxx`	
21	U+10000	U+1FFFFF	4	`11110xxx`	`10xxxxxx`	`10xxxxxx`	`10xxxxxx`

UTF-8

Bits of code point	First code point	Last code point	Bytes in sequence	Byte 1	Byte 2	Byte 3	Byte 4
7	U+0000	U+007F	1	0xxxxxxx			
11	U+0080	U+07FF	2	110xxxxx	10xxxxxx		
16	U+0800	U+FFFF	3	1110xxxx	10xxxxxx	10xxxxxx	
21	U+10000	U+1FFFFF	4	11110xxx	10xxxxxx	10xxxxxx	10xxxxxx

What is the first Unicode value represented by this sequence?
11101010 10000011 10000111 00111111 11000011 10000000

A. 0000000001101010
B. 0000000011101010
C. 0000001010000111
D. 1010000011000111

UTF-8

Bits of code point	First code point	Last code point	Bytes in sequence	Byte 1	Byte 2	Byte 3	Byte 4
7	U+0000	U+007F	1	0xxxxxxx			
11	U+0080	U+07FF	2	110xxxxx	10xxxxxx		
16	U+0800	U+FFFF	3	1110xxxx	10xxxxxx	10xxxxxx	
21	U+10000	U+1FFFFF	4	11110xxx	10xxxxxx	10xxxxxx	10xxxxxx

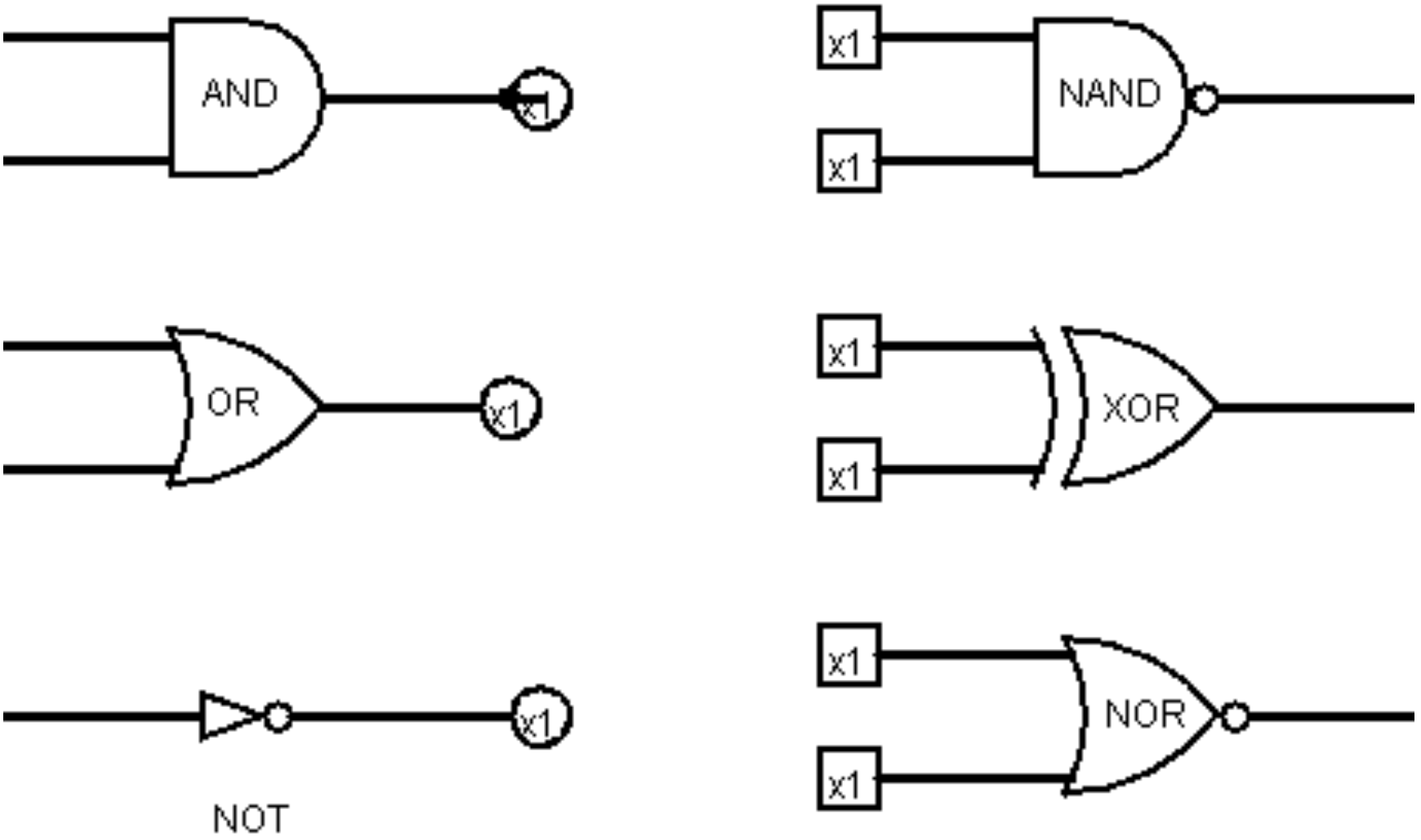
How many Unicode characters are represented by this sequence?
11101010 10000011 10000111 00111111 11000011 10000000

A. 1
B. 2
C. 3
D. 4
E. 5

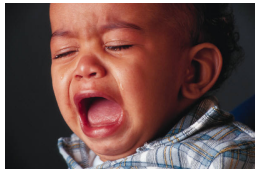
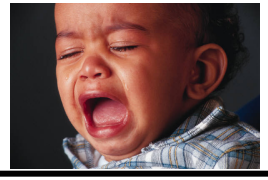
But how do they compute?

- Basic operations are arithmetic, compare, read/write memory, test and jump (to a different memory location for the next instruction).

Fundamental units of computers: Logic Gates



Truth Table for *And (using True and False)*

P	Q	P and Q
True 	True 	True
True 	False 	False
False 	True 	False
False 	False 	False

Truth Table for *And (using 0 and 1)*

P	Q	P and Q
1	1	
1	0	
0	1	
0	0	

Truth Table for *Or (using 0 and 1)*

P	Q	P or Q
1	1	
1	0	
0	1	
0	0	

Truth Table for *Not And (using 0 and 1)*

P	Q	P AND Q	NOT (P AND Q)
1	1	1	0
1	0	0	1
0	1	0	1
0	0	0	1

(NOT P) OR (NOT Q) vs. NOT (P AND Q)

P	Q	NOT P	NOT Q	P AND Q	NOT (P AND Q)	(NOT P) OR (NOT Q)	
1	1	0	0	1	0	0	
1	0	0	1	0	1	1	
0	1	1	0	0	1	1	
0	0	1	1	0	1	1	

(NOT P) OR (NOT Q) vs. NOT (P AND Q)

P	Q	NOT P	NOT Q	P AND Q	NOT (P AND Q)	(NOT P) OR (NOT Q)	
1	1	0	0	1	0	0	
1	0	0	1	0	1	1	
0	1	1	0	0	1	1	
0	0	1	1	0	1	1	

- NOT (P $\wedge$ Q) = NOT P $\vee$ NOT Q
- This is DeMorgan's Law of Boolean Algebra

P and Q or R

- What is P and Q or R if P is true, Q is false, and R is true?
- A. True
- B. False

- There are 10 types of people in this world, those who know binary, and those who do not.

Exclusive-OR == XOR

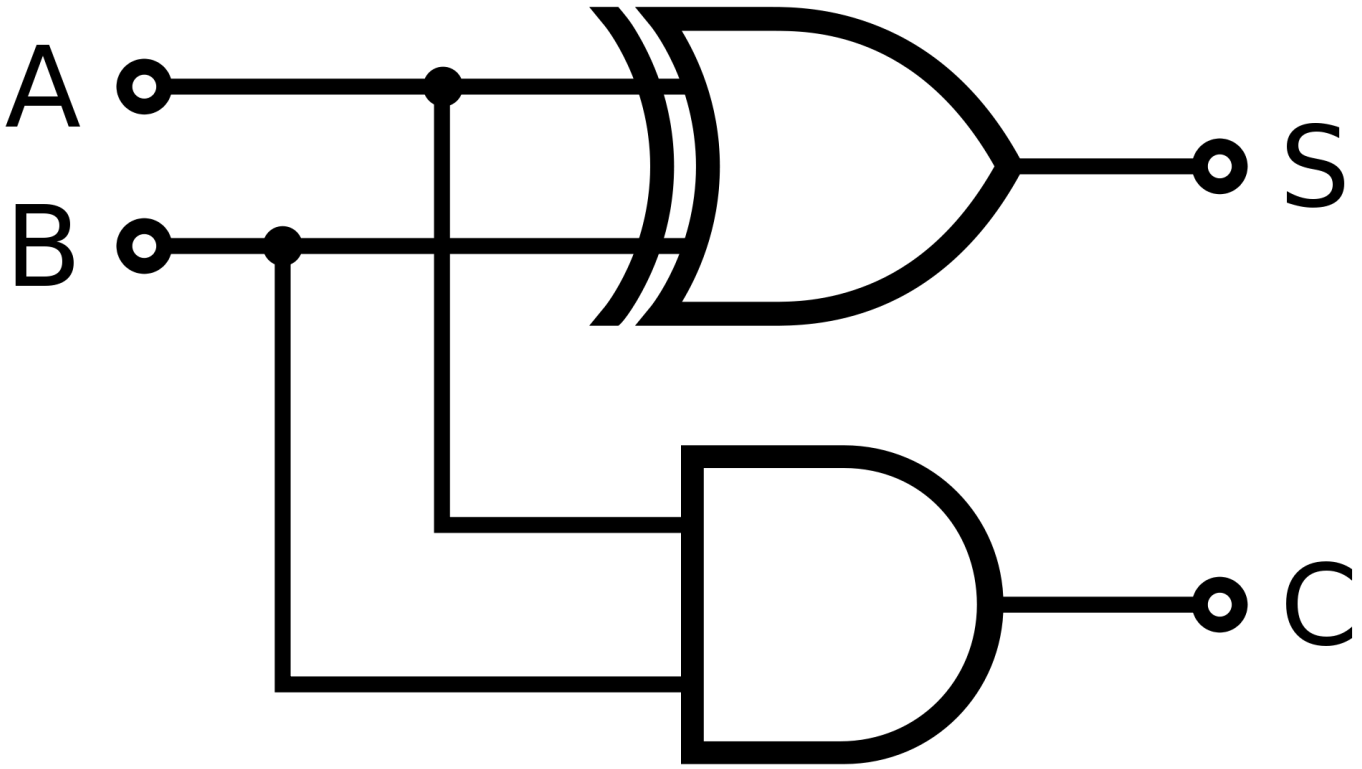
- Consider two propositions, either of which may be true or false
- Exclusive-or is the relationship between them when JUST ONE OF THEM is true.
- It EXCLUDES the case when both are true,so exclusive-or of the two is…
- False when both are true or both are false, and true in the other two cases.

Truth Table for *XOR* *(using 0 and 1)*

P	Q	P xor Q
1	1	
1	0	
0	1	
0	0	

What would you ever want XOR for anyway?

- http://en.wikipedia.org/wiki/Adder_(electronics)
- Binary Addition



Summary: It All Works Because of Digital

- Key principle: information is represented as simply the presence or absence of a phenomenon at a given place and time!
- Phenomenon in computers: Electrical output on a line
  - Hole in punch card, early example.
- Logic Gates
  - Charge on line
  - No charge