#### Induction

• How did your parents represent you to the world when you were born?

• Why do you think we use base 10 as our "normal" counting system?

# How Computers Represent Data

THE FIRST ABSTRACTION



## Binary State







#### **Binary Thinking**

## There are 10 types of people in the world...

Those who understand binary, and those who don't.

#### Numbers

# $\begin{array}{c} 2^{7}2^{6}2^{5}2^{4}2^{3}2^{2}2^{1}2^{0} \\ {}^{128}6432168421 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 = 5 \end{array}$

#### Why this matters...



#### We missed.

28 soldiers died,100 others were injured.

The software fix arrived at Dhahran the next day.

Feb 25, 1991 Dhahran, Saudi Arabia

### Teletype





The popular Teletype Model 33 used 7-bit ASCII code (with an eighth parity bit) instead of Baudot.

## Teletype





#### Pick Something (It's Arbitrary)

#### IBM

EBCDIC Code Table																				
88					0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
B7→					0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
B6					0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
B5→					0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
<b>B4</b>	<b>B</b> 3	B2	BI	HEX-0	0	1	2	3	4	5	6	7	8	9	A	B	С	D	E	F
+	+	ł	ł	HEX-1														3		
0	0	0	0	0	NUL	DLE	DS		SP	*	-		1.0						1 KU	0
0	0	0	1	1	SOH	SBA	sos				1		a	i			A	J		1
0	0	1	0	2	STX	EUA	FS	SYN	155				b	k	s		B	K	S	2
0	0	1	1	3	ETX	IC							c	1	+		C	L	T	3
0	1	0	0	4	PF	RES	BYP	PN					d	m	U		D	M	U	4
0	1	0	1	5	PT	NL	LF	RS						n	v		E	N	V	5
0	1	1	0	6	LC		ETB	UC	10				f	0	w		F	0	W	6
0	1	1	1	7	DEL	IL	ESC	EOT					9	P	×		G	P	X	7
1	0	0	0	8		CAN							h	P	Y		н	Q	Y	8
1	Ó	0	1	9		EM			1				i	r	z		1	R	Z	9
1	0	1	0	A	SMM	cc	SM		¢	1	1	4								
1	0	1	1	В	VT				•	\$	1	#			-					
1	1	0	0	C	FF	DUP		RA	<	×	*	•		in the						
1	1	0	1	D	CR	SF	ENQ	NAK	(	)	-	2								
1	1	1	0	E	so	FM	ACK		+	;	>	=								
11	1	1	1	F	SI	ITB	BEL	SUB	1	-	2						1253			

## Everyone Else

American National Standards Institute (ANSI)

American Standard Code for Information Interchange (ASCII)

(IBM Lost)

# 7-bit encoding standard $2^7 = 128$ characters available

0

ullet

#### ASCII

- Capital Letters 26 •
- Small Letters 26
- o 9 Digits 10
- **Punctuation Marks** 32 •
- Space bar 01 •

- Codes left over 33
- 31 ullet2

Control Codes Special Reserved  $\mathbf{o} = \mathsf{NULL}$ 127 = DEL

#### Teletype

o = No Holes (End of data block) 127 = All Holes (Erase All Data)



# ASCII Digits

011 0001 1
011 0010 2
011 0011 3
011 0100 4

#### ASCII **Upper Case Letters**

- 100 0001 A 100 0010
- 100 0011 100 0100

B  $\mathbf{C}$ D

#### ASCII Lower Case Letters

- 110 0001 a 110 0010 b • 110 0011 C d
- 110 0100

# ASCII Changing Case = Flipping a Bit

100 0001 A
110 0010 a

100 1010 J
110 1010 j

#### ASCII

## Computers use a byte for an ASCII char. The 8-bit position is always O

#### Extended ASCII

- Uses the 8-th position bit for additional characters.
- Started in the 1970's
- Rapidly became proprietary:
  - IBM, Apple, DEC, TRS-80
- Replaced by Unicode

#### UNICODE

- Uses 1-4 bytes
- Started as far back as 1987
- Covers most major languages
- Complex, supports ASCII/Extended ASCII and has a roadmap team.