

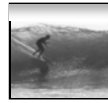


## Announcements

Midterm on Friday: Chapters 1-5,7,8,11  
Bring: pen/pencil, Photo-Id, alert mind

**Tip of the Day:** Studying for the MT is best done w/ book+notes, not online.

1



## Digital Representation

*Everyone knows computers use bits and bytes ... but what are they?*

© Lawrence Snyder, 2004



## Info Representation

Digitization: representing information by any fixed set of symbols

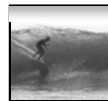


The representation associates one item with each symbol ... encode the telephone keypad using ten colors




What number is:  ?

3



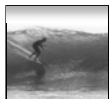
## Creating Symbols

Often, there are many things to digitize, but too few symbols available

- \* The solution is to create more symbols by composing patterns ...
- \* Three patterns make three symbols: 
- \* Pairing them makes 9 symbols; when they are triples, 27 symbols, and ...



4



## An Encoding

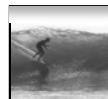
Encode the Latin alphabet

Three pattern  triples = 27 symbols



Digitize -- encode with symbols

5



## Info in the Physical World

Physical world:

- \* The most fundamental representation of information is presence/absence of a phenomenon
  - matter, light, magnetism, flow, charge, ...
- The PandA representation**
- detect: "Is the phenomenon present?"
- set: make phenomenon present or absent

Any controllable phenomenon works: define it right



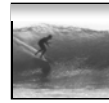
## Info in the Logical World

Logical World:

- \* Information, reasoning, computation are formulated by true/false and logic
  - All men are mortal
  - Aristotle is a man
  - Aristotle is mortal

True and false can be the patterns for encoding information

7



## Connect Physical/Logical

The miracle of IT is that physical and logical worlds can be connected

Present represents true / Absent represents false

-- or maybe vice versa, if everyone agrees--

Pavement Memory



false true false false false true true false true false true false false false

0 1 0 0 0 1 1 0 1 0 1 0 0 0

8



## Bits

Panda is a *binary representation* because it uses 2 patterns

Bit -- it's a contraction for "binary digit"

-- a position in space/time capable of being set and detected in 2 patterns

Sherlock Holmes's *Mystery of Silver Blaze* -- a popular example where "absent" gives information ... the dog didn't bark, that is the phenomenon wasn't detected

9



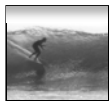
## Bytes

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

0100 0110  
0100 1001  
0101 0100

10



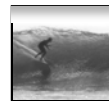
## ASCII

ASCII	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	
0000																																	
0001																																	
0010																																	
0011																																	
0100																																	
0101																																	
0110																																	
0111																																	
1000																																	
1001																																	
1010																																	
1011																																	
1100																																	
1101																																	

0100 0110  
0100 1001  
0101 0100

0100 1000|0111 0101|0111 0011|0110 1011|0110 1001|0110 0101|0111 0011|0010 0001

12



## Encoding Information

Bits and bytes encode the information, but that's not all

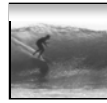
- \* Tags encode format and some structure in word processors
- \* Tags encode format and some structure in HTML
- \* In the *Oxford English Dictionary* tags encode structure and some formatting



## OED Entry For Byte

**byte** (baɪt). *Computers*. [Arbitrary, prob. influenced by bit *sb*<sup>4</sup> and bite *sb*<sup>5</sup>.] A group of eight consecutive bits operated on as a unit in a computer. **1964 Blaauw & Brooks in *IBM Systems Jnl.* III, 122** An 8-bit unit of information is fundamental to most of the formats [of the System/360]. A consecutive group of *n* such units constitutes a field of length *n*. Fixed-length fields of length one, two, four, and eight are termed bytes, halfwords, words, and double words respectively. **1964 *IBM Jnl. Res. & Developm.* VIII, 97/1** When a byte of data appears from an I/O device, the CPU is seized, dumped, used and restored. **1967 P. A. Stark *Digital Computer Programming* xix, 351** The normal operations in fixed point are done on four bytes at a time. **1968 *Dataweek* 24 Jan. 1/1** Tape reading and writing is at from 34,160 to 192,000 bytes per second.

<e><hg><hw>byte</hw> <pr><ph>baɪt</ph></pr></hg>. <la>Computers</la>. <etym>Arbitrary, prob. influenced by <xt><x>bit</x></xr> <ps>n.<hm>4</hm></ps>and <xt><x>bite</x></ps>n.</ps> </xt></etym> <s4>A group of eight consecutive bits operated on as a unit in a computer.</s4> <q><q><q>1964</q><as>Blaauw</a> & <a>Brooks</a> <bib>in</bib><w>IBM Systems Jnl.</w> <lc>III, 122</lc> <qt>An 8-bit unit of information is fundamental to most of the formats <ed>of the System/360</ed>. &es. A consecutive group of <i>n</i> such units constitutes a field of length <i>n</i>. &es. Fixed-length fields of length one, two, four, and eight are termed bytes, halfwords, words, and double words respectively.</qt></q><q><q>1964</q><w>IBM Jnl. Res. & Developm.</w> <lc>VIII, 97/1</lc> <qt>When a byte of data appears from an I/O device, the CPU is seized, dumped, used and restored.</qt></q><q><q>1967</q><as>P. A. Stark</a> <w>Digital Computer Programming</w> <lc>xix, 351</lc> <qt>The normal operations in fixed point are done on four bytes at a time.</qt></q><q><q>1968</q><w>Dataweek</w> <lc>24 Jan. 1/1</lc> <qt>Tape reading and writing is at from 34,160 to 192,000 bytes per second.</qt></q></ps></e>



## Summary

IT joins physical & logical domains so physical devices do our logical work

- \* Symbols represent things 1-to-1
- \* Create symbols by grouping patterns
- \* PandA representation is fundamental
- \* Bit, a place where 2 patterns set/detect
- \* ASCII is a byte encoding of Latin  $\alpha$ bet
- \* In addition to content, encode structure

14