Announcements

Midterm on Monday: Chapters 1-5,7,8
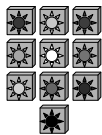
Digital Representation

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

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

Creating Symbols

Usually there are not enough symbols
* The solution is to create more symbols by composing patterns …
* Three patterns alone could be symbols, but grouping them together makes 9 when they are pairs, 27 as triples, and …
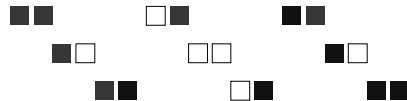
An Encoding

Encode the Latin alphabet

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

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

FIT100

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

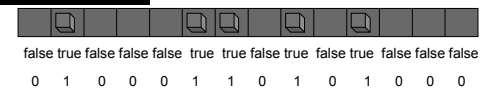True and false can be the patterns for encoding information

---

Connect Physical/Logical

FIT100

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

Present represents true / Absent represents false

Pavement Memory

or maybe vice versa

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

Bits

FIT100

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

Mystery of Silver Blaze -- popular example where "absent" gives information

---

Bytes

FIT100

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

Demonstration

FIT100

I need 8 volunteers, 4 women & 4 men

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Encoding Information

FIT100

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.*[4] and bite *sb.*] A group of eight consecutive bits operated on as a unit in a computer. **1964 Blaauw & Brooks** in *IBM Systems Jrnl.* 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 Jrnl. 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 <xr><x>bit</x></xr> <ps>n.<hm>4</hm></ps>and <xr><x>bite</x> <ps>n.</ps> </xr></etym> <s4>A group of eight consecutive bits operated on as a unit in a computer.</s4> <qp><q><qd>1964</qd><a>Blaauw &amp. <a>Brooks <bib>in</bib> <w>IBM Systems Jrnl.</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><qd>1964</qd> <w>IBM Jrnl. Res. &amp. 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><qd>1967</qd> <a>P. A. Stark <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><qd>1968</qd> <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></qp></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 αbet
* In addition to content, encode structure