How we represent bits, numbers, letters?

**Communicating in the Blink of an Eye**

*Lawrence Snyder*

*University of Washington, Seattle*
Announcements

- How’s it going?
- After Image Survey ... today by 5:00
- The threads of this class

Abilities
Lightbot 2.0
  functional abstraction
  lab 1: symbolic lightbot
  lab 2: functions with lightbot
  Processing
  lab 4: practice w/Processing

Facts
0s and 1s Glories
  lab3: binary & hex -- numbers
  today lecture -- letters

Computing World
Social Network
We have written two kinds of Processing programs –

- static, which only draw a picture
- dynamic, which keep drawing a picture

```java
size(100, 100);
background(255);
fill(0, 0, 255);
ellipse(50, 50, 30, 30);

void setup() {
    size(100, 100);
    background(255);
}
void draw() {
    ellipse(50, 50, 30, 30);
}
void mousePressed() {
    fill(0, 0, 255);
}
```

What's The Difference?
A Curious Story...

The Diving Bell and the Butterfly
Jean-Dominique Bauby
Asking Yes/No Questions

- A protocol for Yes/No questions
  - One blink == Yes
  - Two blinks == No
- PandA implies that this is not the fewest number of blinks ... really?
Asking Letters

In English ETAOINSHRDLU...
How many questions to encode:

Plus ça change, plus c'est la même chose?

Asking in Frequency Order:

ESARINTULOMDPCFBVHGJQZYXKW

9 12
How many questions to encode:

*Plus ça change, plus c'est la même chose?*

- Asking in Frequency Order:
  ESARINTULOMDPCFBVHGJQZYXKW
- Asking in Alphabetical Order:
  ABCDEFGHIJKLMNOPQRSTUVWXYZ

12 16
How many questions to encode: 
*Plus ça change, plus c'est la même chose?*

- Asking in Frequency Order: 247
  ESARINTULOMDPCFBVHGVJQZYXKW
- Asking in Alphabetical Order: 324
  ABCDEFGHIJKLMNOPQRSTUVWXYZ
An Algorithm

- Spelling by going through the letters is an algorithm
- Going through the letters in frequency order is a program (also, an algorithm but with the order specified to a particular case, i.e. FR)
- The nurses didn’t look this up in a book ... they invented it to make their work easier; they were thinking computationally, though they probably didn’t know it
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
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 | 0000 | 0001 | 0010 | 0011 | 0100 | 0101 | 0110 | 0111 | 1000 | 1001 | 1010 | 1011 | 1100 | 1101 | 1110 | 1111 |
|-------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
|       | N_  | S_  | S_  | E_  | E_  | E_  | A_  | B_  | B_  | H_  | L_  | Y_  | F_  | F_  | C_  | S_  |
|       | D_  | D_  | D_  | D_  | Y_  | E_  | C_  | E_  | S_  | E_  | M_  | E_  | S_  | E_  | G_  | R_  |
| !"#$%&'()'*+,-./| 0 1 2 3 4 5 6 7 8 9 ;| < = > ?| `abcdefg| pqrstuvwxyz| { } ~| ^ _| αβγδεζηθικλμνξ| αβγδεζηθικλμνξ| корет| αβγδεζηθικλμνξ|
| 0100 0011| 0101 0011| 0101 0000| 0100 1000| 0111 0111| 1010 1010| 1100 1100| 1111 1111| 0100 1000| 0111 0111| 0110 1110| 0101 0101| 0110 0101| 0111 0011| 0010 0001|
UTF-8

Uniform Transformation Format for bytes (UTF-8) is universal ... all characters have a place: 1,2,3,4 B
UTF-8

Uniform Transformation Format for bytes (UTF-8) is universal ... all characters have a place: 1,2,3,4 B

can you read this?
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
- Tags are one form of meta-data: *meta-data* is information about information
**OED Entry For Byte -- Metadata**

**byte** (balt). *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 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.
This week we have seen ...

- Bits encode numbers using the binary representation 11 1110 0111
- Bits encode letters using ASCII for North American and Western European languages

This suggests an principle we will soon argue:
- All information can be represented with bits
Computers join physical & logical domains so physical devices do our logical work
- Symbols represent things 1-to-1: 0, 1
- Create symbols by grouping patterns: 0101 0111
- PandA representation is fundamental: present?
- Bit, a place where 2 patterns set/detect
- ASCII is a byte encoding of Latin αbet
- In addition to content, encode structure: meta