

---

# Data Representation

Interpreting bits to give them meaning

## Part 3: Media - Text and Pictures

---

Notes for CSC 100 - The Beauty and Joy of Computing  
The University of North Carolina at Greensboro

---

---

---

---

---

---

---

---

---

## Reminders

---

Blown to Bits: Chapter 4 discussion over the next week

Homework 3:

- Questions?
- African fractals lessons are ready
- Goal: At least watch the video by Friday

For Friday:

- Do Pre-Lab work for Lab 10
  - Project goal: Have an informal idea and perhaps a team by Friday
- 

---

---

---

---

---

---

---

---

## Data is more than just numbers!

---

Data is stored using bits but represents many things:

- Documents
- Pictures
- Sound/music
- Video
- ...

How does this work?

- File formats: Structure bits in such a way that mapping between bits and what they represent is unambiguous
    - Standardized or open file formats
      - Specified so that anyone can write programs for them (JPEG, MPEG (and MP3), OpenDocument, HTML, ...)
      - "Open" and "standardized" doesn't mean "free" (MP3, GIF, ...)
  - A data capture or creation program builds the file in the appropriate format
  - A rendering program converts the file format to a recognizable form (image viewer, web browser, video player, ...)
- 

---

---

---

---

---

---

---

---



# Formatted Text

HTML

ASCII provides letters - what about fonts, sizes, etc?

One option: HTML - HyperText Markup Language

- The "language of web pages"
- "Markup" indicates formatting/style
- All characters are just regular character set (like ASCII) - including markup
- Must be *rendered* to convert character-based markup to formatted text
- A lot of formatting is now in CSS - Cascading Style Sheets
- Much more involved than these examples!

### HTML Source

```
This is formatted text, which can be <b>bold</b> or <i>italic</i> or <u>underlined</u> or <span style="font-size: 150%">big</span> or <span style="font-size: 50%">small</span> or ...
```

### Rendered Text

This is formatted text, which can be **bold** or *italic* or underlined or **big** or or ...

---

---

---

---

---

---

---

---

---

---

# Pictures

Grayscale



Grayscale images have levels of intensity, but no color

- More information than bi-tonal black and white (like fax machines or most printers)
- Less information than color

---

---

---

---

---

---

---

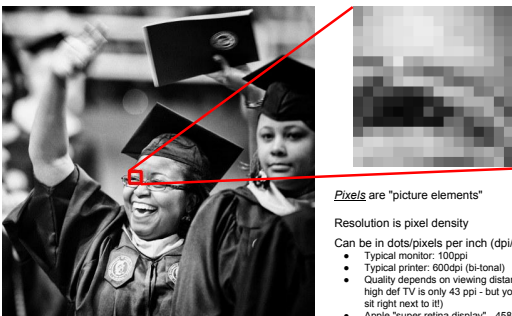
---

---

---

# Pictures

Grayscale - Pixels



*Pixels* are "picture elements"

Resolution is pixel density

Can be in dots/pixels per inch (dpi/ppi)

- Typical monitor: 100ppi
- Typical printer: 600dpi (bi-tonal)
- Quality depends on viewing distance (52" high def TV is only 43 ppi - but you don't sit right next to it)
- Apple "super retina display" - 458 ppi

---

---

---

---

---

---

---

---

---

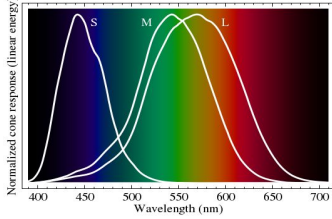
---



## Pictures

*Why does this work?*

Bottom line: If humans can only perceive three colors (red, green, and blue) then reconstructing just those three colors allow us to perceive everything just as in an original.



Interesting question: What if someone were born with a mutation that gave them purple and yellow receptors?

---

---

---

---

---

---

---

---

## Summary of Part 3

Files just store bits

- Bits are bits: no different for text or images or ...
- Rendering program makes all the difference
- Text - encodings defined in standards
  - ASCII, Unicode, HTML
- Image formats take advantage of biology
  - Images aren't "accurate" but we perceive them that way

---

---

---

---

---

---

---

---