COMPSCI 145

Representing, Storing, and Retrieving Information

LECTURE #1

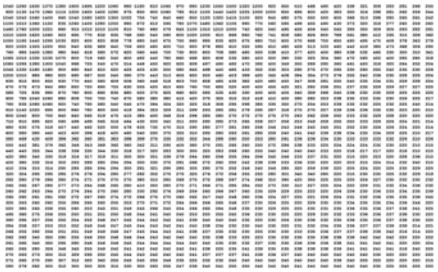
Professor William T. Verts

Data Representation

- What does "Representation" mean?
- What is a Representation?
- Why do we care?

Copyright (C) 2018 -- Dr. William T. Verts

Here's an Example:



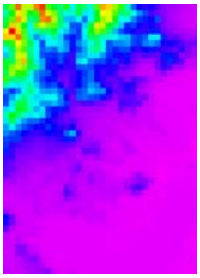
Copyright (C) 2018 -- Dr. William T. Verts

What did you get from the numbers?

- Not much, probably.
- We're not very good at handling these kinds of data ourselves.
- A computer could easily compute:
  - Maxima,
  - Minima,
  - Averages,
  - Standard Deviation,
  - Etc.
- …but that still doesn't mean much to us.

Copyright (C) 2018 -- Dr. William T. Verts

How about this view of the data?



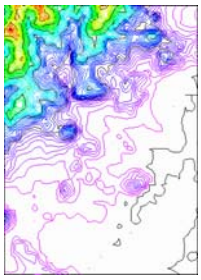
Copyright (C) 2018 -- Dr. William T. Verts

What did you get from the image?

- By color coding values spectrally (Red-Orange-Yellow-Green-Blue-Violet, where Red=High and Violet=Low),
- We get a real sense of where the high numbers are located versus where the low numbers are located.
- We could have done this by computing with the numbers directly, but only with a lot of effort.
- The *change in representation* allows us to exploit our visual systems to understand more of what the numbers tell us.

Copyright (C) 2018 -- Dr. William T. Verts

How about this view of the data?



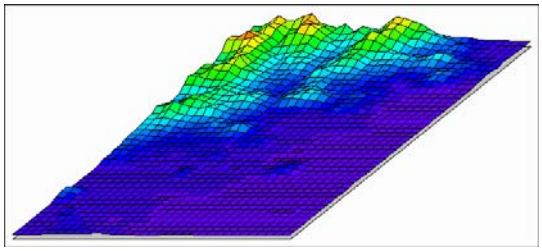
Copyright (C) 2018 -- Dr. William T. Verts

What did you get from the image?

- By contouring the data, we get a set of isolines (lines of the same numeric value),
- Combining the isolines with the color spectrum, we also get the same visual information as from the color-only version.
- This is much harder to do manually.
- Again, the *change in representation* allows us to see things that are not immediately obvious.

Copyright (C) 2018 -- Dr. William T. Verts

How about this view of the data?



Copyright (C) 2018 -- Dr. William T. Verts

What did you get from the image?

- By interpreting the numbers as altitudes, we get another view of the same numbers.
- High numeric values correlate with high altitude values, giving us still another view of the same information.
- This is very hard to do manually.
- And yet again, the *change in representation* allows us to see things that are not immediately obvious.

Copyright (C) 2018 -- Dr. William T. Verts

Here's someone else's Representation

Copyright (C) 2018 -- Dr. William T. Verts

And yet Another (My Photograph)

Copyright (C) 2018 -- Dr. William T. Verts

What do we get from all this?

- The representation we choose determines and constrains:
  - How we think about solving problems,
  - The tools we use to solve those problems,
  - The types of solutions we can obtain with those tools,
  - The way we understand the solutions to those problems.

Copyright (C) 2018 -- Dr. William T. Verts

Keep this in mind

How does a particular representation facilitate computation, the storage of information, or the retrieval of information?

Copyright (C) 2018 -- Dr. William T. Verts

Dr. Bill's Three Laws of Representation

- ***1. Representation facilitates Computation.***
- ***2. All Representations are Compromises.***
- ***3. Compromises represent Human choices.***

Engrave these three laws on your heart. Know them, love them, live them.

Copyright (C) 2018 -- Dr. William T. Verts

Why would we choose one Representation over another? (1/4)

- Speed
- Memory
- Complexity
- Implementation Time
- Cost

Copyright (C) 2018 -- Dr. William T. Verts

Why would we choose one Representation over another? (2/4)

- Technology
- Robustness
- Durability
- Failure Modes
- Ease of Maintenance

Copyright (C) 2018 -- Dr. William T. Verts

Why would we choose one Representation over another? (3/4)

- Extensibility
- Scalability
- Usability
- Visualization
- Efficiency

Copyright (C) 2018 -- Dr. William T. Verts

Why would we choose one Representation over another? (4/4)

- Hazards
- Pedagogy
- Legal
- Available Knowledge
- Etc., etc., etc.

Copyright (C) 2018 -- Dr. William T. Verts

My Goal for this Course

I want you to be able to look at any artifact, be it a computational aspect of a program you use or an object in the "real world", and to be able to ask yourself:

- "What were the compromises that went in to making that artifact the way it is?"
- "Could it have been done 'better'?" (for some definition of better)

Copyright (C) 2018 -- Dr. William T. Verts

Here's an Example: Why are Fire Hydrant Bolts 5-Sided?

Solving the Riddle

- Question: What do normal bolts/nuts look like?
  - 4-sided
  - 6-sided
- Question: Have any of you ever seen a 5-sided wrench?
  - Probably not.
- Question: Who might have such a wrench?
  - Fire-fighters, obviously.
- Question: How do fire-fighters put a critical resource in an exposed public place without the public taking advantage of it?
  - Make it so most members of the public can't easily open it. (But I do have a machine shop…)

Copyright (C) 2018 -- Dr. William T. Verts

Specific Topics in this Course

- Pre-Computer methods of Representing Information
- Analog versus Digital, and why it matters
- Representing Numbers – a month wallowing in binary
- Curves and Surfaces, with Interpolation
- Compression – how it works
- Encryption and Steganography – Hiding Information
- Information Theory – Error Detection and Correction

Copyright (C) 2018 -- Dr. William T. Verts

A Note about Data

- Most people misuse the word "Data", particularly computer science people.
- The phrase "This data shows…" is always wrong!
- Data are plural! Datum is singular!
- You properly say "These data show…" or "This datum shows…"
- [You are, however, allowed to say "This data set shows…"]
- Learn to use it correctly!

Copyright (C) 2019 -- Dr. William T. Verts