The Next Generation of Scientists in the Workforce

Bob Zigon

Purdue SECANT Workshop

November 16, 2007
Questions that need answered

- What do scientists need to know about Computer Science?
- What skill set is expected?
- What computational thinking abilities are required?
What do I do?

- Software technical lead for Flow Cytometry tools
- We study & learn the vocabulary of “flow”
- We study & learn how to control the instrument
- We study & learn to build the instrument
- We put on lab coats and act like flow scientists
- We go into the field and observe their processes
So you’re thinking “What is Flow Cytometry”? 

- It’s a technique for counting and examining characteristics of cells that are suspended in a stream of fluid.
- Used by researchers and clinicians.
- It’s a popular tool used to study blood cancer.
How does flow cytometry work?

- The cells flow down a thin glass tube.
- The cut across a laser beam.
- They fluoresce.
- We capture the fluorescent energy and convert it into a digital value that measures a characteristic of the cell.
What do we supply the scientists with?

- Data Management

  Acquisition, Organization, Querying, Visualization

  8 years of clinical data is 3 tb
  8 years of pharmaceutical data is 185 gb
  8 years of research data is 15 gb

  data is stored in OLTP and OLAP formats

  7 to 10 second query times
What else?

• Data Visualization

  Real time (20 fps) interaction with 750mb data sets

  Build tools for N-Dimensional manipulation and review

  Create data reduction schemes

  Create & implement clustering algorithms
What else?

• Integrate cellular, proteomic and genomic data
• Build usable user interfaces (HCI)
• Study algorithmic complexity with 15M cells
• Electronic signatures and traceability of data
And finally …

- Parallelize algorithms in cost effective ways

- Build wizards for longitudinal and cross sectional studies
Where are the useful nuggets of skill that we need to be teaching?
Questions

- Bob Zigon, robert.zigon@beckman.com