OVERVIEW OF RESEARCH

Dr. Garrett Dancik

Steps to conducting research

1. Read relevant literature, to get necessary background and to learn what else has been done
2. Identify a concrete investigative question
3. Identify an appropriate methodology for answering the question
4. Carry out the actual research project
5. Disseminate results through oral and written communication

Example #1:

1. Question: Does MySQL or MongoDB enable more efficient retrieval of genomic data (or football data, or social media data, …)?
2. Methodology: Load the same data into both databases, use appropriate indexing, and measure performance using relevant queries.

Example #2:

1. Question: Is GD-sort better than quicksort? (Note: better would have to be defined; e.g., better = faster on random data)
2. Methodology:
   - Theoretical running time and memory, big O notation
   - Comparison of GD-sort and quicksort using real data

How to read scientific research

- See **Tips for Reading Journal** articles, on page 3 of https://web.archive.org/web/20230129170522/www.colby.edu/biology/bi319/GuideReadJour.doc

- Example:

  - **Empirical Performance of Internal Sorting Algorithm**
  - Silas, F., Musa, Y., & Joyce, S. A. (2017). *British Journal of Mathematics and Computer Science*, *20*(1), 1-9.

    - https://www.researchgate.net/publication/312168023_Empirical_Performance_of_Internal_Sorting_Algorithm

  - What does the abstract tell us?
  - What does Figure 1 tell us?
  - What does Figure 2 tell us?
  - What does Figure 4 tell us?

Computer Science is a broad discipline

- Association for Computing Machinery (ACM) knowledge areas

• Artificial Intelligence	Distributed Computing
• Algorithms	Data Management
• Architecture and Organization	Software Engineering
• Graphics and Visualization	Software Development
• Human Computer Interaction	Security
• Programming Languages	Ethics
• Networking and Communication	
• Operating Systems	

For full list see: https://csed.acm.org/wp-content/uploads/2024/04/3.1-Body-of-Knowledge-1.pdf

Choosing a Research Topic

1. What aspects of computer science interest you?
2. Explore what others have done by reading the literature
   - Media and review articles can be a good place to start, but you will need to find *research* articles
3. Trending topics
   - Chat-GPT
   - Facial recognition
   - AI and ethics
   - Self-driving cars (image recognition)
   - Virtual reality
   - Quantum computing
   - Targeted advertising
   - Blockchain / cryptocurrency

Computer Science

- Computer Science is the systematic study of the feasibility, structure, expression, and mechanization of the methodical processes (or algorithms) that underlie the acquisition, representation, processing, storage, communication of, and access to information…

- Computer Science is not *just* about building computers or writing computer programs!...Computer science is not about the tools we use to carry out computation. It is about how we use such tools, and what we find out when we do.

- Source: http://www.cs.bu.edu/AboutCS/WhatIsCS.pdf (no longer active)

Computer science has applications in:

Computer Engineering
Information Technology and Information Systems   Computational Statistics
Bioinformatics and Computational Biology   Mathematical Modeling

What is a computer (part 1)?

- A computer is a *platform* that implements varying algorithms and methodologies for storing, retrieving, and analyzing information
  - Ex: Quicksort is a sorting method that exists and can be analyzed independently of any computer.
  - Ex: The properties of a relational database do not depend on the computer used to store the data
- There is a theoretical component to computer science, but these theories can be tested in practice
- Project ideas:
  - Develop a new method and compare with existing methods (difficult)
  - Compare two methods
  - Evaluate how altering a method changes the behavior / performance of the method

What is a computer (part 2)?

- A computer is a *tool* that can be used to
  - Analyze (large amounts of) data
    - Example: analyzing IMDB data: https://www.r-bloggers.com/imdb-movie-analysis/
    - Example: Can Twitter be used to predict crimes? (Answer – yes, to some extent: http://www.citylab.com/tech/2014/03/how-twitter-could-help-police-departments-predict-crime/8651/)
  - Simulate physical or biological system and analyze it
    - Cellular automata models
      - http://mathworld.wolfram.com/CellularAutomaton.html
      - A cellular automata model of traffic flow
        - http://www.ajuronline.org/uploads/Volume%2012/Issue_1/AJURVol12Issue1Aug2014pp39to48.pdf
    - Agent-based models
      - Agents move and interact according to rules executed at discrete time steps
      - In some cases this approach has identified emergent behavior that may be counter-intuitive. Ex: should we put a pillar by the front door of the classroom?
          Ex: "The Game": http://web.archive.org/web/20230208205919/http://www.icosystem.com/labsdemos/the-game/
      - Write a program to download and analyze a large dataset, such as tweets, to answer a specific question
  - Develop a simulation (or modify an existing one) to answer a specific question

Where to find published research

- Library Databases: http://easternct.libguides.com/az.php?s=126319
  (if off-campus, log in with your Eastern userID and password)
  - ABI Inform Complete: The database features thousands of full-text journals, dissertations, working papers, etc. Includes IEEE journals.
  - ACM digital library: http://dl.acm.org
  - ACM surveys (CSUR) http://dl.acm.org/citation.cfm?id=J204
    - Provides comprehensive tutorials and survey papers
  - Academic Search Premier: Multidisciplinary database with full text articles

- Undergraduate research:
  - American Journal of Undergrad Research:
    - http://www.ajuronline.org
  - Student Pulse
    - http://www.studentpulse.com/topics/15/computer-science

- Google Scholar (http://scholar.google.com)