CSci 8002: Intro. to Research in Computer Science - II

Ravi Janardan
janardan@cs.umn.edu
www.cs.umn.edu/~janardan
6-217 Keller Hall
Outline

• Group work
• Formulating research questions
• Wrap-up
What is Research?

“Research is what I’m doing when I don’t know what I’m doing”

– Werner von Braun

“If we knew what we were doing, it wouldn’t be research”

– Albert Einstein

1/26/18
Group Work

Discuss the following [10 min]:

1. What do you understand by the term “research”?

2. Give 1-2 examples of impactful CS&E research that you are aware of

3. Suggest 1-2 research directions that you think would be interesting to pursue (other than settling “P = NP?” 😊)

Remember: There are really no right or wrong answers here
Outcome of group discussion
“Research” (contd)

• **Working definition:** *Research is the process of advancing the frontiers human knowledge through the discovery/invention/application of new concepts, ideas, and techniques—largely for the benefit of society.*

• Basic research vs. Applied research (Basic -> Applied)
  – DNA structure -> genome sequencing, biomedicine
  – Electricity and magnetism -> modern power grid
  – Number theory -> cryptography
  – Finite automata -> compiler technology

• Development (The “D” in “R&D”)

• Incremental vs. Transformative research?
  [https://z.umn.edu/trans-res](https://z.umn.edu/trans-res)
So, how do I start doing research?
Finding research problems

Area (Interest) → Topic → Questions/Hypotheses → Problems

Finding research problems

Area (Interest) → Topic → Questions/Hypotheses → Problems

Theory → Geometric algorithms → Why are high-dimensional problems so hard? → k-Nearest Nbor, Point-location
Finding research problems

Area (Interest) → Topic → Questions/Hypotheses → Problems

Theory → Geometric algorithms → Why are high-dimensional problems so hard? → k-Nearest Nbor, Point-location

HCI → Collaborative systems → Info. in well-known online media is skewed/unreliable (Hyp.) → Gender imbalance in Wikipedia, reliability of e-health info.
Finding research problems

Area (Interest) → Topic → Questions/Hypotheses → Problems

Theory → Geometric algorithms → Why are high-dimensional problems so hard? → k-Nearest Nbor, Point-location

HCI → Collaborative systems → Info. in well-known online media is skewed/unreliable (Hyp.) → Gender imbalance in Wikipedia, reliability of e-health info.

CS + Bio → Proteomics → What is the relationship b/w protein structure and function? → Protein structure similarity and alignment, docking
Some factors in area/topic selection

• Must be compelling and excite you!
• Pass the “Who cares?” test
• Be fertile area (growth potential of at least 5-10 years) with prospect of making impact
• Have moderate learning curve; leverage existing strengths; bring unique skills to the table
• Have collaborator/domain expert (esp. for cross-disciplinary projects)

Must be able to tell a good story
Strategies for question/hypotheses formulation and problem identification

• Read, Read, Read!
  
  – *Recent* conf./workshop papers
  – Journals
  – Books (esp. Handbook of “X”)
  – “Lurk” in reputable online forums/blogs
    (later on: contribute, too)

Read critically (and skeptically), with a question in mind. Identify trends, debates, points of disagreement, what is known, and, most importantly, what is *not* known
Strategies (contd.)

• Write, Write, Write!
  – Take notes of what you read; summarize each paper
  – Distill and synthesize info. from multiple sources into a lit. review.
  – Outlines of questions should start emerging at this point if lit. review has been done properly
  – Maintain a personal web page with links to journals, confs. (+deadlines), key researchers in field, important papers, blogs, your thoughts and write-ups, etc. (“Onestop”)

• Talk to peers, advisor, other faculty, visitors, colleagues at confs., etc.

• Learn by doing. Implement an algorithm or system you read about, try to break it, identify its strengths and weaknesses, improve upon it, etc.
Formulating research questions

- Characteristics of good research questions
  - Must excite you (and keep you awake at night 😊)
  - Must be interesting and compelling to your larger research community
  - Must have long-term research potential (not “one-off”)
  - Answers must yield a significant new contribution to the field and spur further research
  - Not something incremental or a “repackaging” of existing ideas/techniques
  - Other?
Types of research questions

- Discipline/area-dependent. Could have to do with
  - Solving a major open problem (new algorithm, new theorem, disproving a conjecture, ...)
  - Experimental study to establish/verify/refute a hypothesis (or published claim)
  - Building a software system to demonstrate a proof of concept, analyse large datasets, etc. (but maybe not s/w for the sake of s/w?)
  - Other?
Research questions (contd.)

• In general,
  Good research questions => High-quality, impactful research => High-quality publications (and other goodies)

• Helps to know what reputable conferences and journals are looking for and publishing...
“Practical” considerations

• Choose questions where the learning curve is moderate, where you can leverage existing skills, and where you can bring unique skills to the table
  – E.g., linear algebra -> data mining
    graph theory -> networks (social, biological, communication, …)

• Start small and scale up

• Seek collaborators/domain experts (esp. for cross-disciplinary work). Good resource for interesting problems, reality check, real-world data, validation of work, etc.
Guidelines for formulating research questions

• 2-3 research questions based on your readings for lit. review
  – Must be substantial and meaningful
  – Must have downstream research potential
  – Must be (largely) of your own creation (rather than taken directly from a paper)

• For each question, write a paragraph stating
  – What the question is
  – Why it is important
  – What it might mean to answer/refute the question
  – Prior related work (can cross-ref. lit. review)
  – Outline any plausible initial approaches

• Provide a concluding paragraph on how questions relate to each other (if appropriate)

• If needed, talk to me or (better still) to your advisor

• Read and follow detailed guidelines posted. Questions are due Feb. 23
A great book!

• Booth et al. *The Craft of Research*

  – 3rd edition freely available online from U library ([www.lib.umn.edu](http://www.lib.umn.edu))

  – 4th edition available at various retailers (paperback < $20)
Wrap-up

• Questions/discussion?
• Next class
  – Guest lecture (Prof. Sattar).
  – Paper posted on class web page
  – Guidelines for lecture summary posted
  – Read and come prepared to participate
• 8001 scores
  – Posted on 8002 Moodle page (accessible from class web page)