INFO 4990: Information Technology Research Methods

Overview of Research: Definition and Process
Lecture by A. Fekete,
based in part on materials by J. Davis and others

Research

- What is research?
- Key components
  - A question of interest
  - A claim
  - Evidence
  - Argument (link evidence to claim)
- What are some of the different kinds of research?
- How to do research

Definition

- As a general term, research is gathering information that answers a question and so solves a problem.
  - from Booth, Colomb, & Williams "The Craft of Research"
- Eg journalist finds out who contributed to election campaign fund, to understand political decisions
- Eg repairer finds out what procedure to use to fix a washing machine

Academic Research

- In academic research, you must not only answer a question, but you must find something *new* and *interesting*.
- You join a community of researchers.
  - You must *advance* the collective understanding of this community.
- Each community has a cumulative tradition with a set of interesting questions, tools and methods, practices, a style and language for writing up the research.
  - Research is a conversation and ongoing social activity!
- You need critical and careful reading of published research
  - to learn what the community already knows
  - to fit your work into the community
  - to be prepared for your own work to be evaluated

A question

- Every piece of research should address a question of interest to the community
- Each community has a tradition of style of question,
  - what happens?
  - why does it happen?
  - how should one do something?
  - what something should one do?
- Many questions fit into an on-going agenda
  - eg find data models to represent different sorts of information
  - eg move data and computation in a network to exploit locality

A claim

- Every piece of research makes a claim (the "contribution")
  - this should answer a question of interest
- Claims can be very diverse, among fields and within fields
- This is what happens
  - eg how often is data corrupted when using weak concurrency control
- This is why something happens
  - eg what factors lead to project success in open-source development
- This is a better way to do something
  - eg efficiently recalculate a graph layout after a change to topology
- This is a better something to do
  - eg allow users to see the model of their skills kept in a teaching system

Be explicit about the meaning of "better"

Evidence

- You must back up the claim
  - evidence can be very varied, for examples
    - a prototype implementation to show that a system can be built to achieve claimed functionality
    - a simulation model which is executed to show a system has certain properties
    - measurements of a running system to show it has good performance
    - observations of behaviour in an organisation to show what is happening
    - a mathematical proof to show that some process has desired properties
- Each research method is defined by the sort of evidence that it can produce
  - each community has its own standards of quality and reasonableness

Argument

- You should show that the evidence you offer supports the claim you made
  - It's essential that you deal with natural or obvious objections to the correctness or importance of the work
  - that is, you must think like your readers, and anticipate their reactions
- In systems work, this is often called an "evaluation" of the design

Some types of claim and argument I

- This system design leads to better performance on some metric
  - make sure you limit how much worse this makes other metrics (such as cost!)
  - make sure your measurements are fair (don't compare with "strawman" design but with state-of-the-art)
- This system design offers better functionality for some uses
  - make sure you show it can be implemented with adequate performance

Some types of claim and argument II

- This behaviour can be explained by this theory
  - make sure you don't have confounding factors such as level of experience, or method novelty, or subject expectations ("placebo effect")
- This is what happens
  - make sure you don't interfere too much with what happens when you gather data, or misinterperete it due to observer expectations

Common mistakes I

- Gather lots of data without a focussed question or method
- A collection of facts is not a contribution!
  - it must reveal some pattern or understanding that you make explicit

Common mistakes II

- Build a system without a focused question or planned evaluation
  - eg let's see how to use aspect-oriented programming in a sensor network
- An innovative system is not a contribution!
  - it must be a worthwhile innovation in a sense you make explicit
    - eg better performance
    - eg better functionality

Negative results

- Sometimes, you don't get the result you hoped for
  - You gather data that does not reveal any pattern or understanding
    - eg no factor seems to correlate well with project success
  - You design a system that turns out to be worse than the state-of-the-art
    - eg your machine learning algorithm runs slower than expected
- You can still salvage a thesis
  - but publication will be hard, and it won't have much impact
- Try to find some way to contribute to our understanding, or suggest fruitful directions for further work
  - eg what features of the algorithm make it slow
- Make sure the problem is intrinsic, not just your bad coding/experiment design/etc

Ground-breaking work

- Very rarely, a piece of research will establish a whole new agenda for a field, or even a new field
  - the contribution can be as much in the possibilities for further work, as in the result itself!
- In some sense, this is work that asks a new type of question, or introduces a new method
- I don't recommend this for Hons/MIT/MSc/PhD
  - save the idea till you have time enough, and flexibility enough to deal with inevitable digressions/difficulties

Great scholars do not solve problems; they create them.

-Albert Einstein

The University of Sydney

Idealised Research Process v1

- Finding the right "Question" to seek an answer for
- Making flexible plans
  - Choosing appropriate research method(s)
- Gathering the data, doing the experiment, building the prototype etc. ("Evidence")
- Analysis, interpretation, and conclusions ("Claim")
- Writing up the report ("Argument").
  - Importance of "writing" (aided by thinking from the point of view of your readers)

Idealised Research Process v2

- Finding the right "Question" to seek an answer for
- Having an idea of an possible answer ("Claim")
- Making flexible plans
  - Choosing appropriate research method(s)
- Gathering the data, doing the experiment, building the prototype etc. ("Evidence")
- Analysis, interpretation, and conclusions ("Argument")
- Writing up the report.

Actual Research Process

- Research is a non-linear process!
  - it is normal in v1 for argument to lead to changes in the claim
  - it is normal in v2 for the process of gathering evidence to lead to changes to the claim
    - sometimes one refines the claim
      - eg limit the scope (from "this has higher throughput" to "this has higher throughput if the contention rate is low")
    - sometimes one must change the claim entirely
  - sometimes while gathering evidence, one finds new questions which look worth answering!
- New claims or questions need further evidence, revised plans, maybe even different methods

Finding a question

- Especially when you are learning to do research, it may be already chosen for you by supervisor
  - or supervisor may suggest an area, and leave you to find the question
- A question may arise naturally from some previous work
  - incremental "delta" research
- A question may come from the interaction of previous works
  - eg reconcile apparent contradictions
- A question may arise due to new technology
  - eg how to use new hardware
  - eg revisit design choices as speed, bandwidth, cost etc change
- A question may come from a great mind's simple curiosity

Suitable Research Questions

Main Considerations:

-Specificity and answerability– can the questions be answered through research?

- Scale and Scope in relation to needs and available resources.
- Resource Adequacy in relation to available time.

- Often start with broad topic space, then narrow in to a specific question

Research Plan Structure

- State the question
- Evidence gathering methods
  - Type of method to be used.
  - Type of data to be obtained.
  - Pilot Study
- Data analysis or evaluation methods
- Budget and timetable
  - make sure you have internal deadlines, and fallback options!
  - leave time for writing up the results

Robust planning

- Risk-driven planning: Don't make a plan where if something goes wrong, you end up with failure
  - Eg what if the supplier of a crucial device goes broke before delivering, or if the software you plan to use doesn't work on the hardware that's available, or if a data source can't be obtained
  - check key dependencies early enough
  - Have a fallback position!
- Make sure that there are preliminary goals (perhaps small, unsurprising) that are safe (don't require luck/genius/etc)
  - plan to get these accomplished early
  - Then enhance/expand/complicate, to make it more interesting

Defining the Research Problem

- You need several clear, concise and succinct statements of the research problem of different lengths
  - eg one minute (elevator) pitch
  - eg ten minutes introduction to full seminar
- Issues you must deal with:
  - Can it be understood by others without too much background?
  - Does it demonstrate a good understanding of the research community?