What the class is about:
- how to do research
- project helps you demonstrate that you’ve learned the methods

Is there consensus about proper research methods in computer science?

→ No, but we’re getting there.
I will try to teach what I believe are the emerging standards.
1. Establish relationships or other facts
2. Organize these relationships into a body of knowledge

**Relationship between X and Y**

- **X**: design of cache
- **Y**: performance
- **X**: Java experience
- **Y**: errors in C code
- **X**: test suite
- **Y**: ability of generation method
- **X**: size of button
- **Y**: errors and speed of clicking
X: new network protocol
Y: latency and bandwidth
Two Ways to Establish Relationships

(1) Induction
   (1) Observe a specific situation proposed theory
   (2) Draw conclusions that emerge out of your data.
       → possible to induce something false from true statements.

(2) Deduction
   (1) start with a general theory.
   (2) deduce what the theory says about a specific situation → hypothesis
   (3) design a study to build
evidence for your hypothesis.

- correlational observation
- experimentation
Induction Techniques

How can you observe a situation?

- ethnography
- contextual inquiry
- measure quantitative data about a situation

Can’t observe directly?

Other approaches:
- interviews
- surveys
Correlational Observation

- measure X and Y

Imaginary Scenario

X = time of day when Calculus 101 taken

Y = grade in Calculus 101

Result: Later time of day correlated with lower grade

We know there is a correlation, but we don’t know causation.

1 Directionality problem
   - does taking calculus later in the day lead to low grades?

OK
OR
-do people with low grades prefer to take classes later in the day?

(2) Third Variable problem
What if students who work have to take calculus at a later time and have little study time and thus get low grades?
Deduction

1. Start with a general theory.
2. Deduce what the theory says about a specific situation → hypothesis
3. Design a study to build evidence for your hypothesis → correlational observation or experiment

“Scientific method”
Experimentation

- strongest method for testing a causal hypothesis.

Does a change in \( X \) CAUSE a change in \( Y \)?

\( X \): independent variable
\( Y \): dependent variable

An experiment needs \( \geq 2 \) levels of the independent variable—also called TREATMENTS

\( \rightarrow \) solves directionality problem
An experiment should have most other variables “under control” — solv...