Research Methodology
Computer Science as Experimental Science

Lecture, 5. October, 2007
Research in Computer Science

1. Pick a relevant research question
2. Work on it and make some progress
3. Make sure your work is solid and well supported
4. Write scientific paper about work
5. Submit paper to conference, workshop, journal,...
6. If paper is accepted
   • Update CV, Present Paper, goto Step 1 or 2
7. Else Go To Step 1, 2, 3, 4 or 5
Shared experiment data

• Challenge problems
  – E.g. DIMACS

• Competitions
  – E.g., DARPA Grand Challenge

• Problem sets within field
  – E.g., large sets of SAT problems

• Randomly generated data
DIMACS Challenge - SAT

- Part of Second DIMACS challenge
- Problem sets
  - Generated by experts
  - Designed to push technology
- Tremendous impact
  - Kicked off massive technology leap
- Tricky to stay “real”
  - In this case, okay, as SAT is everywhere
Challenge problem – DARPA

• First attempt 2004
  – Drive from LA to Las Vegas (sort of)
  – Result: All vehicles failed within 1km

• Second attempt 2005
Real world challenge problems

• Real data from “industry”
  – Hard to come by
  – Very useful

• Example:
  – Aircraft assembly problem
  – Sanitized (but was really MD11)
  – Kicked off major improvements in scheduling technology
Key issues in methodology

• Sources of data – especially random
• Implementation differences
• Quality vs speed comparisons
• Variance in results
• Statistical analysis
• Usability and subjective analysis
• CS does not follow standard methods
The silver lining

• Most CS experiments work well
  – Can be repeated fairly easily
    • Or at least should be
  – Observation rarely impacts outcome

• Moving in the right direction
  – More rigorous methods appearing
  – More standardized approaches being used
Push-pull research