Research Methodology
Computer Science as Experimental Science

Lecture, 1. 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
Research in Computer Science

- Can we do X?
- If not, why not?
- If we can, how?
- How fast can we do X?
- How well can we do X?
- Theoretical analysis
- Experimental analysis
Role of theoretical analysis

• Determine bounds
  – What can be computed
  – What is an upper or lower bound

• Determine correctness
  – Ensure computations are correct

• And more
Experimental approach needed

• Theoretical analysis incomplete
  – Upper or lower bounds are loose
  – Based on ideal notions
  – Ignores constant factors

• Real problem set is different
  – Real-world problems not average or all
  – Constant factors and degree matter
Experimental methods

• Side-by-side comparisons
  – Apply multiple methods to same problems

• Absolute comparisons
  – Apply new technique to problem set
  – Compare to established reference

• Can-do demonstrations
  – Show something is possible

• Subjective evaluations
Shared experiment data

• Challenge problems
  – E.g. DIMACS
• Competitions
  – E.g., Robo-CUP
• Problem sets within field
  – E.g., large sets of SAT problems
• Randomly generated data
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