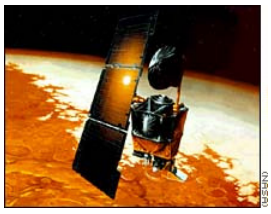
CMSC838M: Advanced Topics in Software Testing

- Atif M. Memon (atif@cs.umd.edu)
- 4115 A.V.Williams building
- Phone: 301-405-3071
- Office hours
  - Tu.Th. (3:15pm-5:00pm)
- Don't wait, don't hesitate, do communicate!!
  - Phone
  - E-mail
  - Office hours
- Course page
  - www.cs.umd.edu/~atif/Teaching/Fall2003

Mars Climate Orbiter

- Purpose: to relay signals from the Mars Polar Lander once it reached the surface of the planet
- Disaster: smashed into the planet instead of reaching a safe orbit
- Why: Software bug - failure to convert English measures to metric values
- $165M

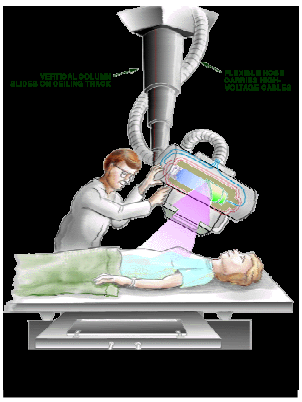


Shooting Down of Airbus 320

- 1988
- US Vicennes shot down Airbus 320
- Mistook airbus 320 for a F-14
- 290 people dead
- Why: Software bug - cryptic and misleading output displayed by the tracking software

THERAC-25 Radiation Therapy

- THERAC-25, a computer-controlled radiation-therapy machine
- 1986: two cancer patients at the East Texas Cancer Center in Tyler received fatal radiation overdoses
- Why: Software bug - mishandled race condition (i.e., miscoordination between concurrent tasks)



London Ambulance Service

- London Ambulance Service Computer Aided Dispatch (LASCAD)
- Purpose: automate many of the human-intensive processes of manual dispatch systems associated with ambulance services in the UK
  - functions: Call taking
- Failure of the London Ambulance Service on 26 and 27 November 1992

"Nice of You to Turn Up"

- Load increased
- Emergencies accumulated
- System made incorrect allocations
  - more than one ambulance being sent to the same incident
  - the closest vehicle was not chosen for the emergency
- At 23:00 on October 28 the LAS eventually instigated a backup procedure, after the death of at least 20 patients

More…

- "Software and its Impact: A Quantitative Assessment," by B.W. Boehm, Datamation, 19(5), 48-59 (1973)
  - Errors in medical software have caused deaths

More…

- "The development of software for ballistic-missile defense," by H. Lin, Scientific American, vol. 253, no. 6 (Dec. 1985), p. 48
  - British destroyer H.M.S. Sheffield; sunk in the Falkland Islands war; ship's radar warning system software allowed missile to reach its target
  - An Air New Zealand airliner crashed into an Antarctic mountain
  - North American Aerospace Defense Command reported that the U.S. was under missile attack; traced to faulty computer software - generated incorrect signals
  - Manned space capsule Gemini V missed its landing point by 100 miles; software ignored the motion of the earth around the sun

More…

- "Software Engineering: Report on a Conference Sponsored by the NATO Science Committee, Brussels, NATO Scientific Affairs Division," 1968, p. 121
  - An error in an aircraft design program contributed to several serious air crashes
- "Ghost in the Machine," Time Magazine, Jan. 29, 1990. p. 58
  - Dallas/Fort Worth air-traffic system began spitting out gibberish in the Fall of 1989 and controllers had to track planes on paper

More…

- Software Reliability: Principles & Practice, p. 25, by G. J. Myers
  - Apollo 8 spacecraft erased part of the computer's memory
  - Eighteen errors were detected during the 10-day flight of Apollo 14
  - An error in a single FORTRAN statement resulted in the loss of the first American probe to Venus

More…

- An Airbus A320 crashes at an air show
- A China Airlines Airbus Industries A300 crashes on April 26, 1994 killing 264
- Ariane 5 satellite launcher malfunction was caused by a faulty software exception routine resulting from a bad 64-bit floating point to 16-bit integer conversion

More…

- ACM SIGSOFT Software Engineering Notes, vol. 6, no. 2
  - F-18 fighter plane crashed due to a missing exception condition
- ACM SIGSOFT Software Engineering Notes, vol. 9, no. 5
  - F-14 fighter plane was lost to uncontrollable spin, traced to tactical software

More…

- Internet Risks Forum NewsGroup (RISKS), vol. 19, issue 56
  - CyberSitter censors "menu */ #define" because of the string "nu...de"
- London's Docklands Light Railway - train stopped in the middle of nowhere due to future station location programmed in software
- *ACM SIGSOFT Software Engineering Notes*, vol. 12, no. 3
  - Chicago cat owners were billed $5 for unlicensed dachshunds. A database search on "DHC" (for dachshunds) found "domestic house cats" with shots but no license

More…

- and many many more ….

CNN.com

PRINT THIS
Powered by clickability

Russia: Software bug made Soyuz stray

STAR CITY, Russia (AP) --A computer software error likely sent a Russian spacecraft into a rare ballistic descent that subjected the three men on board to chest-crushing gravity loads that made it hard to breathe, space experts said Tuesday.

"For me, for a moment, it felt like I was Atlas and I had the weight of the whole world on my shoulders," astronaut Donald Pettit, still a little woozy, told reporters at a crowded news conference.

CNN.com

PRINT THIS
Powered by clickability

Spread of buggy software raises new questions

NEW YORK (AP) --When his dishwasher acts up and won't stop beeping, Jeff Seigle turns it off and then on, just as he does when his computer crashes. Same with the exercise machines at his gym and his CD player.

"Now I think of resetting appliances, not just computers," says Seigle, a software developer in Vienna, Virginia.

Malfunctions caused by bizarre and frustrating glitches are becoming harder and harder to escape now that software controls everything from stoves to cell phones, trains, cars and power plants.

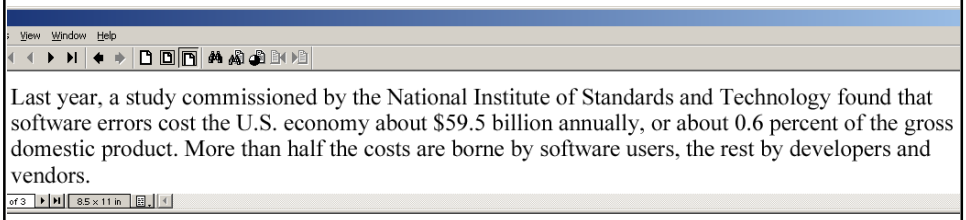
--A poorly programmed ground-based altitude warning system was partly responsible for the 1997 Korean Air crash in Guam that killed 228 people.

--Faulty software in anti-lock brakes forced the recall of 39,000 trucks and tractors and 6,000 school buses in 2000.

--The $165 million Mars Polar Lander probe was destroyed in its final descent to the planet in 1999, probably because its software shut the engines off 100 feet above the surface.

Economic Impact

- NIST study
  - On CNN.com - April 27, 2003

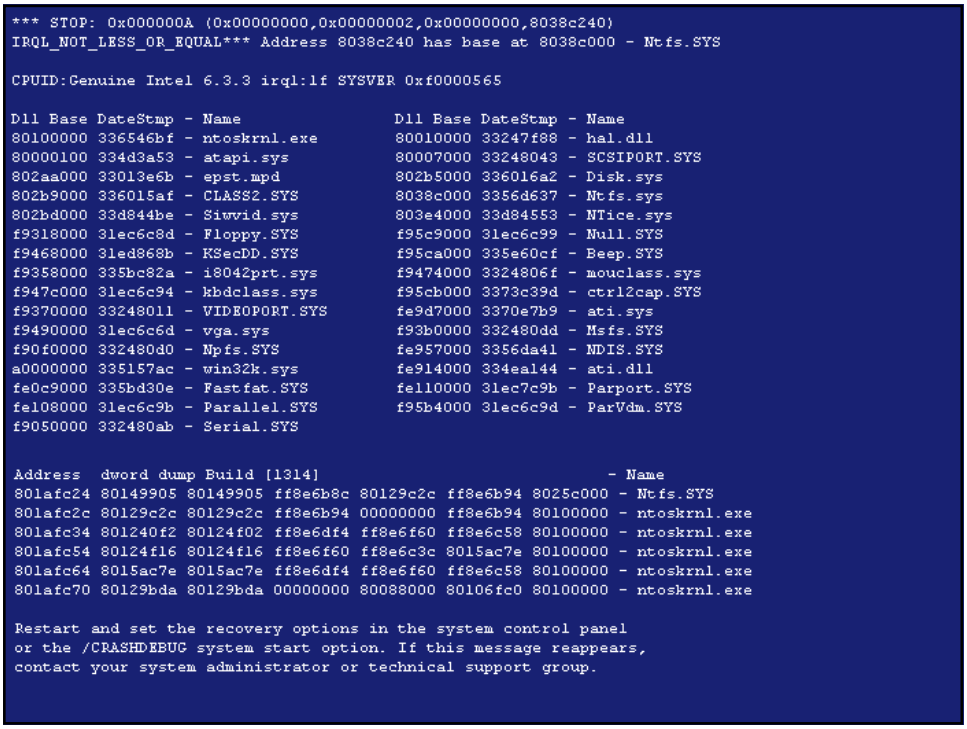

View Window Help

Last year, a study commissioned by the National Institute of Standards and Technology found that software errors cost the U.S. economy about $59.5 billion annually, or about 0.6 percent of the gross domestic product. More than half the costs are borne by software users, the rest by developers and vendors.

of 3 8.5 x 11 in

Expectation

- Can't we expect software to execute correctly?
- Carefully made programs
  - 5 faults/1000 LOC
  - 1M LOC will have 5000 faults
- Windows XP has 45M LOC
  - How many faults?
  - 45 x 5000 = 225,000
- Why not remove the faults?



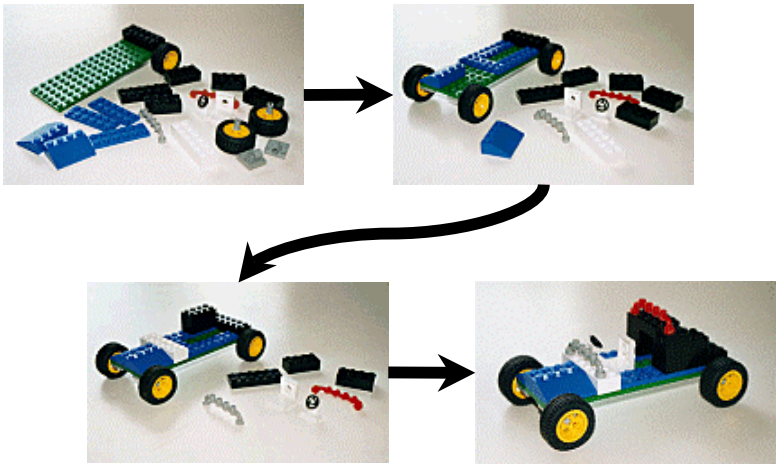
Joke?

- "If the automobile industry had developed like the software industry, we would all be driving $25 cars that get 1,000 miles to the gallon."
- "Yeah, and if cars were like software, they would crash twice a day for no reason, and when you called for service, they'd tell you to reinstall the engine."

How Cars Are Engineered (A Simple View)

- User requirements
  - Engine power, all-wheel, seating, comfort, MP3 player!!
- Detailed design
  - Blueprints, design documents
- Verify design
  - Simulation, prototyping
- Develop parts (components)
  - Test each component
  - Components may be reused
  - Mass produced
- Assemble the car
  - Test the car (Front/side crash tests, Stability tests)
  - Usability testing (Feedback from drivers/passengers)

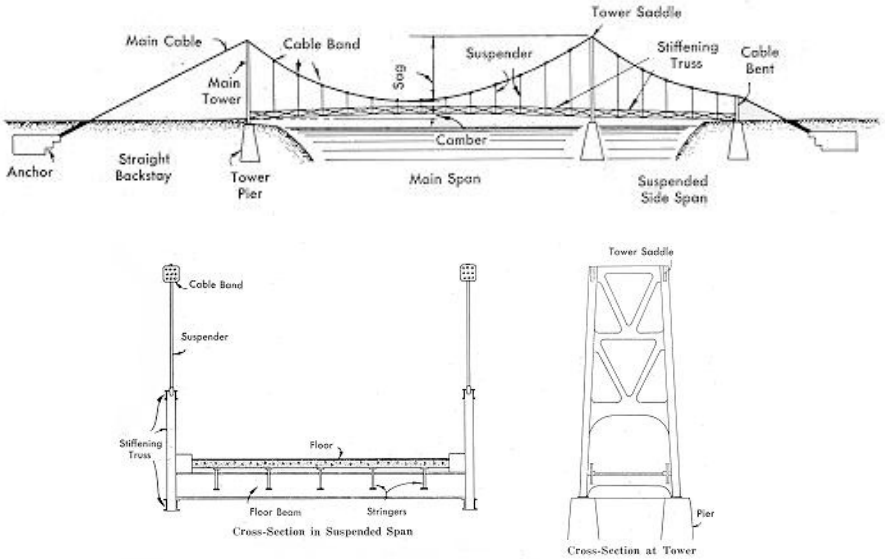
How Cars Are REALLY Engineered (A Detailed View)



But Seriously

- Features of many LEGO parts
  - Modularity
  - Reusability
    - Each part can be used in different places and ways
  - Flexibility of design
  - Compatibility
    - With other LEGO sets
- Building-blocks

Similar Techniques Used by Builders: Bridges



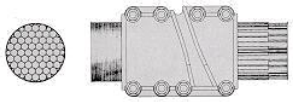
Detailed Design and Specifications

Galvanized Bridge Wire for Parallel Wire Bridge Cables. Recommended diameter .196 inch.

Galvanized Bridge Strand--consists of several bridge wires, of various diameters twisted together.

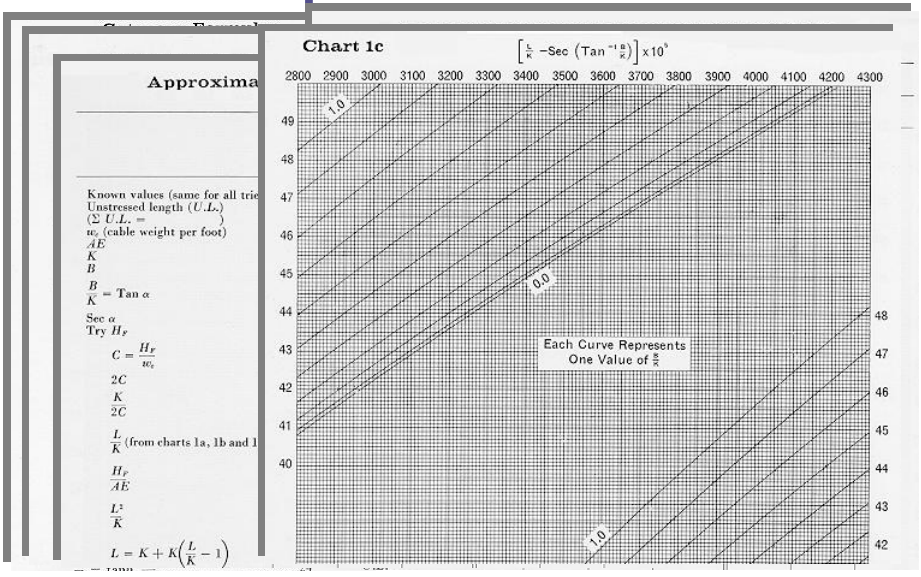
Galvanized Bridge Rope--consists of six strands twisted around a strand core.

Parallel Wire Cable



Detail of Main Cable and Cable Band. The wrapping wire is omitted at the right for clarity. Note the closed construction and aluminum fillers.

More Detailed Design and Specifications



Tacoma Narrows Bridge Disaster

They Make Mistakes Too!

- Even good design cannot guarantee a perfect product
- Need testing of all products including software

Interests

- Testing
- Apply it to different types of software (such as web)
- Student introductions and their interests!

Goals of the Course

- Discuss *advanced* software testing techniques
- Two parts of the course
  - Review testing fundamentals
  - State-of-the-art & emerging techniques
- What do I expect from students?

MS and Ph.D. Qualifying

- Is the course is valid for PhD qualifying coursework?
  - Yes (Software Engineering/Programming Languages)
- Is the course is valid for MS qualifying coursework?
  - Yes (Software Engineering/Programming Languages)
- Is the course is valid for MS comps?
  - Yes (Both Midterm and Final exams count towards the MS comps.)

Assessment

- 25% Mid-term Exam
- 25% Final Exam
- 20% Topic Presentation
  - 40 minutes
- 5% Project Presentation
  - 10 minutes
- 25% Term Project
  - chosen by the student and approved. May be a team (2 students) project, depending on the scope of the project

Exam Contents

- Midterm
  - Everything discussed in class
- Final exam
  - Everything discussed/presented after midterm

Student Presentations

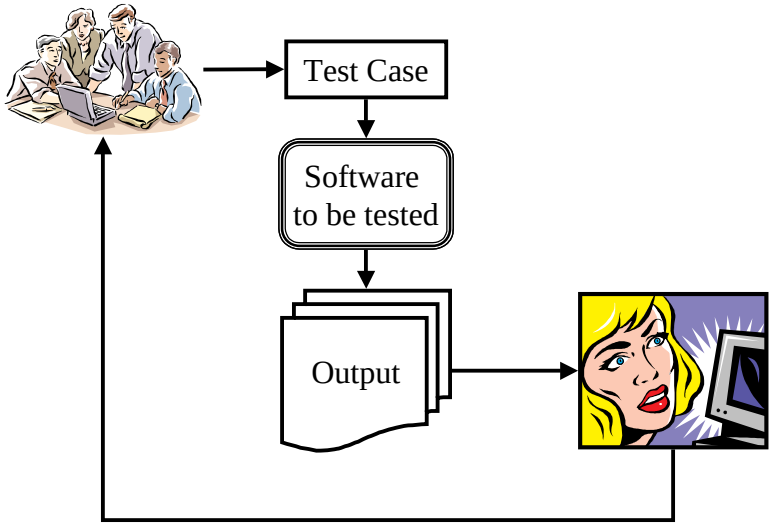
- Students must present a paper (perhaps from the list of suggested papers) on one of the listed topics
- 40 minutes
- Contents
  - Problem definition/motivation
  - What are the challenges?
  - Background literature surveyed by the authors
  - Specific technique developed in this paper
  - Weaknesses of the technique

Student Projects

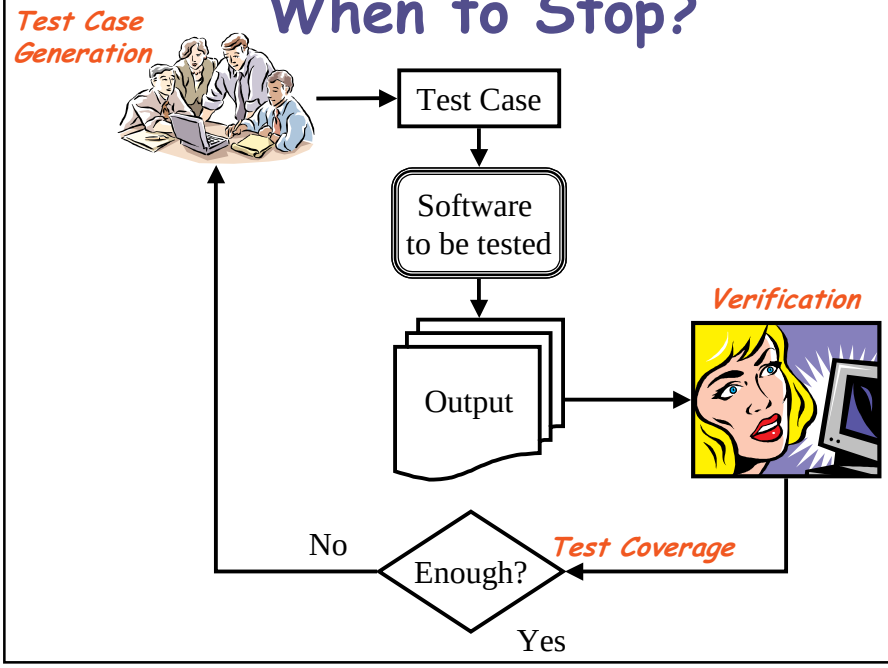
- Develop a new testing technique to test a particular type of software and experimentally demonstrate that it works (or does not work!!)
- Project Proposal Presentations
- Project Presentations

- Possible future research opportunities, publications, and conference talks
- Questions?

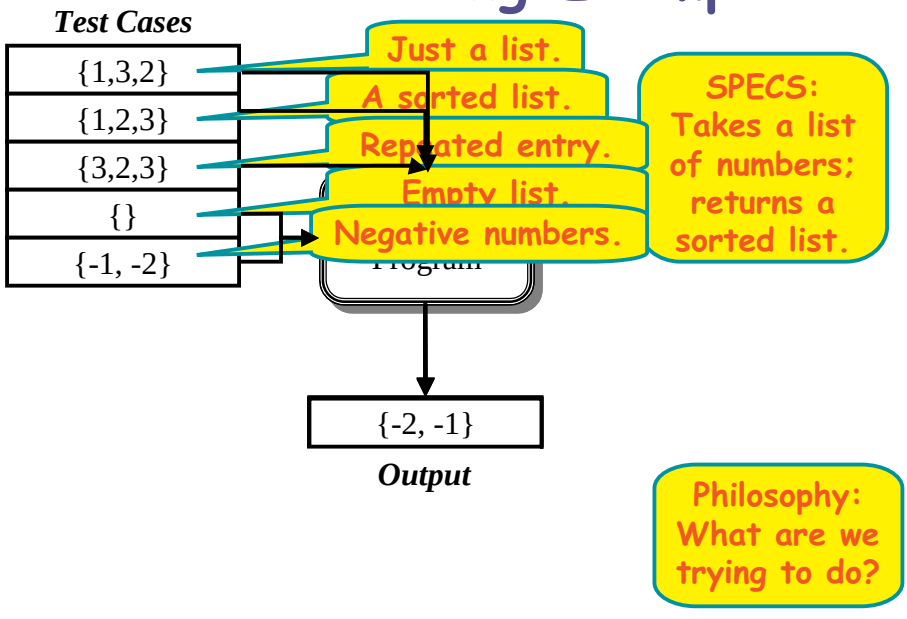
Testing: Our Experiences

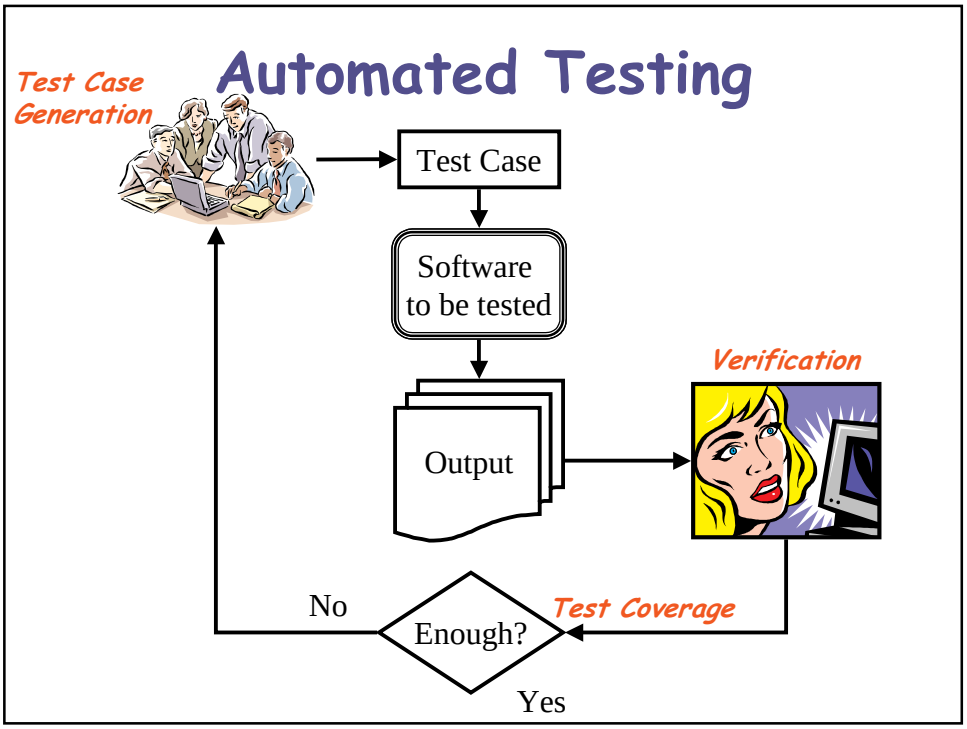


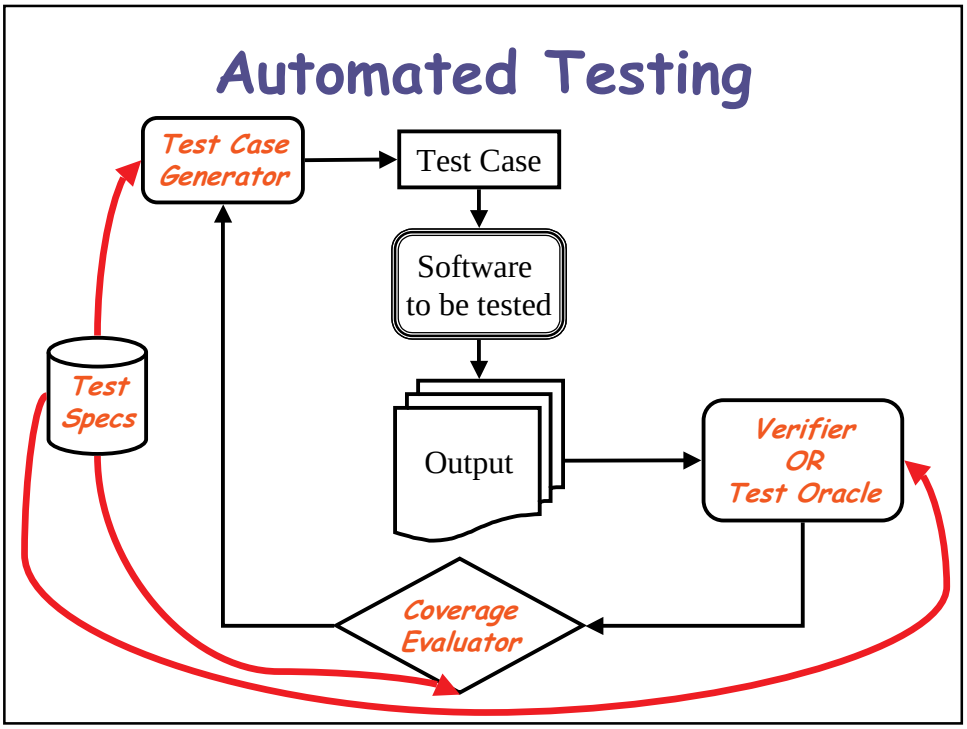
When to Stop?



A Real Testing Example




Automated Testing
Test Case Generation
Test Case
Software to be tested
Output
Verification
No
Enough?
Test Coverage
Yes


Automated Testing
Test Case Generator
Test Case
Software to be tested
Test Specs
Output
Verifier OR Test Oracle
Coverage Evaluator

Testing the New Version

Regression Testing

Discussion

- Different Software Types
  - Object-oriented
  - Component-based
  - Concurrent
  - Distributed
  - Graphical-user Interfaces
  - Web
- Different goals of testing
  - Usability
  - Security
  - Correctness
  - Performance …