



Informatics 43

LECTURE 14

“HOW DO WE KNOW THE SOFTWARE WORKS? (PART 2)”

Last Time

- Many failures are caused by a lack of good quality assurance (QA) practices
- QA = All activities designed to measure and improve quality in a product
 - Validation & verification

Today's lecture – **How do we know the software works?**

- Testing is the most common QA activity
 - Different levels: unit/functional/system
 - Goal: find and fix faults/errors/failures
- How to choose test cases
- How to know if the program's behavior is correct
- Bugs
- How do we know when we are done testing?

Testing Terminology



From IEEE610.12-90 (IEEE Standard Glossary of Software Engineering Terminology):

- ▶ **Mistake**: A human action that produces an incorrect result (“mistake” applies to both coder and user)
- ▶ **Fault / Bug**: An incorrect step, process, or data definition in a computer program (something you can point to in the code)
- ▶ **Error**: A difference between a computed result and the correct result (could be internal or external)
- ▶ **Failure**: The [incorrect] result of a fault (externally visible unexpected behavior or output)

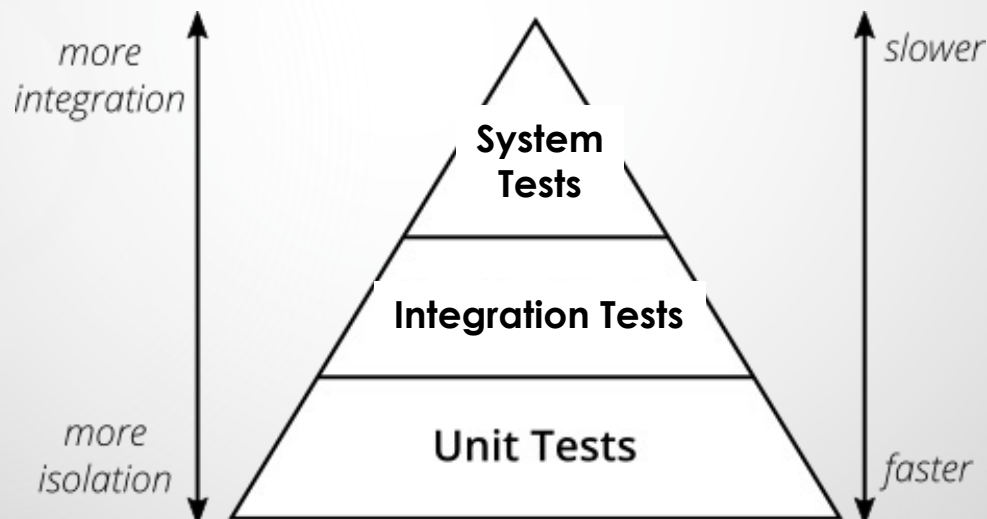
The common term “defect” usually means fault.

Who does the testing?

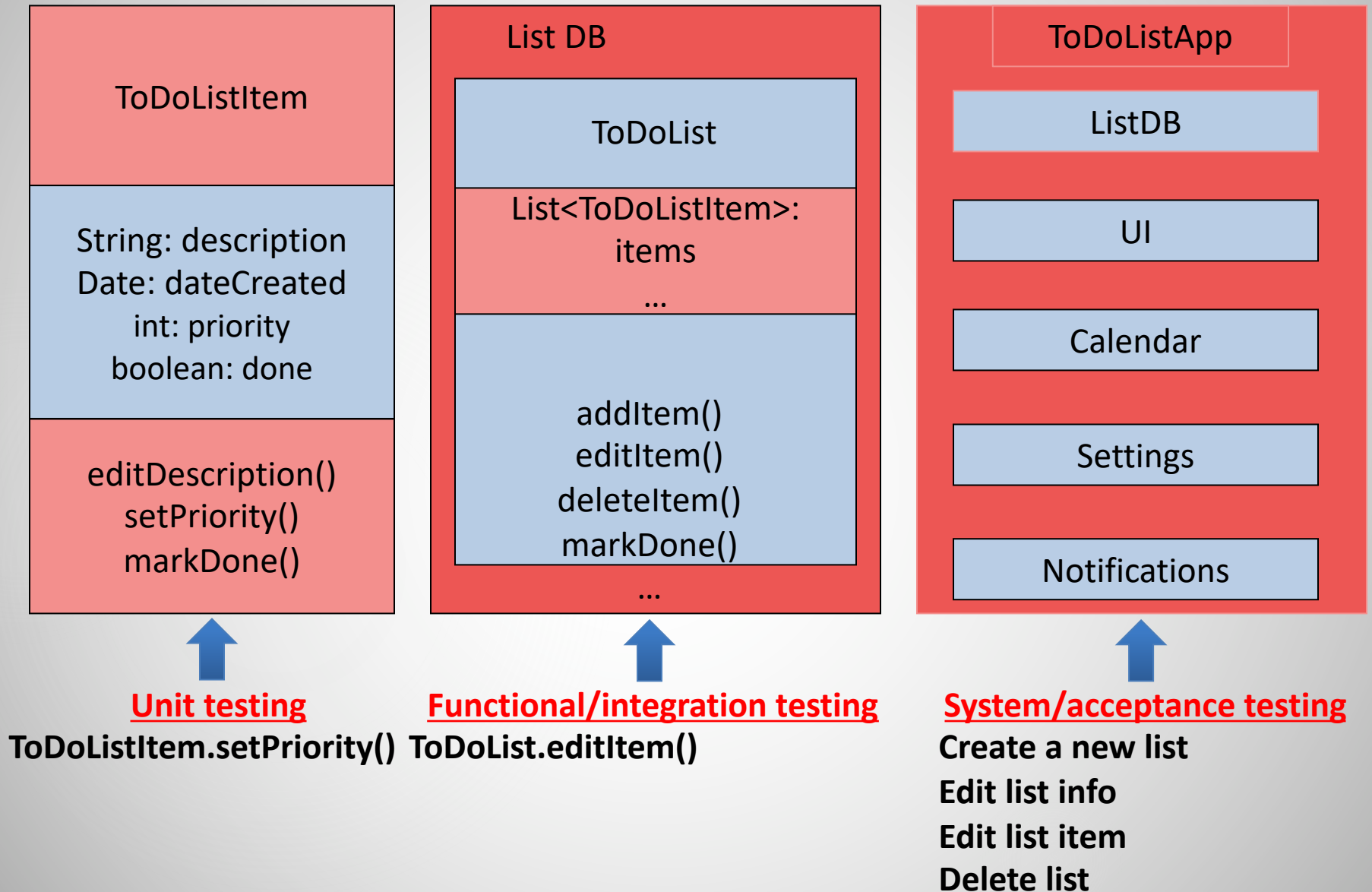
- Programmers
- Testers
- Users
 - Acceptance testing
 - Alpha testing
 - Beta testing
 - Crowdsourcing/bug bounties

Levels of Testing

- Unit testing
 - Testing of a single code unit
- Functional/integration testing
 - Testing of interfaces among integrated units
- System/acceptance testing
 - Testing of complete system for satisfaction of requirements



Levels of Testing



Today's lecture – **How do we know the software works?**

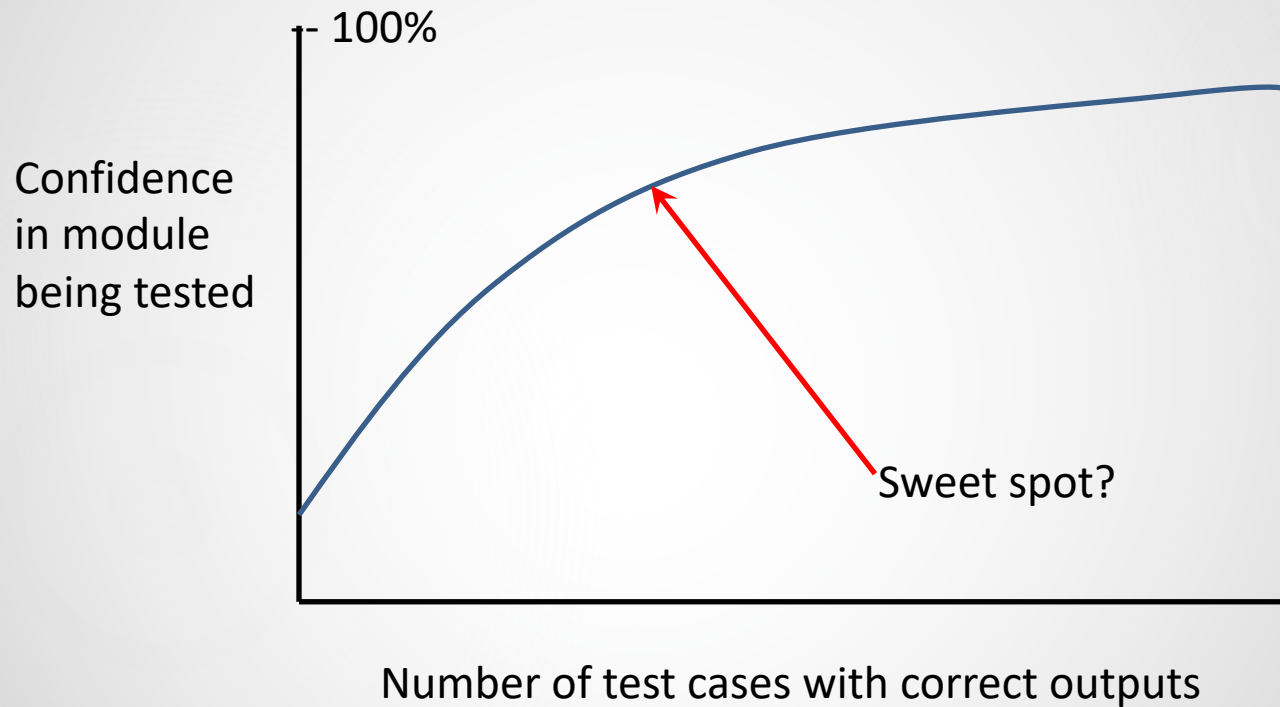
- How to choose test cases
- How to know if the program's behavior is correct
- Bugs
- How do we know when we are done testing?

How to choose test cases?

- There is usually an infinite number of possible test cases, so we must take a small sample

```
int multiplier(int a, int b) {  
    return a * b;  
}
```

Choosing test cases



Ways to choose test cases

- Intuition
- Specification (black-box testing)
 - Equivalence class partitioning
 - Boundary-value analysis
- Code (white-box testing)
 - Path analysis
- Existing test cases (regression testing)
- Faults

Today's lecture – **How do we know the software works?**

- How to choose test cases
- How to know if the program's behavior is correct
- Bugs
- How do we know when we are done testing?

Test Oracles

- A mechanism for deciding whether a test case execution succeeds or fails
- Difficult to automate

Test Oracle



Oracle Example: Cosine

- Your test execution shows $\cos(0.5) = 0.87758256189$
- Oracles?

Attendance Quiz

Systematic Testing and Oracles

Today's lecture – **How do we know the software works?**

- How to choose test cases
- How to know if the program's behavior is correct
- **Bugs**
- How do we know when we are done testing?

All software has bugs!



*"**All nontrivial code has defects**, and the probability of nontrivial defects increases with code size. The more code you use to solve a problem, the harder it gets for someone else to understand what you did and to maintain your code when you have moved on to write still larger programs."*

- Code Inflation, Holzmann (IEEE SW 2015)

All software has bugs!

Q: What is the required period of failure-free operation for conventional takeoffs and landings of the F35 Joint Strike Fighter?

- a) 6,000 years
- b) 600 years
- c) 60 years
- d) 6 hours



All software has bugs!

Q: What is the required period of failure-free operation for conventional takeoffs and landings of the F35 Joint Strike Fighter?

- a) 6,000 years
- b) 600 years
- c) 60 years
- d) **6 hours**



AND a recent government report stated that this target had not yet been realized!

Some bugs are bizarre...



Overview Code **Bugs** Blueprints Translations Answers

file incorrectly labeled as Erlang JAM file (OOo does not print on Tuesdays)

Bug #248619 reported by jgallo on 2008-07-15

This bug affects 2 people

112

Affects	Status	Importance	Assigned to	Milestone
▶ file (Debian)	Fix Released	Unknown	debbugs #514056	
▶ file (Ubuntu)	Fix Released	High	Unassigned	Ubuntu ubuntu-9.10
▶ Hardy	Fix Released	High	Colin Watson	Ubuntu ubuntu-8.04.3

What a fascinating bug!! My wife has complained that open office will never print on Tuesdays!?! Then she demonstrated it. Sure enough, won't print on Tuesday. Other applications print. I think this is the same bug. Here is my guess:

Print to a postscript file. Observe the line:

Source: <https://bugs.launchpad.net/ubuntu/+source/cupsys/+bug/255161>

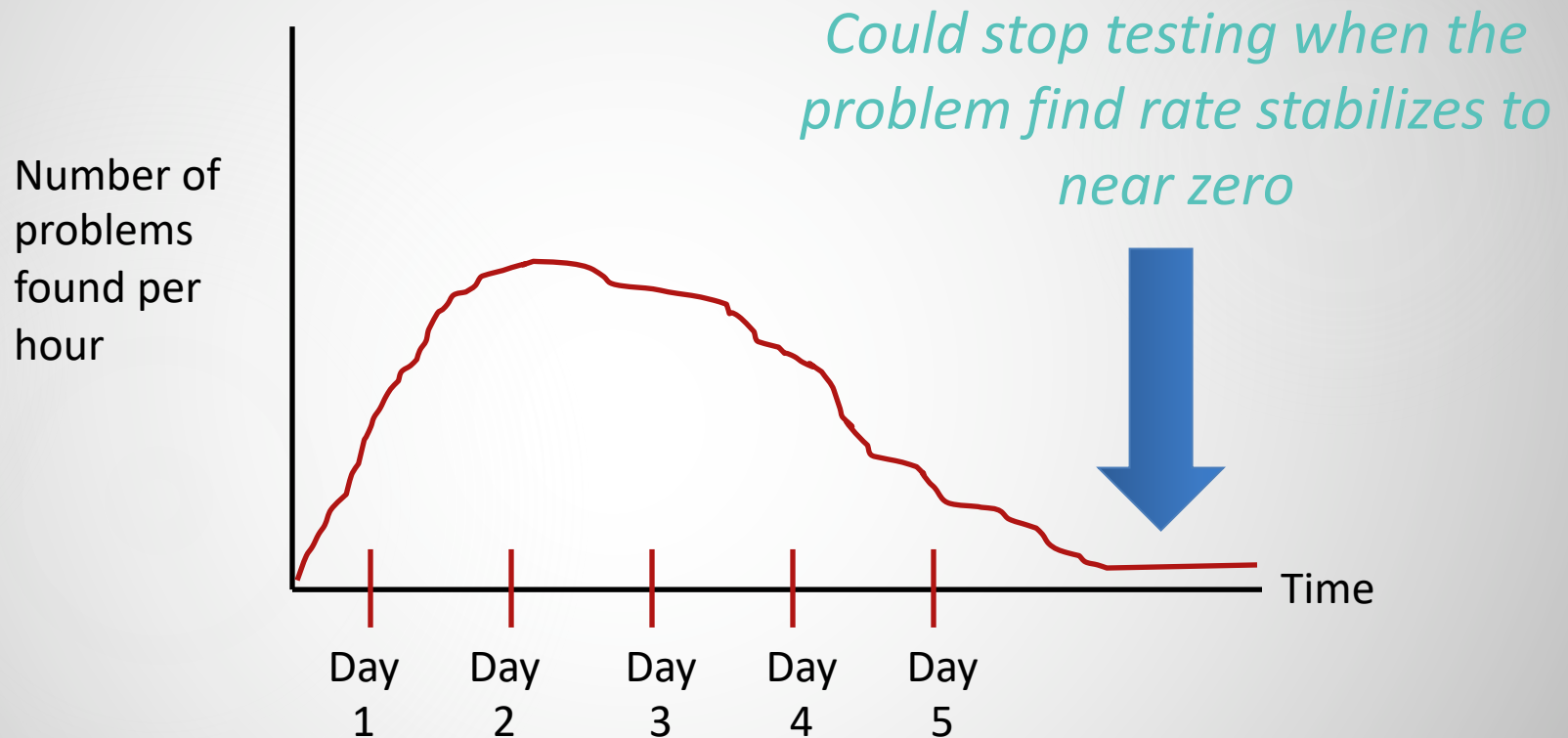
Today's lecture – **How do we know the software works?**

- How to choose test cases
- How to know if the program's behavior is correct
- Bugs
- How do we know when we are done testing?

How do we know when we are done testing (I)?

- Aim to reveal as many faults as possible in a fixed period of time with a fixed budget
 - Target specific areas of the system
- Aim to meet the quality requirements established for the project

How do we know when we are done testing (II)?



How do we know when we are done testing (IV)?

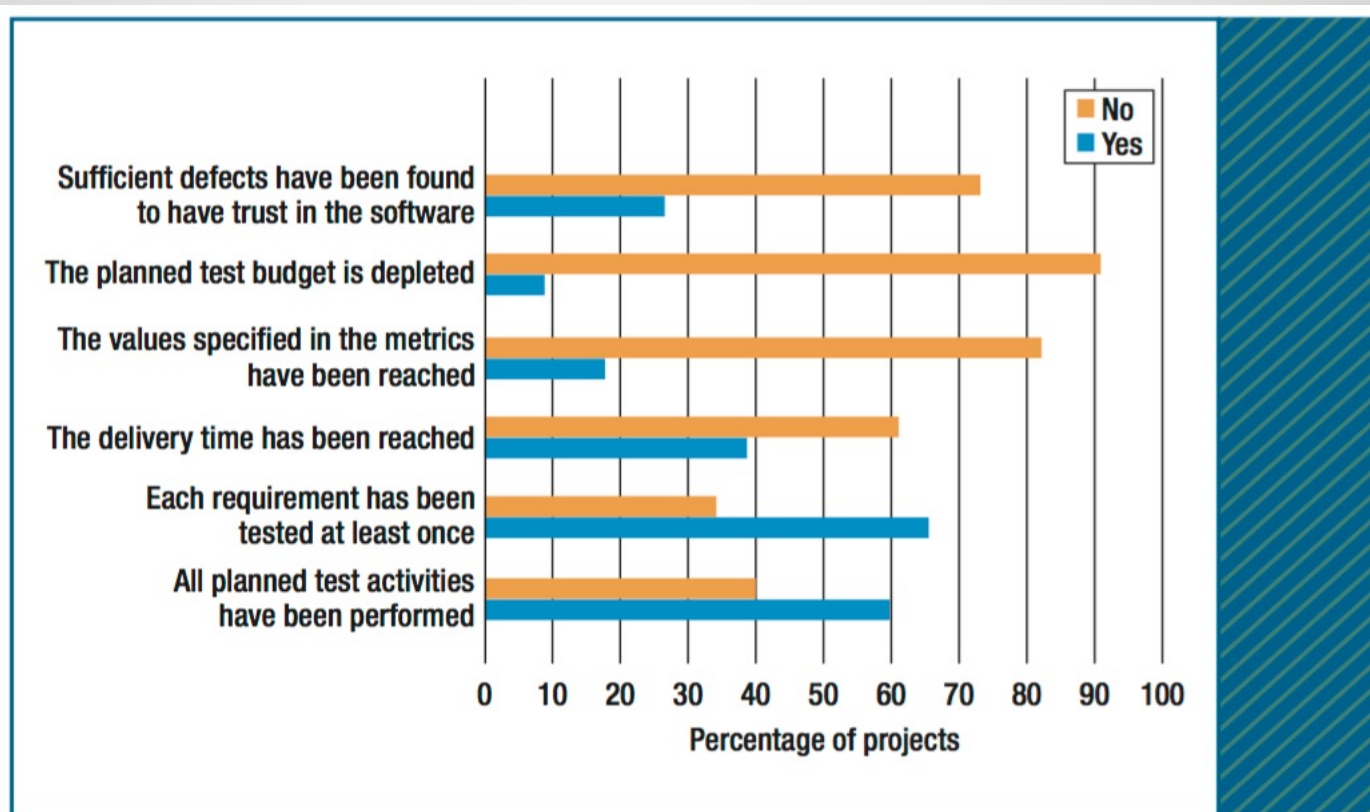


FIGURE 6. Participant responses regarding the exit criteria for test completion (103 responses). The most common criterion was to test each requirement at least once.

Today's lecture – **How do we know the software works?**

- Testing is the most common QA activity
 - Different levels: unit/functional/system
 - Goal: find and fix faults/errors/failures
- How to choose test cases
- How to know if the program's behavior is correct
- Bugs
- How do we know when we are done testing?

Quiz 6 Study Guide

- Testing
 - Validation/verification
 - Error, fault, failure
 - Testing goals, including Dijkstra quote
 - Levels of testing
 - Oracles
 - Different ways to know when we are done testing

Next Time

- Testing, part 3 (Black-box testing)