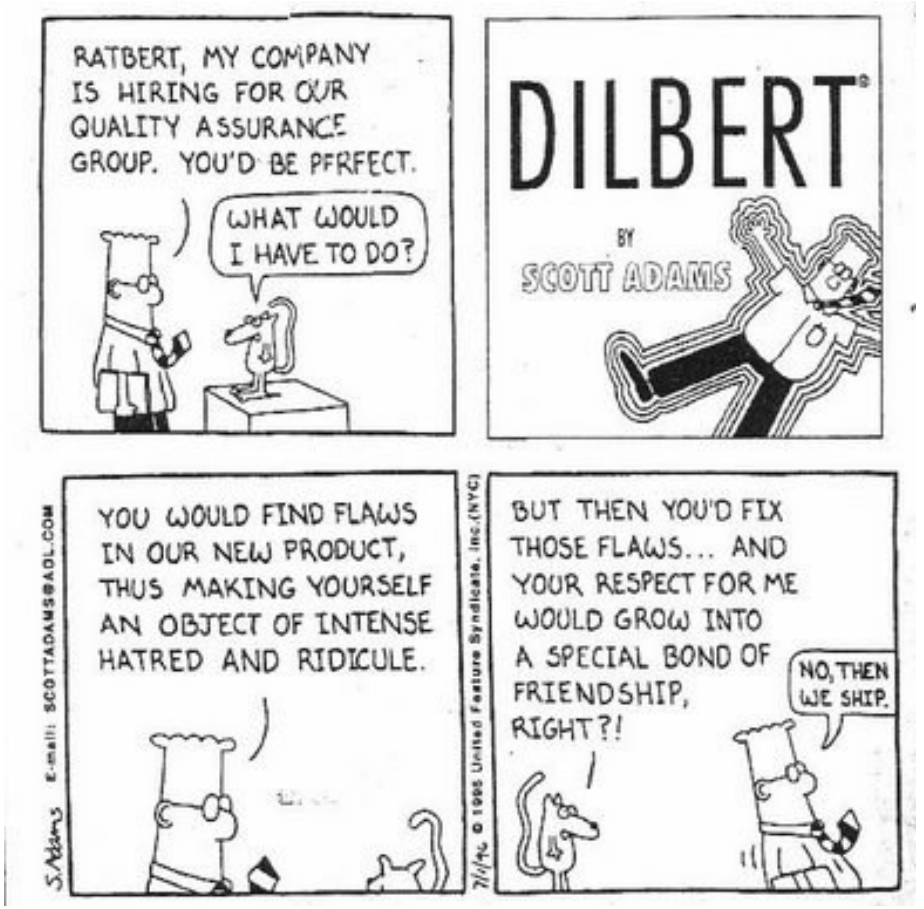
BLG 475E: Software Quality and Testing
Fall 2017-18

Dr. Ayşe Tosun
tosunay@itu.edu.tr

İstanbul Teknik Üniversitesi

Outline

- Software testing
  - Definitions
  - Objectives
  - Strategies
  - Types



Software testing

- 1st Software Quality Assurance Tool
  - Testing was confined to the final stage of the development, after the entire package had been completed.
- Other early SQA tools
  - Walkthroughs, code reviews, unit testing, etc.
- Testing is the most time consuming and expensive SQA.

Cost of software testing

- A survey in 1994 (by Perry)
  - 24% of project development budget is allocated to testing.
  - 32% of project management budget is allocated to testing activities.
  - Time resources
    - 27% of project time is spent to testing.
  - Participants
    - Allocated time to testing is around 45%
    - But pressure towards the release time force managers reduce scheduled testing time!
- According to Brooks
  - 50% of development costs is in unit and system testing.

Definitions

- "Testing is the process of executing a program with intention of finding errors." (Myers, 1979)

- "(1) The process of operating a system or component under specified conditions, observing or recording the results, and making an evaluation of some aspect of the system or component.

  (2) The process of analyzing a software item to detect the differences between existing and required conditions (that is, bugs) and to evaluate the features of the software item." (IEEE Std.610.12, 1990)

Definition by Galin 2004

- Software testing is a **formal** process carried out by a **specialized testing team** in which a software unit, several integrated software units or an entire software package are examined by **running the programs on a computer**. All the associated tests are performed according to **approved test procedures** on **approved test cases**.

Definition: Key Characteristics

- **Formal**
  - Software test plans (not *ad-hoc* examination)
- **Specialized testing team**
  - Independent team or external consultants specialized in testing to eliminate bias and to guarantee effective testing
- **Running the programs**
  - QA activities like code reviews (that does not involve program execution) are not considered as testing.
- **Approved test procedures**
  - Testing procedures approved as confirming to SQA procedures adopted by the developing organization.
- **Approved test cases**
  - Test cases should be approved before testing begins. No editions are allowed.

Objectives of software testing

- **Direct**
  - To identify and reveal as many errors as possible
  - To bring the tested software, after correction of the identified defects and retesting, to an acceptable level of quality.
  - To perform the required tests efficiently and effectively, within the limits of budgets and scheduling.
- **Indirect**
  - To compile a record of software errors for use in error prevention (by corrective and preventive actions)

"If your goal is to show the absence of errors you won't discover many. If your goal is to show the presence of errors, you will discover a large percentage of them." (Myers 1979)

Software test strategies

- Big bang testing
  - Testing in its entirety
- Incremental testing
  - Testing modules as they are completed (unit tests); then testing groups of *tested* modules integrated with newly completed modules (integration tests)
  - Top-down OR Bottom-up
  - Stubs and drivers

Incremental: Top-down

Stage 1: Unit tests of module 11.

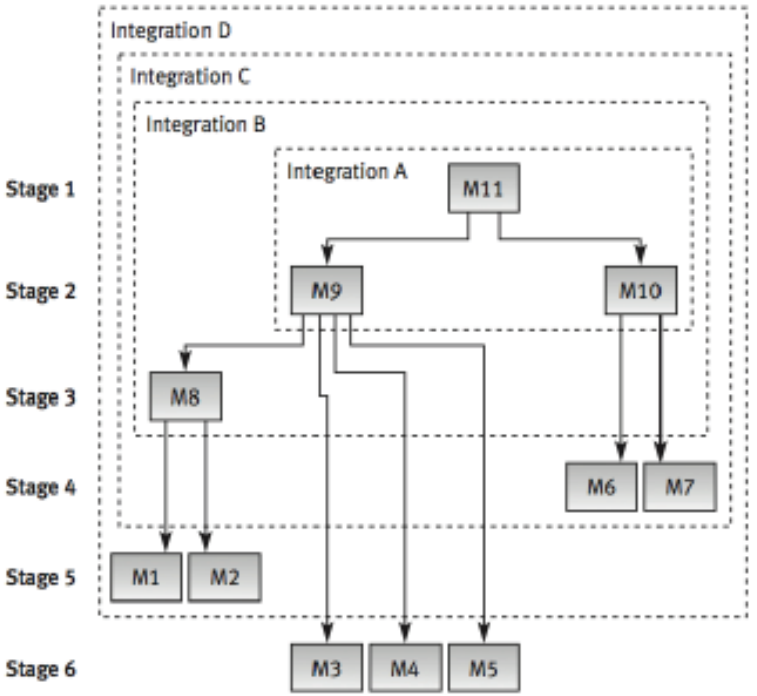
Stage 2: Integration test A of module 11 integrated with modules 9 and 10, developed in the current stage.

Stage 3: Integration test B of A integrated with module 8, developed in the current stage.

Stage 4: Integration test C of B integrated with modules 6 and 7, developed in the current stage.

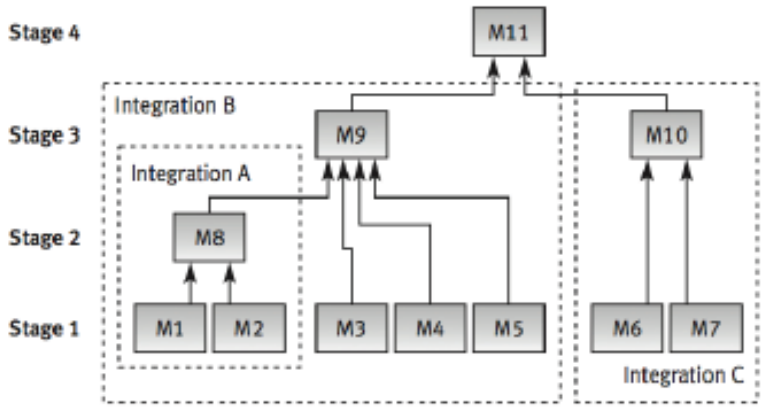
Stage 5: Integration test D of C integrated with modules 1 and 2, developed in the current stage.

Stage 6: System test of D integrated with modules 3, 4 and 5, developed in the current stage.



Galin D., "Software Quality Assurance: From theory to implementation", Addison-Wesley, 2004

Incremental: Bottom-up

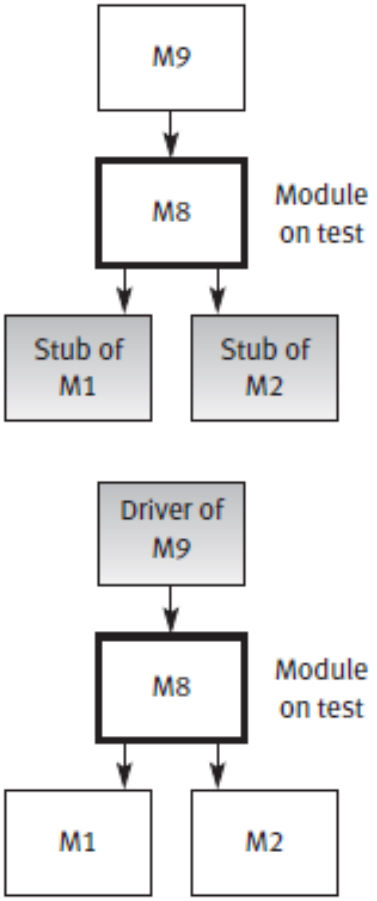


- Stage 1: Unit tests of modules 1 to 7.
- Stage 2: Integration test A of modules 1 and 2, developed and tested in stage 1, and integrated with module 8, developed in the current stage.
- Stage 3: Two separate integration tests, B, on modules 3, 4, 5 and 8, integrated with module 9, and C, for modules 6 and 7, integrated with module 10.
- Stage 4: System test is performed after B and C have been integrated with module 11, developed in the current stage.

Galin D., "Software Quality Assurance: From theory to implementation", Addison-Wesley, 2004

Incremental: Stubs

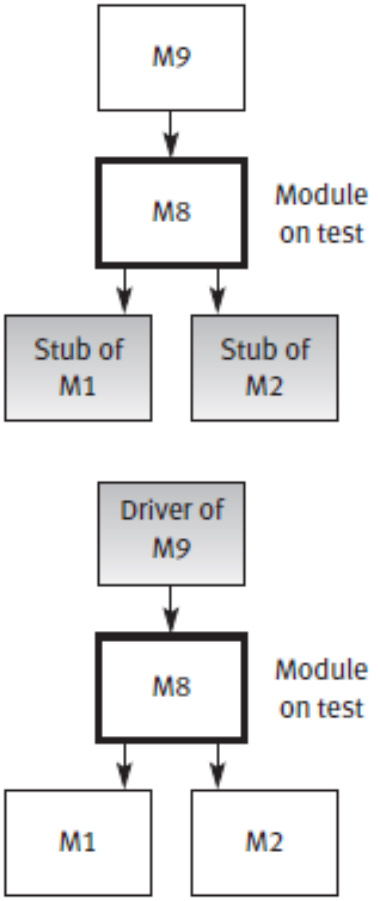
- Stubs are dummy modules
- Replacing unavailable lower level modules
- Required for top-down testing of incomplete systems
- Stub provides the results of subordinate module, yet to be coded.



From Galin, 2004

Incremental: Drivers

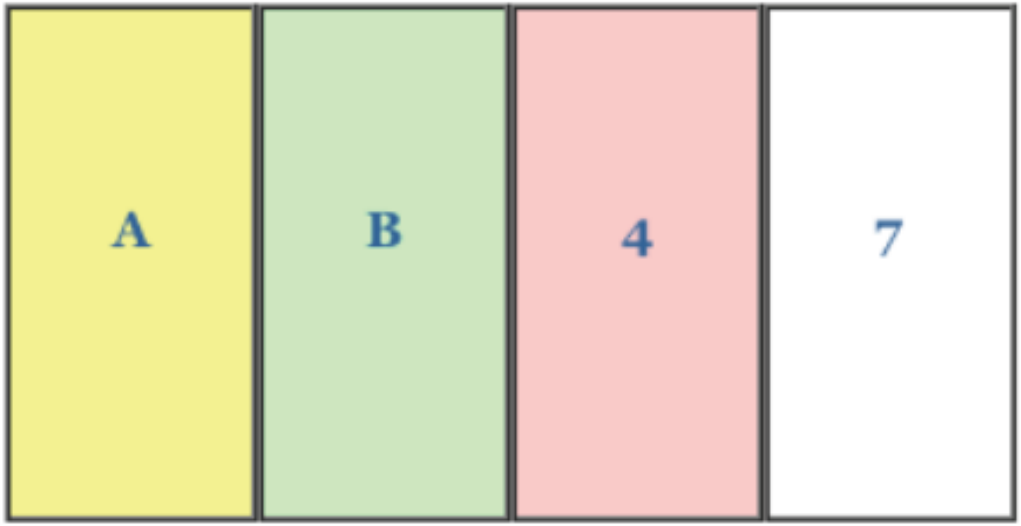
- Required in bottom-up testing
- Substitute module, but of the upper-level module



From Galin, 2004

Big bang vs. Incremental

- Big bang
  - Effective in small and simple programs
  - Prospects of keeping on schedule and within budget are substantially reduced.
  - Despite vast resources invested, effectiveness of the approach is insufficient.
- Incremental
  - Usually on small software modules, as unit or integration tests.
  - Easier to identify higher % errors. → prevents mitigation of errors to later stages in a lifecycle
  - Identification and correction of errors are much simpler with fewer resources (limited volume of software).
  - Increased quantity of programming resources for preparation of stubs/drivers.
  - Increased number of testing operations for a single program



"If a card has a vowel on one side, then it has an even number on the other side."

One side indicates the top side that is visible now.

Which card(s) must you turn over to determine the following statement is FALSE?

a) A and 4?
b) B and 4?
c) A and 7?
d) B and 7?
e) A and B?
f) 4 and 7?

Software test classifications

- According to testing concept
  - Black-box → erroneous outputs
  - White-box → glass-box
- According to requirements
  - Based on McCall's quality model
    - For "correctness"
    - Output correctness tests
      - Documentation tests
      - Availability tests
      - Data processing and calculations correctness tests

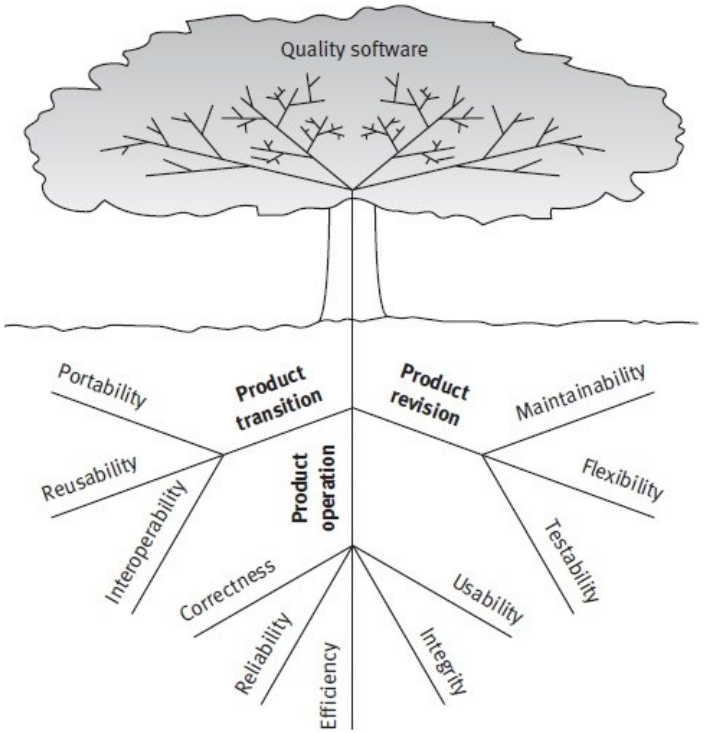


Table 9.1: Software quality requirements and test classification

Factor category	Quality requirement factor	Quality requirement sub-factor	Test classification according to requirements
Operation	1. Correctness	1.1 Accuracy and completeness of outputs, accuracy and completeness of data	1.1 Output correctness tests
		1.2 Accuracy and completeness of documentation	1.2 Documentation tests
		1.3 Availability (reaction time)	1.3 Availability (reaction time) tests
		1.4 Data processing and calculations correctness	1.4 Data processing and calculations correctness tests
		1.5 Coding and documentation standards	1.5 Software qualification tests
	2. Reliability		2. Reliability tests
	3. Efficiency		3. Stress tests (load tests, durability tests)
	4. Integrity		4. Software system security tests
	5. Usability	5.1 Training usability 5.2 Operational usability	5.1 Training usability tests 5.2 Operational usability tests
Revision	6. Maintainability 7. Flexibility 8. Testability		6. Maintainability tests 7. Flexibility tests 8. Testability tests
Transition	9. Portability 10. Reusability 11. Interoperability	11.1 Interoperability with other software 11.2 Interoperability with other equipment	9. Portability tests 10. Reusability tests 11.1 Software interoperability tests 11.2 Equipment interoperability tests

From Galin 2004

White-box or black-box test for different requirements

Test classification according to requirements	White box testing	Black box testing
1.1 Output correctness tests		+
1.2 Documentation tests		+
1.3 Availability (reaction time) tests		+
1.4 Data processing and calculations correctness tests	+	
1.5 Software qualification tests	+	
2. Reliability tests		+
3. Stress tests (load tests and durability tests)		+
4. Software system security tests		+
5.1 Training usability tests		+
5.2 Operational usability tests		+
6. Maintainability tests	+	+
7. Flexibility tests		+
8. Testability tests		+
9. Portability tests		+
10. Reusability tests	+	
11.1 Software interoperability tests		+
11.2 Equipment interoperability tests		+

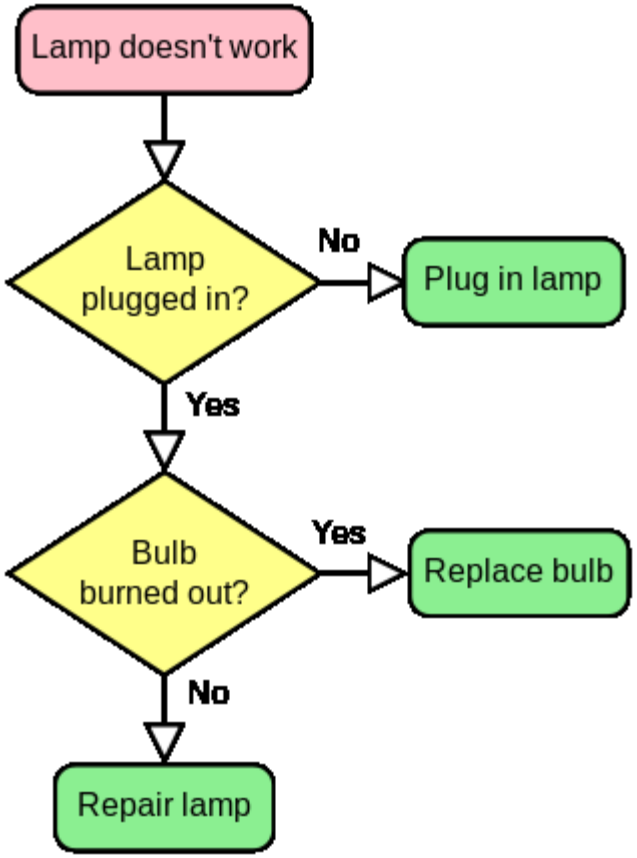
From Galin 2004

White box testing (structure-based)

+ Statement-by-statement checking of code
+ Performance of line coverage follow-up
  - LOC that have not yet been executed
+ Performance of path coverage follow-up
  - If-then-else or do-while
  - (more difficult to achieve than line coverage)
+ Quality of coding
- Vast resources utilized (above those in black-box)
- Inability to test software performance in terms of availability, reliability, other testing factors.

Flow charts to represent basic path testing

- Diamonds: Options covered by conditional statements
- Rectangles: Software sections connecting those conditional statements

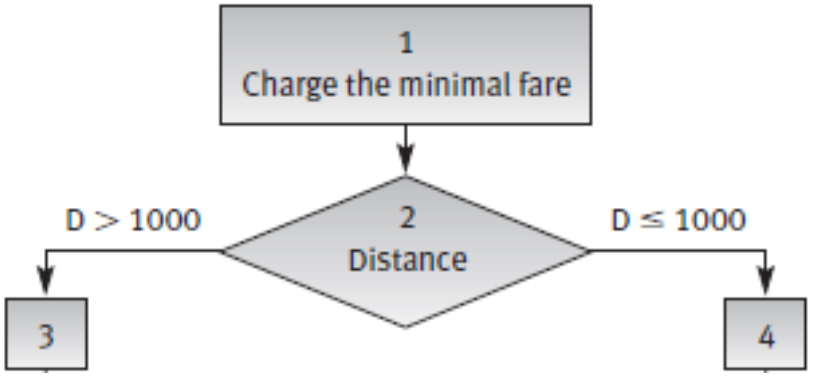


Example: Imperial Taxi Services (ITS) taximeter

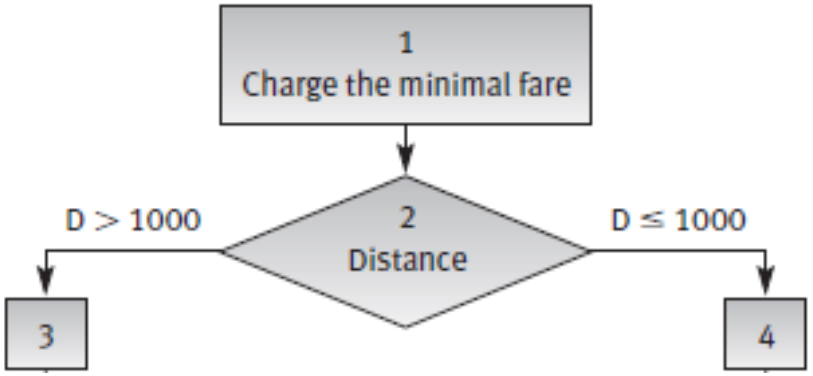
- Serves one-time passengers and regular clients.

1) Minimal fare: $2. This fare covers the distance traveled up to 1000 yards and waiting time (stopping for traffic lights or traffic jams, etc.) of up to 3 minutes.
2) For every additional 250 yards or part of it: 25 cents.
3) For every additional 2 minutes of stopping or waiting or part thereof: 20 cents.
4) One suitcase: no charge; each additional suitcase: $1.
5) Night supplement: 25%, effective for journeys between 21.00 and 06.00.
6) Regular clients are entitled to a 10% discount and are not charged the night supplement.

1) Minimal fare: $2. This fare covers the distance traveled up to 1000 yards and waiting time (stopping for traffic lights or traffic jams, etc.) of up to 3 minutes.
2) For every additional 250 yards or part of it: 25 cents.

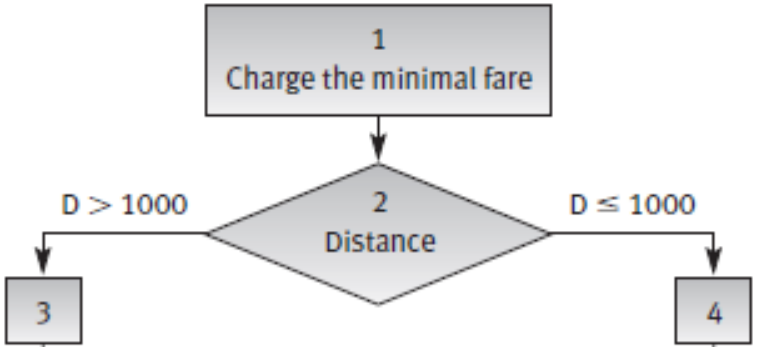


1) Minimal fare: $2. This fare covers the distance traveled up to 1000 yards and waiting time (stopping for traffic lights or traffic jams, etc.) of up to 3 minutes.
2) For every additional 250 yards or part of it: 25 cents.



Quiz: Complete the flow graph

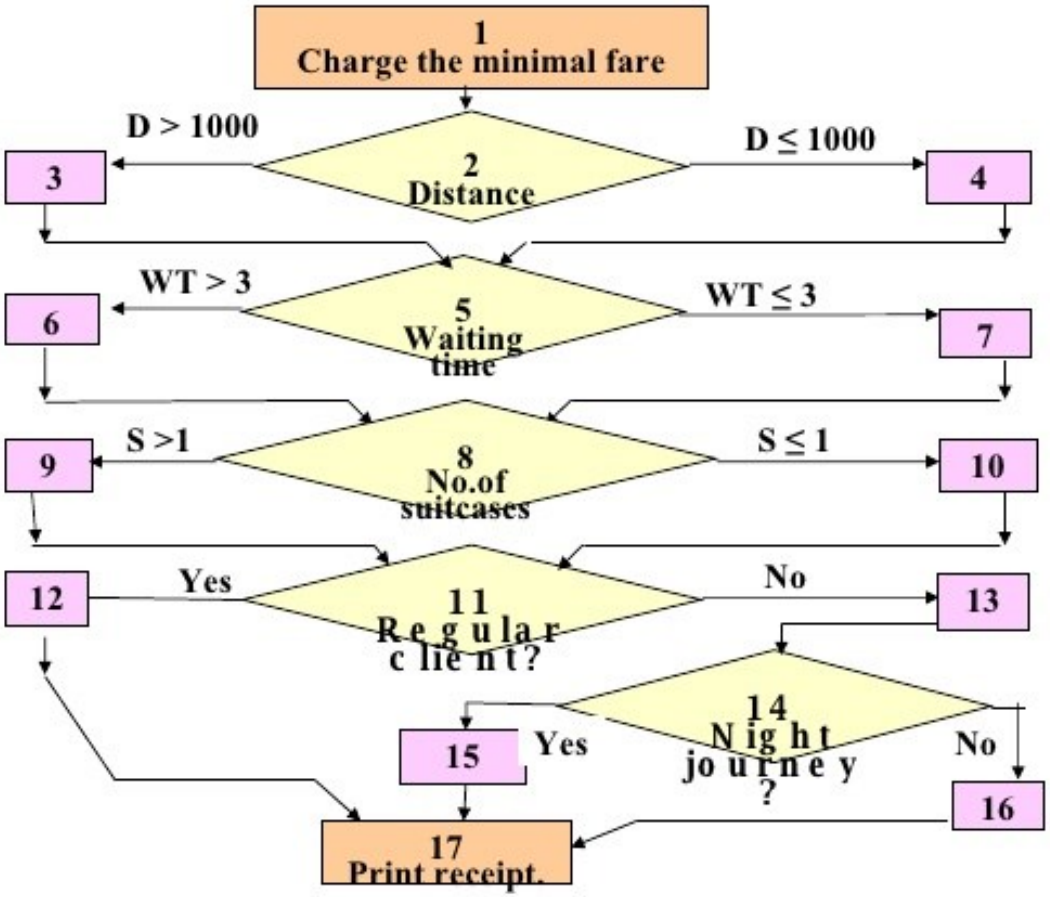
QUIZ: Complete the testing flow graph



1) Minimal fare: $2. This fare covers the distance traveled up to 1000 yards and waiting time (stopping for traffic lights or traffic jams, etc.) of up to 3 minutes.
2) For every additional 250 yards or part of it: 25 cents.
3) For every additional 2 minutes of stopping or waiting or part thereof: 20 cents.
4) One suitcase: no charge; each additional suitcase: $1.
5) Night supplement: 25%, effective for journeys between 21.00 and 06.00.
6) Regular clients are entitled to a 10% discount and are not charged the night supplement.

OHT 9.11

ITS - Flow chart



Galin, *SQA from theory to implementation*

Black-box testing

- Output correctness tests
- Apart from output correctness and maintainability tests, most of the other testing classes are unique to black box testing.

+ Fewer resources

- Coincidental aggregation of several errors may produce the correct response for a test case.

Equivalence class partitioning (black-box)

Output correctness: Greater consumption of test resources

- Equivalence class (EC)
  - Set of input variable values that produce the same output results or that are processed identically.
- Valid EC: only valid states
- Invalid EC: only invalid states

Rules:

- Test cases are defined for *valid ECs* and *invalid ECs.*
- Test cases are added as long as there are uncovered ECs.
- Total number of test cases for valid ECs <= valid ECs
- One test case for each invalid EC.

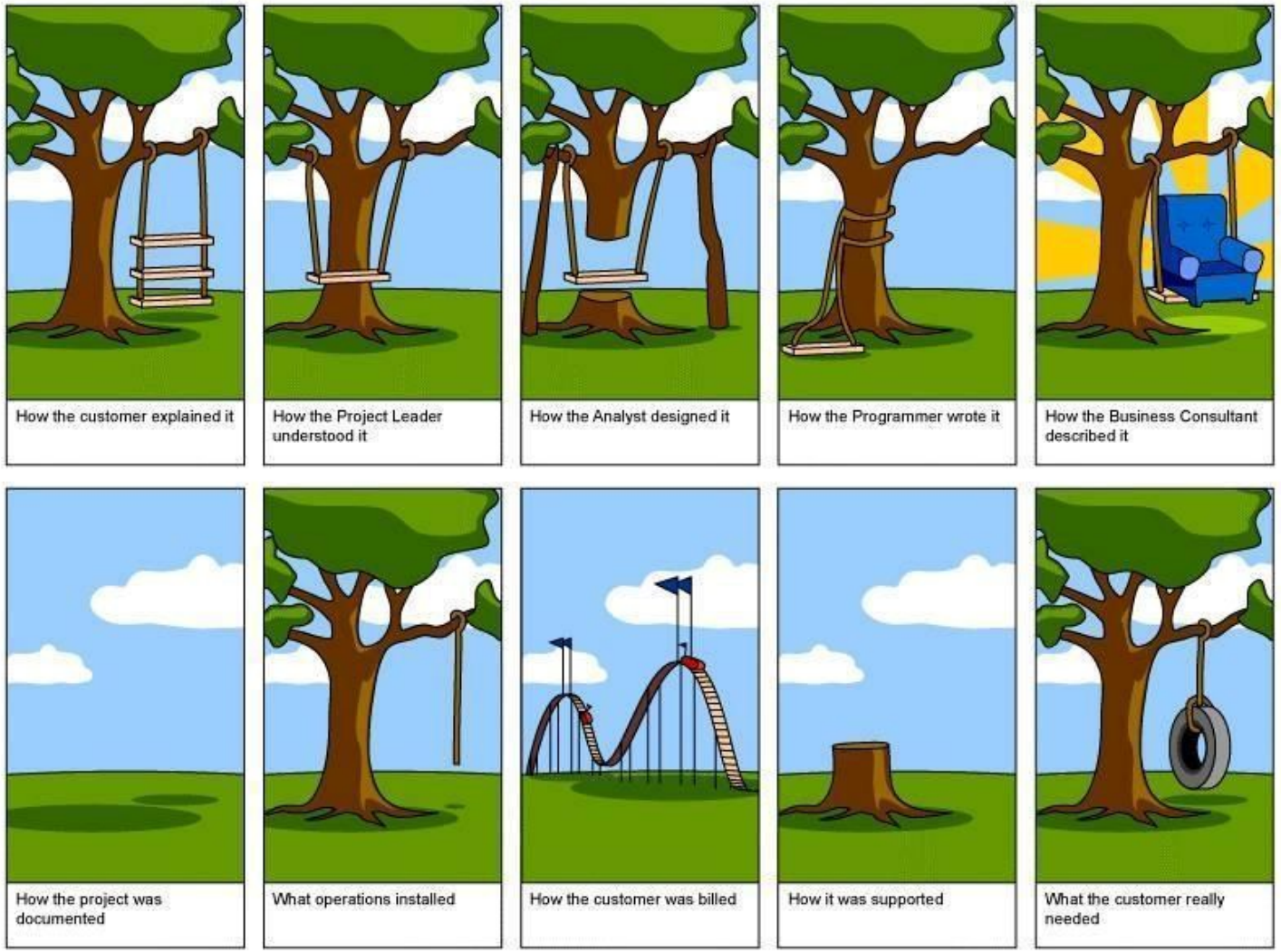
Examples

- Valid input is a month number (1-12)
  - Equivalence classes are: [-∞..0], [1..12], [13.. ∞]
- Valid input is one of ten strings representing a type of fuel
  - Equivalence classes are
    - 10 classes, one for each string
    - A class representing all other strings
- **Example**: The landing gear must be deployed whenever the plane is within 2 minutes from landing or takeoff, or within 2000 feet from the ground. If visibility is less than 1000 feet, then the landing gear must be deployed whenever the plane is within 3 minutes from landing or lower than 2500 feet
  - Equivalence classes? Valid EC, Boundaries and Invalid EC

Equivalence classes

- Additional hints
  - (1) Look for *extreme* range of variables
    - Try very small numbers to check if sign works
    - Try very large 16-bit and 32-bit integers
  - (2) Look for *maximum size of memory* variables
    - e.g. for a depth-first search, set the length of the path to be very long and check for memory overflow.
  - (3) One EC must include *all members of a group*
    - e.g. for a program accepting usernames as input
      - All string of alphabet (starting with capital or lower case letters)
      - Limit or extend the size of the character
      - Check for non-ascii characters

Test Levels


How the customer explained it
How the Project Leader understood it
How the Analyst designed it
How the Programmer wrote it
How the Business Consultant described it
How the project was documented
What operations installed
How the customer was billed
How it was supported
What the customer really needed

Types of testing

- Unit test
- Integration test
- System test
- Acceptance (α,β) test
- Regression test
- Others (load, stress etc)

- What to test?
- Which sources to use for test cases?
- Who is to perform the tests?
- Where to perform the tests?
- When to terminate the tests?

Test types

- Unit
  - Internal behavior of individual units against specifications (functionality)
- Integration
  - Test interactions between modules based on design/interface specifications (stubs and drivers)
- System
  - Verification of the completed system functionality as a whole against requirement specifications.

Unit testing

- Testing individual units of source code in isolation and validating that each works properly
- A unit is the smallest testable part of an application
  - Examples
    - In procedural programming: a function or an individual program
    - In OOP: a method or a class
- Example unit testing frameworks
  - JUnit for Java
  - CPPUnit or GoogleTest for C++

Unit-Test Considerations

- Unit interface is tested to ensure that the information properly flows into and out of the unit under test
- Local data structures are examined to ensure that data stored maintains its integrity during all steps in an algorithm execution
- All independent paths through the control structure are exercised to ensure that all statements in a unit have been executed at least once
- Boundary conditions are tested to ensure that the unit operates properly at boundaries established to limit or restrict processing
- All error handling paths

Integration Testing

- Testing units of source code together and validating that they together work properly
  - Big bang
  - Top down OR bottom up
- Integration testing more difficult
  - Intended behavior of program more difficult to characterize than its parts
- Errors of specification come up during integration testing
  - Each unit does its part, but parts put together don't do the overall job
  - Parts were erroneously specified

Who performs the tests?

- Unit tests
  - Software development team (developer task)
- Integration tests
  - Often, integration test team
- System tests
  - Independent testing team (internal or external)
  - Vendor
- Acceptance
  - Customer

Acceptance tests

- Definition by ISQTB: '*…formal testing with respect to user needs, requirements, and business processes conducted to determine whether a system satisfies the acceptance criteria and to enable the user, customers or other authorized entity to determine whether or not to accept the system*'
- Also known as user acceptance test
- Two types
  - α-test: environment controlled by vendor
  - β-test: real settings of customer

Regression Testing

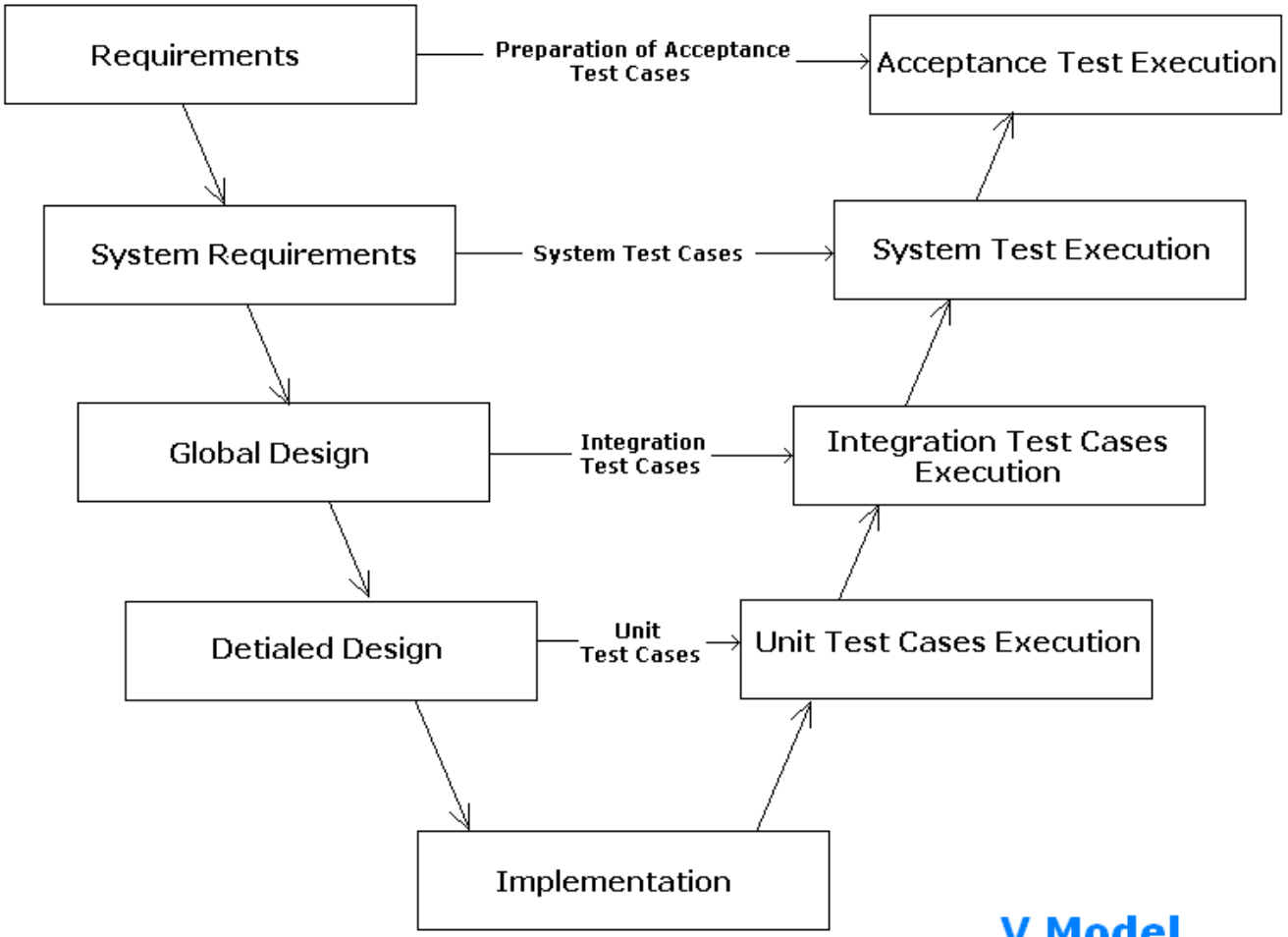
- Regression testing is the re-execution of some subset of tests that have already been conducted to ensure that changes have not propagated unintended side effects
- Whenever software is corrected, some aspect of the software configuration (the program, its documentation, or the data that support it) is changed.
- Regression testing helps to ensure that changes (due to testing or for other reasons) do not introduce unintended behavior or additional errors.

Regression testing

- It tends to be far too expensive to re-run every single test case every time a change is made to software.
- Hence only a *subset of the previously-successful test cases* is actually re-run.
- This process is called *regression testing*.
  - The tests that are re-run are called regression tests.
- Regression test cases are carefully selected to cover as much of the system as possible.

Smoke Testing

- A common approach for creating “daily builds” for product software
- Smoke testing steps:
  - Software components that have been translated into code are integrated into a “build.”
    - A build includes all data files, libraries, reusable modules, and engineered components that are required to implement one or more product functions.
  - A series of tests is designed to expose errors that will keep the build from properly performing its function.
    - The intent should be to uncover “show stopper” errors that have the highest likelihood of throwing the software project behind schedule.
  - The build is integrated with other builds and the entire product (in its current form) is smoke tested daily.
    - The integration approach may be top down or bottom up.


Requirements
Preparation of Acceptance Test Cases
Acceptance Test Execution
System Requirements
System Test Cases
System Test Execution
Global Design
Integration Test Cases
Integration Test Cases Execution
Detialed Design
Unit Test Cases
Unit Test Cases Execution
Implementation

V Model

Other tests

- Availability: reaction time
- Stress: load and durability
- Security: Access control, database backup, file recovery, logging of transactions
- More will be in the next lectures..

References

- Galin, D. 2004. Software Quality Assurance: From theory to implementation, Addison Wesley.
- Lethbridge, T., Laganiere, R. 2005. Object-Oriented Software Engineering: Practical Software Development using UML and Java, McGraw-Hill.
- Turhan, B. 2011. Software Quality and Testing course, Lecture Notes 7.
- Bener, A. 2006. Software testing: Test case design, Lecture Notes.
- Far, B.H. Software Reliability & Software Quality, Chapter 11: Preparing and Executing Test, Dept. of Electrical and Computer Engineering, University of Calgary.