

Méthodes de Test

Cours 1 — Introduction to Test-Driven Development

Stefano Zacchioli
zack@pps.univ-paris-diderot.fr

Laboratoire PPS, Université Paris Diderot

2013-2014

URL <http://upsilon.cc/zack/teaching/1314/methstest/>
Copyright © 2013 Stefano Zacchioli
License Creative Commons Attribution-ShareAlike 3.0 Unported License
<http://creativecommons.org/licenses/by-sa/3.0/>



Outline

- 1 Development processes
- 2 Test-Driven Development
- 3 xUnit & jUnit
- 4 Bibliography

Test-Driven Development (TDD)

Test-Driven Development, or TDD, is an iterative **software development process** which uses **very short development cycles** and leverages **tests** to provide constant **feedback** to software developers.

Goal: “clean code that works”, i.e. develop better software, less stressfully.

Test-Driven Development (TDD)

Test-Driven Development, or TDD, is an iterative **software development process** which uses **very short development cycles** and leverages **tests** to provide constant **feedback** to software developers.

Goal: “clean code that works”, i.e. develop better software, less stressfully.

The “Test-Driven Development” expression is often (ab)used to talk about 2 distinct things:

- the TDD development process
- the **xUnit** family of **testing frameworks**
 - ▶ e.g. JUnit, [Python] unittest, cppUnit, OUnit, NUnit, PHPUnit, ...which have been designed to support the TDD development process, but can also be used when adopting different development process

Outline

- 1 Development processes
- 2 Test-Driven Development
- 3 xUnit & junit
- 4 Bibliography

Reminder — Development process

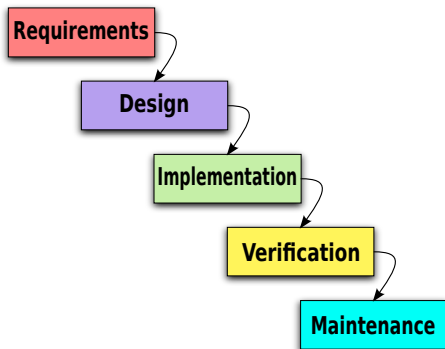
Definition (Software development process)

A *software development process* is a structured set of activities which lead to the production of some software.

Some software development **activities**:

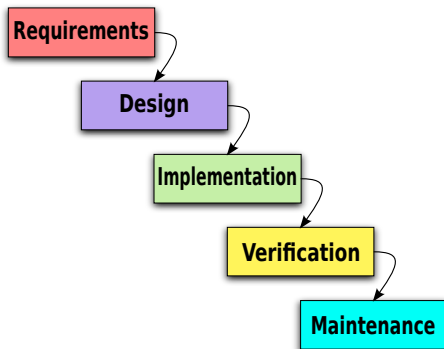
- 1 requirement specification
- 2 design
- 3 implementation
- 4 verification
- 5 maintenance

Waterfall model



[https://en.wikipedia.org/wiki/File:
Waterfall_model_\(1\).svg](https://en.wikipedia.org/wiki/File:Waterfall_model_(1).svg)

Waterfall model

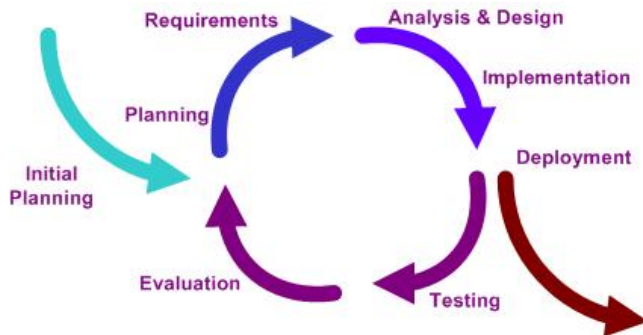


- that's the theory
- in practice:
 - ▶ feedback loops
 - ▶ increasingly more costly
 - ▶ that need to be prevented

[https://en.wikipedia.org/wiki/File:Waterfall_model_\(1\).svg](https://en.wikipedia.org/wiki/File:Waterfall_model_(1).svg)

Iterative development

A family of models where development happens **incrementally**, through repeated **iterations** of development activities.

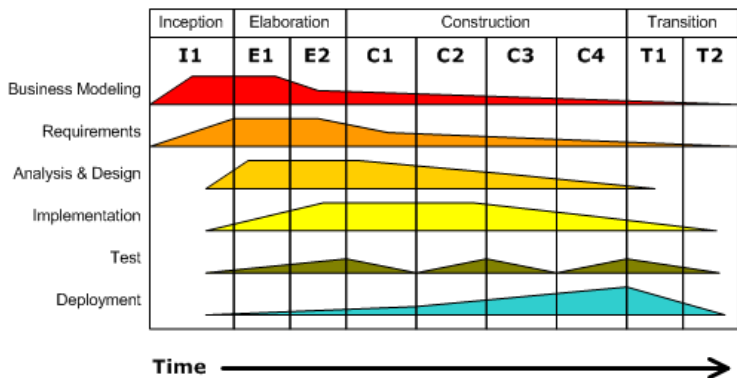


https://en.wikipedia.org/wiki/File:Iterative_development_model_V2.jpg

Key benefit: **feedback** loop.

Rational Unified Process (RUP)

IBM's Rationale Unified Process is a well-established example of an iterative development model, tailored for OOP.

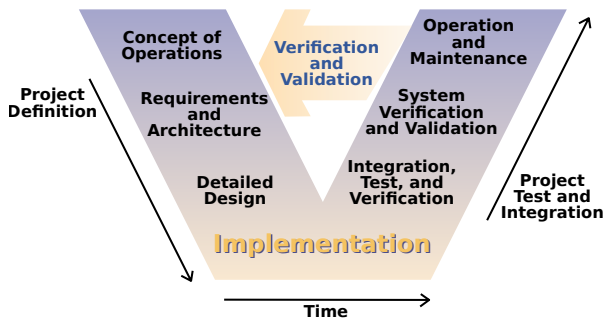


<https://en.wikipedia.org/wiki/File:Development-iterative>

- phases vs work-flow
- notable work-flows: **deployment, testing**

V-Model

Not really a software development model.



https://en.wikipedia.org/wiki/File:Systems_Engineering_Process_II.svg

Rather a (simplistic) **view on the waterfall model** that correlates the initial “definition” phases with the final “delivery” ones.

The V-model helps to think about **test purposes**.

A hierarchy of tests

Disclaimers:

- there are other hierarchies/taxonomies, on different angles
- (as it often happens in SWE) **terminology** is not clear cut
- the **granularity trend**—from small to big—however matters and is agreed upon

Test hierarchy

acceptance	<i>does the whole system work?</i>
integration	<i>does our code work against (other) code (we can't change)?</i>
unit	<i>do our code units^a do the right thing and are convenient to work with?</i>

^ain a broad sense: might be classes, objects, modules, etc. depending on the available abstraction mechanisms

Acceptance test

Does the whole system work?

Acceptance tests represent **features** that the system should have. Both their lack and their misbehaviour imply that the system is not working as it should. Intuition:

- 1 feature → 1+ acceptance test(s)
- 1 user story → 1+ acceptance test(s) (when using **user stories**)

Exercise (name 2+ acceptance tests for this “user login” story)

After creating a user, the system will know that you are that user when you login with that user's id and password; if you are not authenticated, or if you supply a bad id/password pair, or other error cases, the login page is displayed. If a CMS folder is marked as requiring authentication, access to any page under that folder will result in an authentication check.

<http://c2.com/cgi/wiki?AcceptanceTestExamples>

Preview: we will use acceptance tests to guide feature development

Integration test

Does our code work against (other) code (we can't change)?

“Code we can't change” =

- 3rd party libraries/framework
 - ▶ be them proprietary or open source
- code developed by other teams that we don't “own”
 - ▶ (strict code ownership is bad, though)
- code that we do not want to/cannot modify in the current phase of development, for whatever reason

Example

- our `BankClient` should not call the `getBalance` method on `BankingService` before calling `login` and having verified that it didn't throw an exception
- `xmlInitParser` should be called before any other parsing function of `libxml2`
- the `DocBook` markup returned by `CMSEditor.save` should be parsable by `PDFPublisher's` constructor

Unit test

Do our code units do the right thing and are convenient to work with?

Before implementing any unit of our software, we have (to have) an idea of **what the code should do**. Unit tests show **convincing evidence** that—in a limited number of cases—it is actually the case.¹

Example (some unit tests for a List module)

¹remember: tests reveal bugs, but don't prove their absence!

Unit test

Do our code units do the right thing and are convenient to work with?

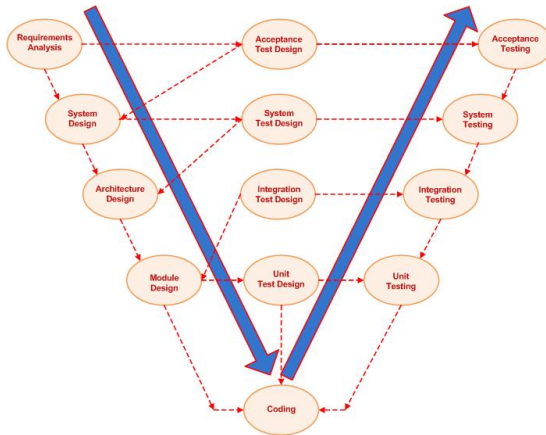
Before implementing any unit of our software, we have (to have) an idea of **what the code should do**. Unit tests show **convincing evidence** that—in a limited number of cases—it is actually the case.¹

Example (some unit tests for a List module)

- calling `List.length` on an empty list returns 0
- calling `List.length` on a singleton list returns 1
- calling `List.last` after `List.append` returns the inserted element
- calling `List.head` on an empty list raises an exception
- calling `List.length` on the concatenation of two lists returns the sum of the respective `List.length`s
- ...

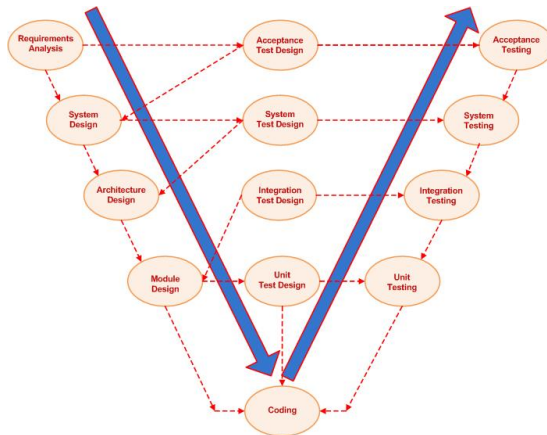
¹remember: tests reveal bugs, but don't prove their absence!

Tests in the V-Model



<https://en.wikipedia.org/wiki/File:V-model.JPG>

Tests in the V-Model



<https://en.wikipedia.org/wiki/File:V-model.JPG>

For TDD we will “hack” unit, integration, acceptance tests, and use them in an arguably more clever way...

Outline

- 1 Development processes
- 2 Test-Driven Development**
- 3 xUnit & junit
- 4 Bibliography

Development as learning

- every software development project tries *something* that has **never been tried before**
 - ▶ otherwise you wouldn't be doing it. . .
- due to constraints, developers often use technologies they don't completely master
 - ▶ new technologies, old technologies used in unfamiliar contexts, etc
- all stakeholders (developers, managers, customers) **learn as the project progresses**

Problem: as we don't know everything at the beginning, there will be **unexpected changes** during the project.

How do we cope with them?

Nested feedback loops

Idea

- 1 we use **empirical feedback** to learn about the system
- 2 we **store what we learn** in the system itself, for future use

To do so, we organize development as **nested feedback loops** with increasing time periods and scopes in the organization (file, unit, product, team, etc.), e.g.:

- pair programming period: seconds
- unit tests seconds–1 minute
- acceptance tests minutes
- daily meeting 1 day
- iterations 1 day–1 week
- releases 1 week–months

We want feedback as quickly as possible. If something slips through an inner loop, it will (hopefully) be caught by an outer one.

Expecting the unexpected

Practices that (empirically) help coping with unexpected changes:

- **constant testing**

- ▶ when we change something we might introduce **regressions**
- ▶ to avoid that we need to constantly test our system
- ▶ doing it manually doesn't scale ⇒ **automated testing**

- **simple design**

keep the code as simple as possible
optimize for simplicity

- ▶ as we will have to change it, we want **code that is easy to understand and modify**
- ▶ empirical studies show that **developers spend more time reading code** than writing it
- ▶ clean design doesn't come for free, to achieve it we must constantly **refactor**
- ▶ **test suites give you courage to refactor**, and do other changes, thanks to their tight feedback loop

TDD principle

So we have test suites. Why do we need TDD?

TDD principle

So we have test suites. Why do we need TDD? Because:

- developers don't like writing tests
- testing is often seen as a 2nd class development activity

TDD principle

So we have test suites. Why do we need TDD? Because:

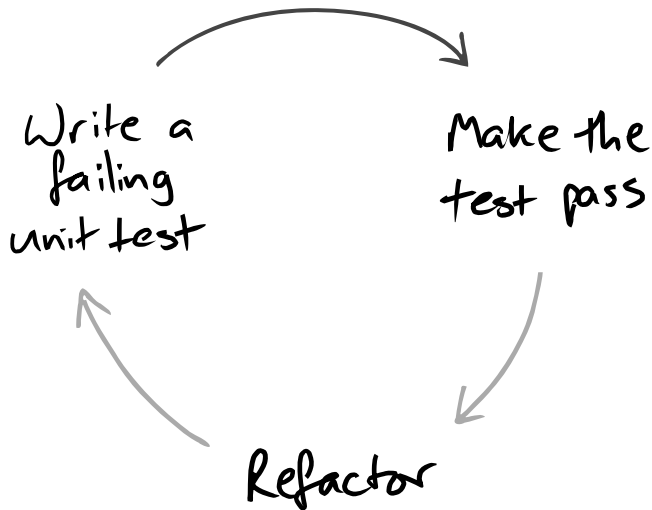
- developers don't like writing tests
- testing is often seen as a 2nd class development activity

TDD idea (i.e. a judo move on the above problem)

write test *before* code

- don't write test only to verify code after it's done
leverage **testing as a design activity**
- write tests to clarify our ideas about what the code should do
- *I was finally able to separate logical from physical design. I'd always been told to do that but no one ever explained how.* — Kent Beck
- write tests to get **rapid feedback about design idea**
 - ▶ if a test is difficult to write, design is often wrong

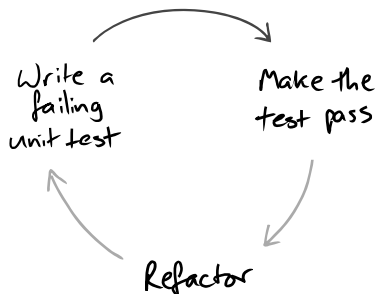
TDD in a nutshell



GOOS, Figure 1.1

The TDD development cycle

- 1 **red**: write a failing test
- 2 **green**: write code that makes it pass
- 3 **refactor** the code to be as simple as possible



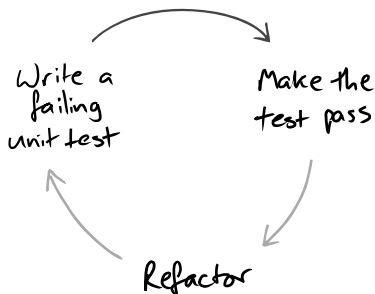
GOOS, Figure 1.1

The TDD development cycle (cont.)

- 1 **red**: write a failing test
- 2 **green**: write code that makes it pass
- 3 **refactor** the code to be as simple as possible

By **writing the test** we:

- clarify **acceptance criteria**
- are pushed to design **loosely coupled components**
 - ▶ otherwise they are difficult to test
- **document** the code, via an executable description of it
- incrementally build a **regression suite**



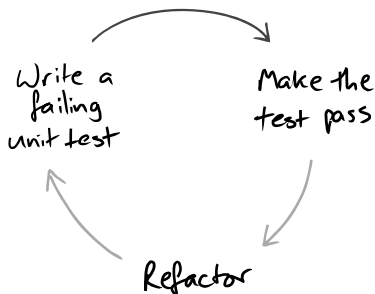
GOOS, Figure 1.1

The TDD development cycle (cont.)

- 1 **red**: write a failing test
- 2 **green**: write code that makes it pass
- 3 **refactor** the code to be as simple as possible

By **running the test** we:

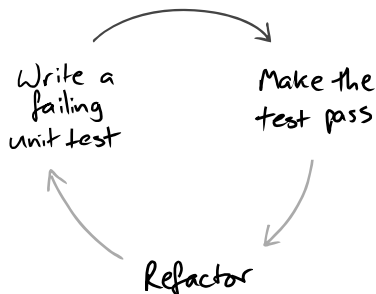
- **detect errors** when the context is fresh in our mind
- have a **measure of progress**, know when to stop (i.e. when we are “done enough”)
 - ▶ avoid over-coding, “gold plating”



GOOS, Figure 1.1

The TDD development cycle (cont.)

- 1 **red**: write a failing test
- 2 **green**: write code that makes it pass
- 3 **refactor** the code to be as simple as possible



GOOS, Figure 1.1

TDD golden rule

Never write a new functionality without a failing test.

Step 1 — write a failing unit test

This looks easy. But requires some care:

- write **the test you'd want to read**, ideally in 3 steps
 - 1 prepare test environment (input data and/or context)
 - 2 invoke the logic under testing
 - 3 verify that the results are correct

If it cannot be done in a few lines (≈ 3), write helper methods.

Remember: **tests are documentation** too.

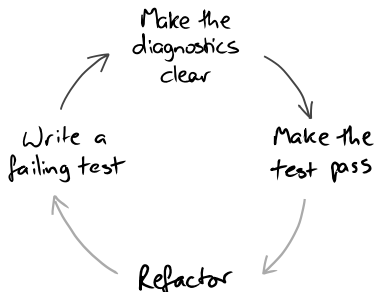
Step 1 — write a failing unit test

This looks easy. But requires some care:

- write **the test you'd want to read**, ideally in 3 steps
 - 1 prepare test environment (input data and/or context)
 - 2 invoke the logic under testing
 - 3 verify that the results are correct

If it cannot be done in a few lines (≈ 3), write helper methods.
Remember: **tests are documentation** too.

- **watch the test fail** *before* making it pass
 - ▶ otherwise you're not sure about your assumptions
 - ▶ if the tests fail in unexpected ways, fix it
 - ▶ if the **diagnostic isn't clear**, fix it



GOOS, Figure 5.2

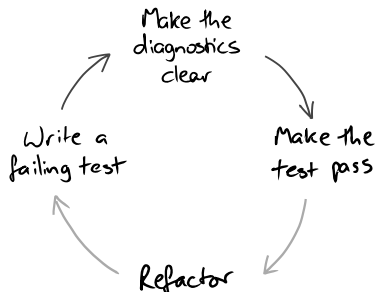
Step 1 — write a failing unit test

This looks easy. But requires some care:

- write **the test you'd want to read**, ideally in 3 steps
 - 1 prepare test environment (input data and/or context)
 - 2 invoke the logic under testing
 - 3 verify that the results are correct

If it cannot be done in a few lines (≈ 3), write helper methods.
Remember: **tests are documentation** too.

- **watch the test fail** *before* making it pass
 - ▶ otherwise you're not sure about your assumptions
 - ▶ if the tests fail in unexpected ways, fix it
 - ▶ if the **diagnostic isn't clear**, fix it
- unit **test behavior**, not methods
 - ▶ for TDD **test coverage** is less important than readable tests



GOOS, Figure 5.2

Step 2 — make the test pass

(you won't like this)

To make the test pass we allow ourselves to **take shortcuts**.

Common strategies to make the test pass:

- **fake it** — all sort of dirty tricks
 - ▶ e.g. return the **constant value** the test expects

To make the test pass we allow ourselves to **take shortcuts**.

Common strategies to make the test pass:

- **fake it** — all sort of dirty tricks
 - ▶ e.g. return the **constant value** the test expects
- **obvious implementation** — just type in the “obviously right” implementation
 - ▶ it takes experience to tune your confidence
 - ▶ too confident: you will have bad surprises
 - ▶ too prudent: you'll fake it too often
 - ▶ tip: use confidence increasingly, **fall back when you get an unexpected “red bar”** (i.e. test failure)

To make the test pass we allow ourselves to **take shortcuts**.

Common strategies to make the test pass:

- **fake it** — all sort of dirty tricks
 - ▶ e.g. return the **constant value** the test expects
- **obvious implementation** — just type in the “obviously right” implementation
 - ▶ it takes experience to tune your confidence
 - ▶ too confident: you will have bad surprises
 - ▶ too prudent: you'll fake it too often
 - ▶ tip: use confidence increasingly, **fall back when you get an unexpected “red bar”** (i.e. test failure)
- **triangulation** — when you have more than 2-3 tests that use the same implementation, factor out a common one
 - ▶ corollary: triangulation is commonly applied after several applications of the previous techniques

Step 3 — refactor

At this point: we have a test, some new code, and we are reasonably convinced that **it is that code that makes the test pass**.

We can now improve the code design, using **tests as a safety net**.

The goal of **refactoring** is to improve the design of existing code, without altering its external behavior (see Fowler 1999). We only give some of its intuitions here:

Code smells

- duplicate code
- long methods / large class
- too many parameters
- inappropriate intimacy
- Liskov principle violation
- complex conditionals
- ...

Techniques

- encapsulate field
- generalize type
- conditionals → polymorphism
- extract class / method
- rename method / field
- pull up / push down
- ...

TDD cycle — example

Goal: get a list of the elements contained in a binary tree

- 1 write a unit test that (pseudo-code)
 - 1 creates a binary tree containing the elements 2, 3, 1
 - 2 invokes a `toList` method on it
 - 3 asserts that `toList`'s return value = `[2; 3; 1]`
- Run all tests and ensures the new one *fails as we expect*
- ▶ i.e. with a compilation error due to the lack of `toList`

TDD cycle — example

Goal: get a list of the elements contained in a binary tree

① write a unit test that (pseudo-code)

- ① creates a binary tree containing the elements 2, 3, 1
- ② invokes a `toList` method on it
- ③ asserts that `toList`'s return value = `[2; 3; 1]`

Run all tests and ensures the new one *fails as we expect*

- ▶ i.e. with a compilation error due to the lack of `toList`

② implement `toList`

- ▶ either by faking it: `return([2; 3; 1])`
- ▶ or by writing the implementation you consider obvious

Run all tests and ensures the new one *succeeds*

TDD cycle — example

Goal: get a list of the elements contained in a binary tree

1 write a unit test that (pseudo-code)

- 1 creates a binary tree containing the elements 2, 3, 1
- 2 invokes a `toList` method on it
- 3 asserts that `toList`'s return value = `[2; 3; 1]`

Run all tests and ensures the new one *fails as we expect*

- ▶ i.e. with a compilation error due to the lack of `toList`

2 implement `toList`

- ▶ either by faking it: `return([2; 3; 1])`
- ▶ or by writing the implementation you consider obvious

Run all tests and ensures the new one *succeeds*

3 refactor, e.g.:

- ▶ write a proper implementation if you faked it
- ▶ clean up, clean up, clean up

Run all tests and ensures the new one it *still succeeds*

TDD cycle — exercise

Exercise (bug fixing work-flow)

You have adopted TDD as the development process for your project. Describe the work-flow you would use to fix a bug.

?

TDD cycle — exercise

Exercise (bug fixing work-flow)

You have adopted TDD as the development process for your project. Describe the work-flow you would use to fix a bug.

Possible work-flow:

- 0 receive **bug report**
- 1 run all tests to ensure **clean slate**
- 2 create a **new test** that
 - ▶ recreates the context of the (alleged) bug
 - ▶ would succeed if the bug didn't exist
- 3 run all tests
 - ▶ new test fails → **reproducible bug**
 - ▶ new test passes → unreproducible bug → investigate with submitter

TDD cycle — exercise (cont.)

Exercise (bug fixing work-flow)

You have adopted TDD as the development process for your project. Describe the work-flow you would use to fix a bug.

Possible work-flow:

- 5 fix the bug
 - ▶ even with dirty workarounds, to ensure you've cornered it
- 6 run all tests
 - ▶ all test passes → bingo!
 - ▶ new test fails → try again
 - ▶ old tests fail → **regression**, try again
- 7 refactor as needed
 - ▶ from workaround, to proper fix
- 8 release fix (including the new test!)

Outline

- 1 Development processes
- 2 Test-Driven Development
- 3 xUnit & jUnit**
- 4 Bibliography

xUnit

xUnit collectively refers to a set of **frameworks for automated unit testing** which share a **common test coding style**.

Each xUnit framework includes:

test case abstraction used to define tests

test suite abstraction used to organize test in test suites

test fixture mechanisms to factorize test initialization and clean up code

test runner end-user program to discover and run test suites, summarizing their results

²https://en.wikipedia.org/wiki/List_of_unit_testing_frameworks

xUnit

xUnit collectively refers to a set of **frameworks for automated unit testing** which share a **common test coding style**.

Each xUnit framework includes:

test case abstraction used to define tests

test suite abstraction used to organize test in test suites

test fixture mechanisms to factorize test initialization and clean up code

test runner end-user program to discover and run test suites, summarizing their results

xUnit frameworks exist for most languages and platforms, e.g.:²

- SUnit (Smalltalk)
- JUnit (Java)
- CppUnit (C++)
- OUnit (OCaml)
- Test::Unit (Ruby)
- HUnit (Haskell)
- NUnit (.NET)
- unittest (Python)
- Check (C)

²https://en.wikipedia.org/wiki/List_of_unit_testing_frameworks

JUnit

JUnit is the original Java port of SUnit by Kent Beck and Erich Gamma.

It's still the most established xUnit Java test framework.

We will JUnit, and in particular **JUnit 4**, for this lecture examples. The notions we will see are portable to other xUnit frameworks.

JUnit is Free Software, released under the Eclipse Public License; it is available at <http://junit.org/> and in most FOSS distributions (package `junit4` in distros of the Debian family).

JUnit — test discovery

```
import org.junit.Test;
import static org.junit.Assert.*;

public class TreeTest { // no inheritance needed

    @Test
    public void emptyTreeCreation() {
        // prepare
        // execute
        // assert
    }

    @Test
    public void treeAddition() {
        // prepare
        // execute
        // assert
    }
}
```


JUnit — assertions

The `junit.framework.Assert` (whose content you should import) provides a plethora of **assertion methods**, e.g.:

```
assertTrue(String msg, Boolean test)
assertFalse(String msg, Boolean test)
assertNull(String msg, Boolean test)
assertEquals(String msg, Object expected, Object actual)
assertSame(String msg, Object expected, Object actual)
    // note: order does matter
assertNot*(...)
fail(String msg) // fail unconditionally
```

Assertion methods also come in msg-less variants.

```
@Test
public void emptyTreeSize() {
    Tree t = new BinaryTree(); // prepare
    int s = t.size();          // execute
    assertEquals(0, s);        // assert
}
```

JUnit — asserting exceptions

Asserting that a test throws an exception is ugly:

- wrap the whole test in **try ... catch**
- do nothing in the **catch** branch
- `fail()` at the end of the **try** branch

JUnit offers a nicer declarative alternative:

```
@Test(expected = IndexOutOfBoundsException.class)
public void arrayListBoundaries() {
    List<String> l = new ArrayList<String>(); // prepare
    l.get(1); // execute
} // test will fail if it reaches this point
```

Test isolation

The running of a test should not influence that of another.

I.e. **test should be isolated**.

Consequences:

- test execution is **order independent**
- to achieve isolation you need to split your system accordingly
 - ▶ separation of concern, low coupling & high cohesion
 - ▶ once again: tests help good design
- each test should initialize its context (**set up**) before execution and clean it up completely (**tear down**) afterwards

But we don't want **duplications in setup/teardown code!**

Test fixtures to the rescue!

JUnit — fixtures

```
import org.junit.Before;
import org.junit.After;
public class ArrayTest {
    private List<String> l;

    @Before // i.e. call this before each @Test method
    public void setUp() {
        l = new ArrayList<String>();
        l.add("foo"); l.add("bar"); l.add("baz"); }

    @After // i.e. call this after each @Test method
    public void tearDown() { l = null; }

    @Test
    public void removeFirst() {
        assertEquals("foo", l.remove(0)); }

    @Test
    public void removeSecond() {
        assertEquals("bar", l.remove(1)); }
}
```

JUnit — test suites

You can (and should) **organize your tests** in **test suites**.

```
import org.junit.runners.Suite;  
import org.junit.runner.RunWith;
```

```
@RunWith(Suite.class)  
@Suite.SuiteClasses({ // list classes containing @Test  
    Behavior1Test.class,  
    Behavior2Test.class,  
    ... ,  
    BehaviorNTest.class,  
})  
public class BehaviorTestSuite { }
```

Unfortunately, there is no nice way to run all available tests (in a package, for example). External test runners (Eclipse, Ant, etc.) use reflection to address this problem.

JUnit — running tests

Test outcome

Each @Test method has a 3-state outcome:

- success** the test criteria have been met (i.e. all assertions were correct)
- failure** the test criteria have not been met (i.e. one or more assertion were incorrect)
- error** test execution didn't complete properly (e.g. an exception *not under test* interrupted test execution)

Running tests will **collect outcomes** and provide a **summary**.

- on the command line:

```
java -cp /usr/share/java/junit4.jar \  
    org.junit.runner.JUnitCore \  
    ClassName
```

- in Eclipse: *Run* → *Run as* → *JUnit Test* Shift+Alt+X T
- JUnit integration available for build tools, e.g. Ant, Maven, etc.

Outline

- 1 Development processes
- 2 Test-Driven Development
- 3 xUnit & junit
- 4 Bibliography**

Bibliography



Steve Freeman and Nat Pryce

*Growing Object-Oriented Software, Guided by Tests*³
Addison-Wesley, 2009.



Kent Beck

Test Driven Development: By Example
Addison-Wesley, 2002.



Martin Fowler

Refactoring: Improving the Design of Existing Code
Addison-Wesley Professional, 1999.



Kent Beck

Simple smalltalk testing: With patterns
The Smalltalk Report 4.2 (1994): 16-18.
available at <http://www.xprogramming.com/testfram.htm>

³referred to as the “GOOS” book in this course slides