Lecture 11:
Refactoring (and TDD)
Schedule of Lectures

Week 01: Introduction to SE
Week 02: Requirements Engineering I
Week 03: Requirements Engineering II
Week 04: Analysis
Week 05: Development Infrastructure
Week 06: Continuous Development and Integration
Week 07: Architecture and Design I
Week 08: Architecture and Design II
Week 09: Verification and Validation I
Week 10: Verification and Validation II
Week 11: Refactoring (and TDD)
Week 12: Agile/Lean Methods
Week 13: Agile Methods in Industry
Week 14: Course wrap-up, review and exam preparation
Week 15: Reserve time slot (no lecture scheduled as of today)
Structure of Lecture 11

• What is Refactoring and why do it?
• Examples
  • Code Smells
  • Refactoring
What is Refactoring?
What is Refactoring?

Refactoring

Maintenance

Evolution
Beck’s Definition

“A change to the system that leaves its behavior unchanged, but enhances some nonfunctional quality – simplicity, flexibility, understandability, performance”

(Kent Beck, *Extreme Programming Explained*, page 179)
Fowler’s Definition

“A change made to the internal structure of software to make it easier to understand and cheaper to modify without changing its observable behavior”

(Martin Fowler, *Refactoring*, page 53)
Refactoring – Comprehensive Definition

Changes made to the system that
• Do not change observable behavior (all the tests still pass)
• Remove duplication or needless complexity
• Enhance software quality:
  • Make the code simpler and easier to understand
  • Make the code more flexible
  • Make the code easier to change
A Trivial Example

 Consolidate Duplicate Conditional Fragments (page 243); This

    if (isSpecialDeal()) {
        total = price * 0.95;
        send();
    } else {
        total = price * 0.98;
        send();
    }

 becomes this

    if (isSpecialDeal()) {
        total = price * 0.95;
    } else {
        total = price * 0.98;
    }
    send();

(Martin Fowler, Refactoring)
A Trivial Example

- Consolidate Duplicate Conditional Fragments (page 243); This

  ```java
  if (isSpecialDeal()) {
      total = price * 0.95;
      send()
  } else {
      total = price * 0.98;
      send()
  }
  ```

  becomes this

  ```java
  if (isSpecialDeal()) {
      total = price * 0.95;
      send()
  } else {
      total = price * 0.98;
  }
  send();
  ```

  Or:

  ```java
  total = price * 0.98;
  if (isSpecialDeal()) {
      total = price * 0.95;
  }
  send();
  ```

(Martin Fowler, *Refactoring*)
Context – Test-Driven Development (TDD)

User Story (new functionality)

Write a failing test.

Clean up code.

Make the test pass.

Test

Refactor

Code
Test-Driven Development (TDD) – Example

Story1: There is a 2% discount for cash payment

Test1: /* check for cash discount */
price = 1;
total = calcTotal(price)
if isCash(){
    assertEquals(total, price*0.98);}

→ Test1 fails

User Story
(new functionality)

Write a failing test.

Clean up code.

Make the test pass.

Refactor

Test

Code
Test-Driven Development (TDD) – Example

Story1: There is a 2% discount for cash payment

Test1: /* check for cash discount */
price = 1;
total = calcTotal(price)
if isCash(){
    assertEquals(total, price*0.98);}

→ Test1 fails

Code1:
... calcTotal(...){
...
if isCash(){
    total = price*0.98;
    send();}{}

→ Test1 passes

Nothing to refactor (yet)
Test-Driven Development (TDD) – Example

Story2: If there is a special deal, cash discount is 5%

Test2: /* check for cash and special deal */

Price = 1;
total = calcTotal(price)
if isCash(){
    if specialDeal(){
        assertEquals(total, price*0.98);
    }
}

Test 2 → fails

Code 2:
if isCash{
    if (isSpecialDeal()) {
        total = price * 0.95;
        send();
    } else {
        total = price * 0.98;
        send();
    }
}

→ Test2 passes (Does Test1 still pass?)
Test-Driven Development (TDD) – Example

Code 2:
If isCash{
    if (isSpecialDeal()) {
        total = price * 0.95;
        send();
    } else {
        total = price * 0.98;
        send();
    }
}

→ Test2 passes (and also Test1 passes)

Refactoring → Code3:
if isCash() {
    total = price * 0.98;
    if (isSpecialDeal()) {
        total = price * 0.95;
    }
    send();
}

→ Test1 & Test2 pass
Why Refactor?

- Prevent “design decay”
- Clean up messes in the code
- Simplify the code
- Increase readability and understandability
- Find bugs (easier)
- Reduce debugging time
- Build in learning we do about the application
- Redoing things is fundamental to every creative process
When to Refactor?

• “All the time”
• When you duplicate code ➔ Rule of Three
  • Rule of thumb to decide when a duplicated piece of code should be replaced by a new method: code can be copied once, but when the same code is used three times, it should be extracted into a new method.
• When you learn/understand something new about the code
• (When you add functionality)
• (When you fix a bug)
• When the code smells
When **not to** Refactor?

- **When the tests aren’t passing**
- When you should simply rewrite the code from scratch
- When you have impeding deadlines
  - Downside: unfinished refactoring becomes a technical debt

**Potential problems:**

Taken too far, refactoring can lead to incessant tinkering with the code, trying to make it perfect.

Refactoring code when the tests don’t work or tests when the application doesn’t work leads to potentially dangerous situations.

Databases can be difficult to refactor

Refactoring published interfaces can cause problems for the code that uses those interfaces
What is the Refactoring Process?

- Make sure your tests pass
- Find some code that “smells”
- Determine how to simplify this code
- Make the simplifications
- Run tests to ensure things still work correctly
- Repeat the simplify/test cycle until the smell is gone
What is the Refactoring Process?

- Make sure your tests pass
- Find some code that “smells”
- Determine how to simplify this code
- Make the simplifications
- Run tests to ensure things still work correctly
- Repeat the simplify/test cycle until the smell is gone

When does it smell?

How?
Structure of Lecture 11

• What is Refactoring and why do it?
• Examples
  • Code Smells
  • Refactoring
Code Smells

- Indicators that something may be wrong in the code
- Can occur both in production code and test code
Code Smells (Fowler)

Alternative Classes with Different Interfaces
Comments
Data Class
Data Clumps
Divergent Change
Duplicated Code
Feature Envy
Inappropriate Intimacy
Incomplete Library Class
Large Class
Lazy Class

Long Method
Long Parameter List
Message Chains
Middle Man
Parallel Inheritance Hierarchies
Primitive Obsession
Refused Bequest
Shotgun Surgery
Speculative Generality
Switch Statements
Temporary Field
Code Smells (Fowler)

- **Duplicated Code**
  - bad because if you modify one instance of duplicated code but not the others, you (may) have introduced a bug!

- **Long Method**
  - long methods are more difficult to understand
    - performance concerns with respect to lots of short methods are largely obsolete
Code Smells (Fowler)

- Large Class
  - Large classes try to do too much, which reduces cohesion
- Long Parameter List
  - Hard to understand, can become inconsistent
- Divergent Change
  - Related to cohesion
  - Symptom: one type of change requires changing one subset of methods; another type of change requires changing another subset
Code Smells (Fowler)

- **Shotgun Surgery**
  - a change requires lots of little changes in a lot of different classes

- **Feature Envy**
  - A method requires lots of information from some other class
  - move it closer!

- **Data Clumps**
  - attributes that clump together (are used together) but are not part of the same class
Shotgun Surgery – Example: Add logging functionality & then change it

```c
void MyFunc()
{
    ... 
}

void MyFunc2()
{
    ... 
}

... 

void MyFuncN()
{
    ... 
}
```

```c
void MyFunc()
{
    printf("Entering MyFunc()\n");
    ...
}

void MyFunc2()
{
    printf("Entering MyFunc2()\n");
    ...
}

... 

void MyFuncN()
{
    printf("Entering MyFuncN()\n");
    ...
}
```
Code Smells (Fowler)

- **Primitive Obsession**
  - characterized by a reluctance to use classes instead of primitive data types

- **Switch Statements**
  - Switch statements are often duplicated in code; they can typically be replaced by use of polymorphism (let OO do your selection for you!)

- **Parallel Inheritance Hierarchies**
  - Similar to Shotgun Surgery; each time I add a subclass to one hierarchy, I need to do it for all related hierarchies
  - Note: some design patterns encourage the creation of parallel inheritance hierarchies (so they are not always bad!)
    (e.g., Façade design pattern → used to simplify an interface)
Code Smells (Fowler)

- **Lazy Class**
  - A class that no longer “pays its way”
  - e.g. may be a class that was downsized by a previous refactoring, or represented planned functionality that did not pan out

- **Speculative Generality**
  - “Oh I think we need the ability to do this kind of thing someday”

- **Temporary Field**
  - An attribute of an object is only set in certain circumstances; but an object should need all of its attributes
Code Smells (Fowler)

- **Message Chains**
  - a client asks an object for another object and then asks that object for another object etc. Bad because client depends on the structure of the navigation

- **Middle Man**
  - If a class is delegating more than half of its responsibilities to another class, do you really need it? (involves trade-offs, some design patterns encourage this (e.g. Decorator))

- **Inappropriate Intimacy**
  - Pairs of classes that know too much about each other’s private details (loss of encapsulation; change one class, the other has to change)
Code Smells (Fowler)

- **Data Class**
  - These are classes that have fields, getting and setting methods for the fields, and nothing else; they are data holders, but objects should be about data AND behavior

- **Refused Bequest**
  - A subclass ignores most of the functionality provided by its superclass
  - Subclass may not pass the “IS-A” test

- **Comments (!)**
  - Comments are sometimes used to hide bad code
  - “…comments often are used as a deodorant” (!)
Test Code Smells (van Deursen et al.)

- Mystery Guest
- Resource Optimism
- Test Run War
- General Fixture
- Eager Test
- Lazy Test
- Assertion Roulette
- Indirect Testing
- For Testers Only
- Sensitive Equality
- Test Code Duplication

“Refactoring Test Code” by Arie van Deursen, Leon Moonen, Alex van den Bergh, and Gerard Kok, published in Extreme Programming Perspectives
Structure of Lecture 11

- What is Refactoring and why do it?
- Examples
  - Code Smells
  - Refactoring
Signs of Good OO Design

• Short methods
  • Simple method logic

• Few instance variables

• Clear object responsibilities
  • State the purpose of the class in one sentence
  • No super-intelligent objects
The Catalogue of Refactorings (1)

- Add Parameter
- Change Bidirectional Association to Unidirectional
- Change Reference to Value
- Change Unidirectional Association to Bidirectional
- Change Value to Reference
- Collapse Hierarchy
- Consolidate Conditional Expression
- Consolidate Duplicate Conditional Fragments
- Decompose Conditional
- Duplicate Observed Data
- Dynamic Method Definition
- Eagerly Initialized Attribute
- Encapsulate Collection
- Encapsulate Downcast
- Encapsulate Field
- Extract Class
- Extract Interface
- Extract Method
- Extract Module
- Extract Subclass
- Extract Superclass
- Extract Surrounding Method
- Extract Variable

- Form Template Method
- Hide Delegate
- Hide Method
- Inline Class
- Inline Method
- Inline Module
- Inline Temp
- Introduce Assertion
- Introduce Class Annotation
- Introduce Expression Builder
- Introduce Foreign Method
- Introduce Gateway
- Introduce Local Extension
- Introduce Named Parameter
- Introduce Null Object
- Introduce Parameter Object
- Isolate Dynamic Receptor
- Lazily Initialized Attribute
- Move Eval from Runtime to Parse Time
- Move Field
- Move Method
- Parameterize Method
- Preserve Whole Object
The Catalogue of Refactorings (2)

- Pull Up Constructor Body
- Pull Up Field
- Pull Up Method
- Push Down Field
- Push Down Method
- Recompose Conditional
- Remove Assignments to Parameters
- Remove Control Flag
- Remove Middle Man
- Remove Named Parameter
- Remove Parameter
- Remove Setting Method
- Remove Unused Default Parameter
- Rename Method
- Replace Abstract Superclass with Module
- Replace Array with Object
- Replace Conditional with Polymorphism
- Replace Constructor with Factory Method
- Replace Data Value with Object
- Replace Delegation With Hierarchy
- Replace Delegation with Inheritance
- Replace Dynamic Receptor with Dynamic Method Definition
- Replace Error Code with Exception
- Replace Exception with Test
- Replace Hash with Object
- Replace Inheritance with Delegation
- Replace Loop with Collection Closure Method
- Replace Magic Number with Symbolic Constant
- Replace Method with Method Object
- Replace Nested Conditional with Guard Clauses
- Replace Parameter with Explicit Methods
- Replace Parameter with Method
- Replace Record with Data Class
- Replace Subclass with Fields
- Replace Temp with Chain
- Replace Temp with Query
- Replace Type Code with Class
- Replace Type Code with Module Extension
- Replace Type Code With Polymorphism
- Replace Type Code with State/Strategy
- Replace Type Code with Subclasses
- Self Encapsulate Field
- Separate Query from Modifier
- Split Temporary Variable
- Substitute Algorithm
IDEs offer very incomplete suggestions of refactorings and don’t recommend sequences of refactorings … and if they do, give little justification for what they suggest.
Code Smells & Refactoring

- Full document posted on course wiki

(code smells and refactoring guide table)

(list shown here is not complete)
## Typical Refactorings

<table>
<thead>
<tr>
<th>Class Refactorings</th>
<th>Method Refactorings</th>
<th>Attribute Refactorings</th>
</tr>
</thead>
<tbody>
<tr>
<td>add (sub)class to hierarchy</td>
<td>add method to class</td>
<td>add variable to class</td>
</tr>
<tr>
<td>rename class</td>
<td>rename method</td>
<td>rename variable</td>
</tr>
<tr>
<td>remove class</td>
<td>remove method</td>
<td>remove variable</td>
</tr>
<tr>
<td>extract class</td>
<td>push method down</td>
<td>push variable down</td>
</tr>
<tr>
<td></td>
<td>push method up</td>
<td>pull variable up</td>
</tr>
<tr>
<td></td>
<td>add parameter to method</td>
<td>create accessors</td>
</tr>
<tr>
<td></td>
<td>move method</td>
<td>abstract variable</td>
</tr>
<tr>
<td></td>
<td>extract code in new method</td>
<td></td>
</tr>
<tr>
<td></td>
<td>replace parameter with method</td>
<td></td>
</tr>
</tbody>
</table>
Push Method Down

- Behavior on a superclass is relevant only for some of its subclasses.
- Move it to those subclasses.
Push Method Up

- You have methods with identical results on subclasses.
- Move them to the superclass
Add Parameter to Method

• A method needs more information from its caller.
• Add a parameter for an object that can pass on this information.
Replace **switch** Statement with Polymorphism

- **switch** statements are very rare in properly designed object-oriented code
  - Therefore, a **switch** statement is a simple and easily detected “bad smell”
  - Of course, not all uses of **switch** are bad
  - A switch statement should *not* be used to distinguish between various types of objects (object types) and *not* rely on data from other objects
- Refactoring options for this case exist
  - The simplest is the creation of subclasses (exploiting polymorphism)
Encapsulate Field

- Un-encapsulated data is a no-no in OO application design.
- Use properly get and set procedures to provide public access to private (encapsulated) member variables.

```java
public class Course {
    private List students;
    public List getStudents() {
        return students;
    }
    public void setStudents(List s) {
        students = s;
    }
}

int classSize = course.students.size();
```

```java
public class Course {
    private List students;
    public List getStudents() {
        return students;
    }
    public void setStudents(List s) {
        students = s;
    }
}

int classSize = course.getStudents().size();
```
Extract Class

• Break one class into two, e.g. having the phone details as part of the Customer class is not a realistic OO model, and also breaks the Single Responsibility design principle.

• We can refactor this into two separate classes, each with the appropriate responsibility.

```java
public class Customer {
    private String name;
    private String workPhoneAreaCode;
    private String workPhoneNumber;
}

public class Phone {
    private String areaCode;
    private String number;
}
```
Extract Method

• Sometimes we have methods that do too much. The more code in a single method, the harder it is to understand and get right. It also means that logic embedded in that method cannot be reused elsewhere.

• The Extract Method refactoring is one of the most useful for reducing the amount of duplication in code.

```java
public class Customer {
    void int foo() {
        …
        // Compute score
        score = a*b+c;
        score *= xfactor;
    }
}
```

```java
public class Customer {
    void int foo() {
        …
        score = ComputeScore(a,b,c,xfactor);
    }
    int ComputeScore(int a, int b, int c, int x) {
        return (a*b+c)*x;
    }
}
```
Extract Sub-Class

• When a class has features (attributes and methods) that would only be useful in specialized instances, we can create a specialization of that class and give it those features. This makes the original class less specialized (i.e., more abstract), and good design is about binding to abstractions wherever possible.

```java
public class Person {
    private String name;
    private String jobTitle;
}

public class Person {
    protected String name;
}

public class Employee extends Person {
    private String jobTitle;
}
```
Extract Super-Class

• When you find two or more classes that share common features, consider abstracting those shared features into a super-class.

• Again, this makes it easier to bind clients to an abstraction, and removes duplicate code from the original classes.

```java
public class Employee {
  private String name;
  private String jobTitle;
}

gclass Student {
  private String name;
  private Course course;
}
```

```java
public abstract class Person {
  protected String name;
}

gclass Employee extends Person {
  private String jobTitle;
}

gclass Student extends Person {
  private Course course;
}
```
Move Method – Before

• If a method in one class uses (or is used by) another class more than the class in which its defined, move it to the other class

```java
public class Student {
    public boolean isTaking(Course course) {
        return (course.getStudents().contains(this));
    }
}

public class Course {
    private List students;
    public List getStudents() {
        return students;
    }
}
```
Move Method – Refactored

• The student class now no longer needs to know about the Course interface, and the isTaking() method is closer to the data on which it relies - making the design of Course more cohesive and the overall design more loosely coupled

```java
public class Student
{
}

public class Course
{
    private List students;
    public boolean isTaking(Student student)
    {
        return students.contains(student);
    }
}
```
Introduce Null Object

- If relying on null for default behavior, use inheritance instead
- Replace the null value with a null object

```java
public class User {
    Plan getPlan() {
        return plan;
    }
}
```

```java
if (user == null)
    plan = Plan.basic();
else
    plan = user.getPlan();
```

```java
public class NullUser extends User {
    Plan getPlan() {
        return Plan.basic();
    }
}
```
Replace Error Code with Exception

- A method returns a special code to indicate an error is better accomplished with an Exception.

```java
int withdraw(int amount)
{
    if (amount > balance)
        return -1;
    else {
        balance -= amount;
        return 0;
    }
}
```

```java
void withdraw(int amount)
    throws BalanceException
{
    if (amount > balance)
    {
        throw new BalanceException();
    }
    balance -= amount;
}
```
Replace Exception with Test

- Conversely, if you are catching an exception that could be handled by an if-statement, use that instead.

```java
double getValueForPeriod (int periodNumber) {
    try {
        return values[periodNumber];
    } catch (ArrayIndexOutOfBoundsException e) {
        return 0;
    }
}
```

```java
double getValueForPeriod (int periodNumber) {
    if (periodNumber >= values.length) return 0;
    return values[periodNumber];
}
```
Nested Conditional with Guard

• A method has conditional behavior that does not make clear what the normal path of execution is. Use Guard Clauses for all the special cases.

```java
double getPayAmount() {
    double result;
    if (isDead) result = deadAmount();
    else {
        if (isSeparated) result = separatedAmount();
        else {
            if (isRetired) result = retiredAmount();
            else result = normalPayAmount();
        }
    }
    return result;
}
```

```java
double getPayAmount() {
    if (isDead) return deadAmount();
    if (isSeparated) return separatedAmount();
    if (isRetired) return retiredAmount();
    return normalPayAmount();
}
```
Replace Parameter with Explicit Method

- You have a method that runs different code depending on the values of an enumerated parameter. Create a separate method for each value of the parameter.

```java
void setValue (String name, int value) {
    if (name.equals("height")) {
        height = value;
        return;
    }
    if (name.equals("width")) {
        width = value;
        return;
    }
    Assert.shouldNeverReachHere();
}
```

void setHeight(int arg) {
    height = arg;
}

void setWidth (int arg) {
    width = arg;
}
Replace Temp with Query

• You are using a temporary variable to hold the result of an expression. Extract the expression into a method. Replace all references to the temp with the expression. The new method can then be used in other methods and allows for other refactorings.

```java
double basePrice = quantity * itemPrice;
if (basePrice > 1000)
    return basePrice * 0.95;
else
    return basePrice * 0.98;
```

```java
if (basePrice() > 1000)
    return basePrice() * 0.95;
else
    return basePrice() * 0.98;
...

    double basePrice() {
        return quantity * itemPrice;
    }
```
Rename Variable or Method

- Perhaps one of the simplest, but one of the most useful that bears repeating: If the name of a method or variable does not reveal its purpose then change the name of the method or variable.

```java
public class Customer {
    public double getinvcdtlmt();
}
```

```java
public class Customer {
    public double getInvoiceCreditLimit();
}
```
More on Refactoring

• Refactoring Catalog
  • http://www.refactoring.com/catalog

• Java Refactoring Tools
  • NetBeans 4+ – Built In
  • JFactor – works with VisualAge and JBuilder
  • RefactorIt – plug-in tool for NetBeans, Forte, JBuilder and JDeveloper. Also works standalone.
  • JRefactory – for jEdit, NetBeans, JBuilder or standalone

• Eclipse & Visual Studio 2005+ & IntelliJ IDEA
  • Refactoring Built In
    • Encapsulate Field, Extract Method, Extract Interface, Reorder Parameters, Remove Parameter, Promote Local Var to Parameter, more.
Further Reading

• Martin Fowler: Refactoring, 2005
  - Refactoring: Improving the Design of Existing Code (2nd Edition)
  - Signature Series (Fowler))  2nd Edition
  - by Martin Fowler ~ (Author)
  - New Edition in Dec. 2018

• http://refactoring.com/catalog/
Hands-On Example

by

Martin Fowler

(separate PDF on course wiki -- videos)

For self study!
Industry Guest Lecture – Nov 29

"How an idea gets shipped"
by Karl-Aksel Puulmann and Uku Patta
(eAgronom.com)
Industry Guest Lecture – Nov 29

Managers
Farmers
Workers

Farm management simplified

We at eAgronom know that running a farm can be complicated.

Daily tasks include all from every-day farm chores on the field to buying seeds and products, arranging crop sales and managing workers, following government rules and dealing with unexpected situations.

eAgronom provides you with the tools you need to run your farm by delivering complex data in an easy to understand format:

- Have a complete overview of your farm
- Plan, monitor and analyse everyday farm activities
- Manage and onboard workers
Next Lecture

• Date/Time:
  • Friday, 22-Nov, 10:15-12:00

• Topic:
  • Agile/Lean Methods

• Labs:
  • Complete and submit Homework 5
  • Go to next week’s labs: introduction to Homework 6