

Refactoring To Patterns

version 0.17



Joshua Kerievsky
joshua@industriallogic.com
Industrial Logic, Inc.
<http://industriallogic.com>

Table of Contents

<i>Introduction</i>	<i>7</i>
<i>Introductory Chapters</i>	<i>14</i>
<i>Book History</i>	<i>15</i>
<i>Patterns Happy</i>	<i>15</i>
<i>Small Steps</i>	<i>16</i>
<i>Mechanics Matter</i>	<i>17</i>
<i>When I Don't Refactor To Patterns</i>	<i>17</i>
<i>Code Smells</i>	<i>18</i>
<i>Chain Constructors</i>	<i>19</i>
Motivation	20
Mechanics	20
Example	20
Chaining To An Init Method	21
<i>Replace Multiple Constructors with Creation Methods</i>	<i>22</i>
Motivation	22
Mechanics	23
Example	24
Parameterized Creation Methods	26
<i>Encapsulate Classes with Creation Methods</i>	<i>27</i>
Motivation	28
Forces	29
Mechanics	29
Example	29
Encapsulating Inner Classes	31
<i>Extract Creation Class</i>	<i>34</i>
Motivation	35
Mechanics	35
Example	36

<i>Move Object Composition to Creation Method</i>	38
Motivation	38
Prerequisites	38
Mechanics	38
Example	38
<i>Replace Multiple Instances with Singleton</i>	39
<i>Replace Singleton with Object Reference</i>	40
Motivation	40
<i>Replace Singleton with Registry</i>	41
<i>Introduce Polymorphic Creation with Factory Method</i>	43
Motivation	44
Mechanics	45
Example	46
Example 2: Duplication Across Subclasses	48
<i>Replace Conditional Calculations with Strategy</i>	50
Motivation	51
Mechanics	52
Example	53
<i>Replace Implicit Tree with Composite</i>	60
Motivation	60
Mechanics	61
Example	61
<i>Encapsulate Composite with Builder</i>	64
Motivation	64
Mechanics	65
Example	65
Extended Example	68
<i>Form Superset Interface</i>	70
Motivation	71
Mechanics	71
Example	72

<i>Move Embellishment to Decorator</i>	73
Motivation	74
Mechanics	75
Example	77
<i>Move Protection to Proxy</i>	86
Decorated Collections	86
<i>Replace Hard-Coded Notifications with Observer</i>	88
Motivation	89
Mechanics	90
Example	90
<i>Move Accumulation to Collecting Parameter</i>	94
Motivation	94
Mechanics	95
Example	95
JUnit's Collecting Parameter	98
<i>Replace Conditional Searches with Specification</i>	99
Motivation	100
Mechanics	101
Example	102
<i>Replace One/Many Distinctions with Composite</i>	111
Motivation	112
Mechanics	112
Example	113
<i>Compose Method</i>	119
Motivation	120
Mechanics	120
Example 1	121
Example 2	125
Example 3	132
<i>Separate Versions with Adapters</i>	134
Motivation	135

Mechanics	136
Example	136
Adapting with Anonymous Inner Classes	141
Adapting Legacy Systems	141
<i>Adapt Interface</i>	143
Motivation	143
Mechanics	144
Example	144
<i>Replace Enum with Type-Safe Enum</i>	146
Motivation	147
Mechanics	148
Example	148
<i>Replace State-Altering Conditionals with State</i>	154
Motivation	155
Mechanics	157
Example	158
Inner States	167
<i>Replace Invariable Behavior with Template Method</i>	169
<i>Extract Composite</i>	170
Motivation	171
Mechanics	172
Example	173
<i>Replace Subclasses with Visitor</i>	179
Motivation	179
Mechanics	182
Example 1	183
<i>Move Rendering To Visitor</i>	192
<i>Replace Multiple Instantiations with Flyweight</i>	193
Motivation	193
Mechanics	193
Example	193

<i>Replace Singleton with Constant</i>	194
Motivation	194
Mechanics	194
Example	194
<i>References</i>	<i>195</i>
<i>Appendix A – Naming Conventions</i>	<i>196</i>
<i>Appendix B – Loan Terminology</i>	<i>197</i>
<i>Conclusion</i>	<i>198</i>
<i>Acknowledgements</i>	<i>198</i>

Introduction

Patterns are a cornerstone of object-oriented design, while test-first programming and merciless refactoring are cornerstones of evolutionary design. To stop over- or under-engineering, it's necessary to learn how patterns fit into the new, evolutionary rhythm of software development.

The great thing about software patterns is that they convey many useful design ideas. It follows, therefore, that if you learn a bunch of these patterns, you'll be a pretty good software designer, right? I considered myself just that once I'd learned and used dozens of patterns. They helped me develop flexible frameworks and build robust and extensible software systems. After a couple of years, however, I discovered that my knowledge of patterns and the way I used them frequently led me to over-engineer my work.

Once my design skills had improved, I found myself using patterns in a different way: I began refactoring to patterns, instead of using them for up-front design or introducing them too early into my code. My new way of working with patterns emerged from my adoption of Extreme Programming design practices, which helped me avoid both over- and under-engineering.

Zapping Productivity

When you make your code more flexible or sophisticated than it needs to be, you over-engineer it. Some do this because they believe they know their system's future requirements. They reason that it's best to make a design more flexible or sophisticated today, so it can accommodate the needs of tomorrow. That sounds reasonable, if you happen to be a psychic.

But if your predictions are wrong, you waste precious time and money. It's not uncommon to spend days or weeks fine-tuning an overly flexible or unnecessarily sophisticated software design—leaving you with less time to add new behavior or remove defects from a system.

What typically happens with code you produce in anticipation of needs that never materialize? It doesn't get removed, because it's inconvenient to do so, or because you expect that one day the code will be needed. Regardless of the reason, as overly flexible or unnecessarily sophisticated code accumulates, you and the rest of the programmers on your team, especially new members, must operate within a code base that's bigger and more complicated than it needs to be.

To compensate for this, folks decide to work in discrete areas of the system. This seems to make their jobs easier, but it has the unpleasant side effect of generating copious amounts of duplicate code, since everyone works in his or her own comfortable area of the system, rarely seeking elsewhere for code that already does what he or she needs.

Over-engineered code affects productivity because when someone inherits an over-engineered design, they must spend time learning the nuances of that design before they can comfortably extend or maintain it.

Over-engineering tends to happen quietly: Many architects and programmers aren't even aware they do it. And while their organizations may discern a decline in team productivity, few know that over-engineering is playing a role in the problem.

Perhaps the main reason programmers over-engineer is that they don't want to get stuck with a bad design. A bad design has a way of weaving its way so deeply into code that improving it becomes an enormous challenge. I've been there, and that's why up-front design with patterns appealed to me so much.

The Patterns Panacea

When I first began learning patterns, they represented a flexible, sophisticated and even elegant way of doing object-oriented design that I very much wanted to master. After thoroughly studying the patterns, I used them to improve systems I'd already built and to formulate designs for systems I was about to build. Since the results of these efforts were promising, I was sure I was on the right path.

But over time, the power of patterns led me to lose sight of simpler ways of writing code. After learning that there were two or three different ways to do a calculation, I'd immediately race toward implementing the Strategy pattern, when, in fact, a simple conditional expression would have been simpler and faster to program—a perfectly sufficient solution.

On one occasion, my preoccupation with patterns became quite apparent. I was pair programming, and my pair and I had written a class that implemented Java's `TreeModel` interface in order to display a graph of `Spec` objects in a tree widget. Our code worked, but the tree widget was displaying each `Spec` by calling its `toString()` method, which didn't return the `Spec` information we wanted. We couldn't change `Spec`'s `toString()` method since other parts of the system relied on its contents. So we reflected on how to proceed. As was my habit, I considered which patterns could help. The Decorator pattern came to mind, and I suggested that we use it to wrap `Spec` with an object that could override the `toString()` method. My partner's response to this suggestion surprised me. "Using a Decorator here would be like applying a sledgehammer to the problem when a few light taps with a small hammer would do." His solution was to create a small class called `NodeDisplay`, whose constructor took a `Spec` instance, and whose one public method, `toString()`, obtained the correct display information from the `Spec` instance. `NodeDisplay` took no time to program, since it was less than 10 simple lines of code. My Decorator solution would have involved creating over 50 lines of code, with many repetitive delegation calls to the `Spec` instance.

Experiences like this made me aware that I needed to stop thinking so much about patterns and refocus on writing small, simple, straightforward code. I was at a crossroads: I'd worked hard to learn patterns to become a better software designer, but now I needed to relax my reliance on them in order to become truly better.

Going Too Fast

Improving also meant learning to not under-engineer. Under-engineering is far more common than over-engineering. We under-engineer when we become exclusively focused on quickly adding more and more behavior to a system without regard for improving its design along the way. Many programmers work this way—I know I sure have. You get code working, move on to other tasks and never make time to improve the code you wrote. Of course, you'd love to have time to improve your code, but you either don't get around to it, or you listen to managers or customers who say we'll all be more competitive and successful if we simply don't fix what ain't broke.

That advice, unfortunately, doesn't work so well with respect to software. It leads to the "fast, slow, slower" rhythm of software development, which goes something like this:

1. You quickly deliver release 1.0 of a system, but with junky code.
2. You attempt to deliver release 2.0 of the system, but the junky code slows you down.
3. As you attempt to deliver future releases, you go slower and slower as the junky code multiplies, until people lose faith in the system, the programmers and even the process that got everyone into this position.

That kind of experience is far too common in our industry. It makes organizations less competitive than they could be. Fortunately, there is a better way.

Socratic Development

Test-first programming and merciless refactoring, two of the many excellent Extreme Programming practices, dramatically improved the way I build software. I found that these two practices have helped me and the organizations I've worked for spend less time over-engineering and under-engineering, and more time designing just what we need: well-built systems, produced on time.

Test-first programming enables the efficient evolution of working code by turning programming into what Kent Beck once likened to a Socratic dialogue: Write test code to ask your system a question, write system code to respond to the question and keep the dialogue going until you've programmed what you need. This rhythm of programming put my head in a different place. Instead of thinking about a design that would work for every nuance of a system, test-first programming enabled me to make a primitive piece of behavior work correctly before evolving it to the next necessary level of sophistication.

Merciless refactoring is an integral part of this evolutionary design process. A refactoring is a "behavior-preserving transformation," or, as Martin Fowler defined it, "a change made to the internal structure of software to make it easier to understand and cheaper to modify without changing its observable behavior." [Fowler, *Refactoring: Improving the Design of Existing Code* (Addison-Wesley, 1999)].

Merciless refactoring resembles the way Socrates continually helped dialogue participants improve their answers to his questions by weeding out inessentials, clarifying ambiguities and consolidating ideas. When you mercilessly refactor, you relentlessly poke and prod your code to remove duplication, clarify and simplify.

The trick to merciless refactoring is to not schedule time to make small design improvements, but to make them *whenever* your code needs them. The resulting quality of your code will enable you to sustain a healthy pace of development. Martin Fowler et al.'s book, *Refactoring: Improving the Design of Existing Code* (Addison-Wesley, 1999), documents a rich catalog of refactorings, each of which identifies a common need for an improvement and the steps for making that improvement.

Why Refactor To Patterns?

On various projects, I've observed what and how my colleagues and I refactor. While we use many of the refactorings described in Fowler's book, we also find places where patterns can help us improve our designs. At such times, we refactor to patterns, being careful not to produce overly flexible or unnecessarily sophisticated solutions.

When I explored the motivation for refactoring to patterns, I found that it was identical to the motivation for implementing non-patterns-based refactorings: to reduce or remove duplication, simplify the unsimple and make our code better at communicating its intention.

However, the motivation for refactoring to patterns is not the primary motivation for using patterns that is documented in the patterns literature. For example, let's look at the documented Intent and Applicability of the Decorator pattern and then examine Erich Gamma and Kent Beck's motivation for refactoring to Decorator in their excellent, patterns-dense testing framework, JUnit.

Decorator's Intent [*Design Patterns*, page 175]:

Attach additional responsibilities to an object dynamically. Decorators provide a flexible alternative to subclassing for extending functionality.

Decorator's Applicability (GoF, page 177):

- To add responsibilities to individual objects dynamically and transparently, that is, without affecting other objects.
- For responsibilities that can be withdrawn.
- When extension by subclassing is impractical. Sometimes a large number of independent extensions are possible and could produce an explosion of subclasses to support every combination, or a class definition may be hidden or otherwise unavailable for subclassing.

Motivation for Refactoring to Decorator in JUnit

Erich remembered the following reason for refactoring to Decorator:

“Someone added TestSetup support as a subclass of TestSuite, and once we added RepeatedTestCase and ActiveTestCase, we saw that we could reduce code duplication by introducing the TestSetup , Decorator.” [private email]

Can you see how the motivation for refactoring to Decorator (reducing code duplication) had very little connection with Decorator's Intent or Applicability (a dynamic alternative to subclassing)? I noticed similar disconnects when I looked at motivations for refactorings to other patterns. Consider these examples:

Pattern	Intent (GoF)	Refactoring Motivations
Builder	Separate the construction of a complex object from its representation so that the same construction process can create different representations.	Simplify code Remove duplication Reduce creation errors
Factory Method	Define an interface for creating an object, but let the subclasses decide which class to instantiate. The Factory method lets a class defer instantiation to subclasses.	Remove duplication Communicate intent
Template Method	Define the skeleton of an algorithm in an operation, deferring some steps to client subclasses. Template Method lets subclasses redefine certain steps of an algorithm without changing the algorithm's structure.	Remove duplication

Based on these observations, I began to document a catalog of refactorings to patterns to illustrate when it makes sense to make design improvements with patterns. For this work, it's essential to show refactorings from real-world projects in order to accurately describe the kinds of forces that lead to justifiable transformations to a pattern.

My work on refactoring to patterns is a direct continuation of work that Martin Fowler began in his excellent catalog of refactorings, in which he included the following refactorings to patterns:

- Form Template Method (345)
- Introduce Null Object (260)
- Replace Constructor with Factory Method (304)
- Replace Type Code with State/Strategy (227)
- Duplicate Observed Data (189)

Fowler also noted the following:

There is a natural relation between patterns and refactorings. Patterns are where you want to be; refactorings are ways to get there from somewhere else. Fowler, *Refactoring: Improving the Design of Existing Code* (Addison-Wesley, 1999)

This idea agrees with the observation made by the four authors of the classic book, *Design Patterns: Elements of Reusable Object-Oriented Software* (Addison-Wesley, 1994):

Our design patterns capture many of the structures that result from refactoring. ... Design patterns thus provide targets for your refactorings.

Evolutionary Design

Today, after having become quite familiar with patterns, the “structures that result from refactoring,” I know that understanding good reasons to refactor to a pattern are more valuable than understanding the end result of a pattern or the nuances of implementing that end result.

If you'd like to become a better software designer, studying the evolution of great software designs will be more valuable than studying the great designs themselves. For it is in the evolution that the real wisdom lies. The structures that result from the evolution can help you, but without knowing why they were evolved into a design, you're more likely to misapply them or over-engineer with them on your next project.

To date, our software design literature has focused more on teaching great solutions than teaching evolutions to great solutions. We need to change that. As the great poet Goethe said, “That which thy fathers have bequeathed to thee, earn it anew if thou wouldst possess it.” The refactoring literature is helping us reacquire a better understanding of good design solutions by revealing sensible evolutions to those solutions.

If we want to get the most out of patterns, we must do the same thing: See patterns in the context of refactorings, not just as reusable elements existing apart from the refactoring literature. This is perhaps my primary motivation for producing a catalog of refactorings to patterns.

By learning to evolve your designs, you can become a better software designer and reduce the amount of work you over- or under-engineer. Test-first programming and merciless refactoring

are the key practices of evolutionary design. Instill refactoring to patterns in your knowledge of refactorings and you'll find yourself even better equipped to evolve great designs.

Writing Goals

At present, I've written more than a dozen refactorings and have many more in the works. My goal in writing this is to help you learn how to

- refactor *to Patterns* when appropriate and *away from Patterns* when something simpler is discovered
- use Patterns to communicate intention
- know and continue to learn a large body of Patterns
- understand how to implement Patterns in simple and sophisticated ways
- use Patterns to clean, condense, clarify and simplify code
- evolve designs

The form I am using in this work is nearly identical to the one used by Martin in his Refactoring book. I have added the following to this form:

- A section on Communication, Duplication and Simplicity
- Numbered steps in the Mechanics section that correspond to numbered steps in the Examples section.

This is a continuously evolving piece of work. Your feedback is welcome – please send thoughts, comments or questions to joshua@industriallogic.com. This work lives on the internet at the following address: <http://industriallogic.com/xp/refactoring/>

I've also started an email list – called refactoring@yahoogroup.com – which is a good place to discuss refactoring, refactoring to patterns and emerging tools and IDEs that enable automated refactorings.

Introductory Chapters

[ToDo: write chapters on

- * how this book relates to Martin's book
- * being patterns happy (you learn a pattern and immediately find a place to use (abuse) it)
- * lessons learned about refactoring (how flexibility emerges from keeping code clean)
- * compositions of refactorings (or why I needed to write Chain Constructors)
- * categories of refactorings in this book
- * etc.]

[Appendix on the state of refactoring tools?]

Book History

I've written this book for a variety of reasons. I'll explain them in a moment. First, I'd like to explain how this book fits in with the landmark book, *Refactoring: Improving the Design of Existing Code*, by Martin Fowler and his distinguished contributors: Kent Beck, John Brant, William Opdyke and Don Roberts.

Patterns Happy

Did you ever learn a software pattern and then want to use it right away because you liked it so much? If so, you've been Patterns Happy.

Programmers who attend an intensive Design Patterns workshop I teach are cautioned, repeatedly about the effects of being Patterns Happy. To make the point crystal clear, I've taken to giving graduates of the workshop two take-home gifts: a gold seal to show they've made it through the intensive workshop and the following code listing (which some clever person on SlashDot posted [ToDo: look up reference]):

```
interface MessageStrategy {
    public void sendMessage();
}

abstract class AbstractStrategyFactory {
    public abstract MessageStrategy createStrategy(MessageBody mb);
}

class MessageBody {
    Object payload;
    public Object getPayload() {
        return payload;
    }
    public void configure(Object obj) {
        payload = obj;
    }
    public void send(MessageStrategy ms) {
        ms.sendMessage();
    }
}

class DefaultFactory extends AbstractStrategyFactory {
    private DefaultFactory() {
        ;
    }
    static DefaultFactory instance;
    public static AbstractStrategyFactory getInstance() {
        if (instance == null)
            instance = new DefaultFactory();
        return instance;
    }

    public MessageStrategy createStrategy(final MessageBody mb) {
        return new MessageStrategy() {
            MessageBody body = mb;
            public void sendMessage() {
                Object obj = body.getPayload();
                System.out.println((String) obj);
            }
        };
    }
}

public class HelloWorld {
    public static void main(String[] args) {
        MessageBody mb = new MessageBody();
        mb.configure("Hello World!");
        AbstractStrategyFactory asf = DefaultFactory.getInstance();
        MessageStrategy strategy = asf.createStrategy(mb);
    }
}
```

```
        mb.send(strategy);  
    }  
}
```

Do you see why I give away the above code listing? It's because the implementation ought to be so simple my grandmother could understand it. Yet this crazy, patterns-dense, Hello World program offers sophisticated runtime flexibility and configurability that no one will ever need! It's simply the perfect reminder for people learning patterns not to over-engineer with them. Extreme Programmers would look at the above code and whisper, "You Aren't Gonna Need It!"

The Patterns Happy malady isn't limited to beginner programmers. Intermediate and advanced programmers fall prey to it too, particularly after they read sophisticated patterns books or articles. For example, the other day I discovered an implementation of the Closure pattern on a system I was helping develop. It turned out that a programmer on the project had just learned Closure on Ward Cunningham's Wiki (<http://c2.com/cgi/wiki?UseClosuresNotEnumerations>) and had decided to use it on this system. As I studied the Closure implementation in our code, I could not find a good justification for using it in our code. It just wasn't necessary. So I refactored the Closure out of the code, replacing it with simpler code. When I finished I polled the programmers on the team to see if they thought the newer code was simpler than the Closure-code. They agreed that it was. So I used this opportunity to caution them about being Patterns Happy.

Small Steps

There was once a young, bright programmer who was attending an intensive testing and refactoring class I was teaching. Everyone in this class was participating in a coding challenge that involves refactoring code with over 80% of the smells described by Martin Fowler and Kent Beck in *Refactoring*. During this challenge, pairs of programmers must discover a smell, find a refactoring for the smell, announce the smell and associated refactoring and demonstrate the refactoring by programming it live, while others in the class observe the coding on a projector.

At about 5 minutes until noon, the class had been refactoring for nearly an hour. Since lunch had already been brought into the room, I asked if anyone had a small refactoring they'd like to complete before we broke for lunch. Our young programmer raised his hand and said he had one. Without mentioning a specific smell or associated refactoring, this fellow described a rather large problem in the code and explained what he intended to do to fix it. A few students expressed their doubt that such a problem could be fixed in only 5 minutes, but the programmer insisted that he could complete the work so we all agreed to let him and his pair have a try.

5 minutes passes very quickly when you are refactoring.

Our programmer and his pair, who were trying to work quickly, found that after moving and altering some code, many of the unit tests were now failing. Failing unit tests show up as a big red bar in the unit testing tool, which looks awfully big and red when it being projected onto a large screen! As the programmers attempted to fix the broken unit tests, one by one people began to go for lunch at the nearby table. Fifteen minutes later I myself took a break to get lunch, and as I stood on the line to get my food, I watched the programming action on the projector.

Ten minutes later the big red bar still hadn't turned to green (which is the signal that all tests are passing). At that point our young programmer and his pair quickly got up to obtain some food and then returned to the computer. Many of us looked on as the programmer attempted to eat his food with one hand while continuing to refactor with the other. Meanwhile, the minutes were ticking along...

At ten minutes to one (nearly fifty-five minutes after beginning their refactoring), the big red bar finally turned green. The refactoring was complete. As the class re-assembled, I asked everyone what had gone wrong. Our young programmer provided the answer: he had not taken small steps. By combining several refactorings into single, large steps, he thought he would go faster but just the opposite was true. Each big step generated failures in the unit tests, which took

a good deal of time to fix, not to mention that some of the fixes needed to be undone during later steps.

Many of us have had a similar experience – we take steps that are too large and then struggle for minutes or hours to get back to a green bar. The better I get at refactoring, the more I proceed by taking small steps.

Mechanics Matter

Every refactoring in the book, *Refactoring*, along with every refactoring in this book, includes a Mechanics section. The mechanics section describes the exact steps one must take to get from a beginning state to an end state. Once upon a time I didn't care at all about these "mechanics." I knew how to refactor, I reasoned, so I didn't need to follow a recipe to get from one place to another, I just needed to know the destination.

How wrong I was!

What I learned was that the mechanics section matters quite a bit if you happen to not want to end up in a "world of red."

[keep writing!]

When I Don't Refactor To Patterns

Code Smells

How do refactorings to patterns map to smells? Here's a table that includes some old smells and some fresh new ones:

[Note to reviewers: I plan on writing on chapter to summarize the smells, some of which are in Martin's book and some of which are new. The chapter won't cover the smells in detail (as Martin does a fine job of that) but will merely introduce readers to the concept, which is referenced a lot in this book.]

Smell	Refactorings
Conditional Complexity	<i>Move Embellishment to Decorator (73)</i> <i>Replace Conditional Calculations with Strategy (50)</i> <i>Replace State-Altering Conditionals with State (154)</i>
Combinatorial Explosion	<i>Replace Conditional Searches with Specification (99)</i>
Duplicated Code	<i>Chain Constructors (19)</i> <i>Introduce Polymorphic Creation with Factory Method (43)</i>
Inappropriate Intimacy	<i>Encapsulate Composite with Builder (64)</i>
Indecent Exposure	<i>Encapsulate Classes with Creation Methods (27)</i>
Large Class	<i>Extract Creation Class (34)</i>
Long Method	<i>Compose Method (119)</i> <i>Move Accumulation to Collecting Parameter (94)</i>
Long Parameter List	<i>Replace Conditional Calculations with Strategy (50)</i>
Primitive Obsession	<i>Replace Implicit Tree with Composite (60)</i> <i>Replace Conditional Calculations with Strategy (50)</i> <i>Replace Enum with Type-Safe Enum (146)</i> <i>Replace State-Altering Conditionals with State (154)</i> <i>Replace Conditional Searches with Specification (99)</i>
Switch Statement	<i>Replace Conditional Calculations with Strategy (50)</i>

[Add a smell called over-sophistication or some such thing for refactorings like Move Embellishment to Decorator]

Chain Constructors

You have multiple constructors that contain duplicate code

Chain the constructors together to obtain the least duplicate code

```
public class Loan {
    ...
    public Loan(float notional, float outstanding, int rating, Date expiry) {
        this.strategy = new TermROC();
        this.notional = notional;
        this.outstanding = outstanding;
        this.rating = rating;
        this.expiry = expiry;
    }
    public Loan(float notional, float outstanding, int rating, Date expiry, Date maturity) {
        this.strategy = new RevolvingTermROC();
        this.notional = notional;
        this.outstanding = outstanding;
        this.rating = rating;
        this.expiry = expiry;
        this.maturity = maturity;
    }
    public Loan(CapitalStrategy strategy, float notional, float outstanding,
        int rating, Date expiry, Date maturity) {
        this.strategy = strategy;
        this.notional = notional;
        this.outstanding = outstanding;
        this.rating = rating;
        this.expiry = expiry;
        this.maturity = maturity;
    }
}
```



```
public class Loan {
    ...
    public Loan(float notional, float outstanding, int rating, Date expiry) {
        this(new TermROC(), notional, outstanding, rating, expiry, null);
    }
    public Loan(float notional, float outstanding, int rating, Date expiry, Date maturity) {
        this(new RevolvingTermROC(), notional, outstanding, rating, expiry, maturity);
    }
    public Loan(CapitalStrategy strategy, float notional, float outstanding,
        int rating, Date expiry, Date maturity) {
        this.strategy = strategy;
        this.notional = notional;
        this.outstanding = outstanding;
        this.rating = rating;
        this.expiry = expiry;
        this.maturity = maturity;
    }
}
```

Motivation

Code that's duplicated across two or more of a class's constructors is an invitation for trouble. Someone adds a new variable to a class, updates a constructor to initialize the variable, but neglects to update the other constructors, and bang, say hello to your next bug. The more constructors you have in a class, the more duplication will hurt you. It's therefore a good idea to reduce or remove all duplication if possible, which has the added bonus of reducing your system's *code bloat*.

We often accomplish this refactoring with constructor chaining: specific constructors call more general-purpose constructors until a final constructor is reached. If you have one constructor at the end of every chain, I call that your *catch-all* constructor, since it handles every constructor call. This catch-all constructor often accepts more parameters than the other constructors, and may or may not be private or protected.

If you find that having many constructors on your class detracts from its usability, consider applying *Replace Multiple Constructors with Creation Methods* (22).

Communication	Duplication	Simplicity
When constructors in a class implement duplicate work, the code fails to communicate what is specific from what is general. Communicate this by having specific constructors forward calls to more general-purpose constructors and do unique work in each constructor.	Duplicate code in your constructors makes your classes more error-prone and harder to maintain. Find what is common, place it in general-purpose constructors, forward calls to these general constructors and implement what isn't general in each constructor.	If more than one constructor contains the same code, it's harder to see how each constructor is different. Simplify your constructors by making specific ones call more general purpose ones, in a chain.

Mechanics

- Find two constructors that contain duplicate code. Determine if one can call the other such that duplicate code can be safely (and hopefully easily) deleted from one of these constructors.
✓ Compile and test.
- Repeat step 1 for all constructors in the class, including ones you've already touched, in order to obtain as little duplication across all constructors as possible.
- Change the visibility of any constructors that may not need to be public.
✓ Compile and test.

Example

1. We'll go with the example shown in the code sketch. We start with a single Loan class, which has three constructors to represent different types of loans and tons of bloated and ugly duplication:

```
public Loan(float notional, float outstanding, int rating, Date expiry) {
    this.strategy = new TermROC();
    this.notional = notional;
    this.outstanding = outstanding;
    this.rating = rating;
    this.expiry = expiry;
}
```

```
public Loan(float notional, float outstanding, int rating, Date expiry, Date maturity) {
    this.strategy = new RevolvingTermROC();
    this.notional = notional;
    this.outstanding = outstanding;
    this.rating = rating;
    this.expiry = expiry;
    this.maturity = maturity;
}

public Loan(CapitalStrategy strategy, float notional, float outstanding, int rating,
    Date expiry, Date maturity) {
    this.strategy = strategy;
    this.notional = notional;
    this.outstanding = outstanding;
    this.rating = rating;
    this.expiry = expiry;
    this.maturity = maturity;
}
```

I study the first two constructors. They do contain duplicate code, but so does that third constructor. I consider which constructor it would be easier for the first constructor to call. I see that it could call the third constructor, with a minimum amount of work. So I change the first constructor to be:

```
public Loan(float notional, float outstanding, int rating, Date expiry) {
    this(new TermROC(), notional, outstanding, rating, expiry, null);
}
```

I compile and test to see that the change works.

2. I repeat step 1 to remove as much duplication as possible. This leads me to the second constructor. It appears that it too can call the third constructor, as follows:

```
public Loan(float notional, float outstanding, int rating, Date expiry, Date maturity) {
    this(new RevolvingTermROC(), notional, outstanding, rating, expiry, maturity);
}
```

I'm now aware that constructor three is my class's catch-all constructor, since it handles all of the construction details.

3. I check all callers of the three constructors to determine if I can change the public visibility of any of them. In this case, I can't (take my word for it – you can't see the code that calls these methods).

I compile and test to complete the refactoring.

Chaining To An Init Method

Chain Constructors isn't possible in languages like C++, but it is possible by using an init method.

Sometimes your own logic will prevent you from chaining constructors the way you'd like to.

[More to write]

[Init methods are sometimes necessary because you are doing dynamic object loading –
Class.forName.newInstance()]

Replace Multiple Constructors with Creation Methods

Constructors on a class make it hard to decide
which constructor to call during development

*Replace the constructors with intention-revealing
Creation Methods that return object instances*

Loan
+Loan(notional, customerRating, maturity) +Loan(notional, customerRating, maturity, expiry) +Loan(notional, outstanding, customerRating, maturity, expiry) +Loan(capitalStrategy, notional, customerRating, maturity, expiry) +Loan(capitalStrategy, notional, outstanding, customerRating, maturity, expiry)



Loan
-Loan(capitalStrategy, notional, outstanding, customerRating, expiry, maturity) <u>+createTermLoan(notional, customerRating, maturity) : Loan</u> <u>+createTermLoan(capitalStrategy, notional, outstanding, customerRating, maturity) : Loan</u> <u>+createRevolver(notional, outstanding, customerRating, expiry) : Loan</u> <u>+createRevolver(capitalStrategy, notional, outstanding, customerRating, expiry) : Loan</u> <u>+createRCTL(notional, outstanding, customerRating, maturity, expiry) : Loan</u> <u>+createRCTL(capitalStrategy, notional, outstanding, customerRating, maturity, expiry) : Loan</u>

Motivation

Some languages allow you to name your constructors any old way you like, regardless of the name of your class. Other languages, such as C++ and Java, don't allow this: each of your constructors must be named after your class name. If you have one simple constructor, this may not be problem. If you have multiple constructors, programmers will have to choose which constructor to call by studying which parameters are expected and/or poking around at the constructor code. What's wrong with that? A lot. Constructors simply don't communicate intention efficiently or effectively. The more constructors you have, the easier it is for programmers to mistakenly choose the wrong one. Having to choose which constructor to call slows down development and the code that does call one of the many constructors often fails to sufficiently communicate the nature of the object being constructed.

If you think that sounds bad, it gets worse. As systems mature, programmers often add more and more constructors to classes without checking to see if older constructors are still being used. Constructors that continue to live in a class when they aren't being used are dead weight, serving only to bloat the class and make it more complicated than it needs to be. Mature software systems are often filled with dead constructor code because programmers lack fast, easy ways to identify all callers to specific constructors: either their IDE doesn't help them with this or it is too much trouble to devise and execute search expressions that will identify the exact callers of a specific method. On the other hand, if the majority of object creation calls come through specifically-named methods, like `createTermLoan()` and `createRevolver()`, it is fairly trivial to find all callers to such explicitly-named methods.

Now, what does our industry call a method that creates objects? Many would answer "Factory Method," after the name given to a creational pattern in the classic book, *Design Pattern* [GoF]. But are all methods that create objects true Factory Methods? Given a broad definition of the term – i.e. a method that simply creates objects – the answer would be an emphatic "yes!" But given the way the authors of the creational pattern, Factory Method, wrote about it (in 1994),

it is clear that not every method that creates objects offers the kind of loose-coupling provided by a genuine Factory Method. So, to help us all be clearer when discussing designs or refactorings related to object creation, I'm using the term *Creation Method* to refer to a method that creates objects. This means that every Factory Method is a Creation Method but not necessarily the reverse. It also means that you can substitute the term *Creation Method* wherever Martin Fowler uses the term "factory method" in *Refactoring* [F] and wherever Joshua Bloch uses the term "static factory method" in *Effective Java* [Bloch].

Communication	Duplication	Simplicity
Many constructors on a class don't communicate available types very well – communicate type availability clearly by offering access to instances via intention-revealing Creation Methods	There is no direct duplication here; just many nearly identical-looking constructors	Figuring out which constructor to call isn't simple – make it simple by offering up the various types through intention-revealing Creation Methods.

Mechanics

After identifying a class that has copious constructors, it's best to consider applying *Extract Class* [F] or *Extract Subclass* (330) [F] *before* you decide to apply this refactoring. *Extract Class* is a good choice if the class in question is simply doing too much work – i.e. it has too many responsibilities. *Extract Subclass* is a good choice if instances of the class only use a small portion of the class's instance variables. If you apply *Extract Subclass*, also consider applying *Encapsulate Classes with Creation Methods* (27).

1. Identify a class that has copious constructors, is not overburdened with responsibilities and which has instances that use most of its instance variables.
2. Identify the *catch-all* constructor or create one using *Chain Constructors* (19).

Strictly speaking, you can implement this refactoring without having a catch-all constructor, though it's a good idea to create one if doing so eliminates duplicate code.

3. Identify a constructor that clients call to create a *kind* of instance and produce a Creation Method for that kind of instance. Make the Creation Method call your catch-all constructor whenever possible, to enable the elimination of constructors (step 6).

Give your Creation Method an intention-revealing name and make it accept the least number of parameters necessary to produce valid instances. Note that you may create more than one Creation Method for a given constructor.

4. Replace constructor calls that create the kind of instance chosen in step 3 with calls to your Creation Method.

✓ Compile and test.

5. Repeat steps 3 and 4 for every constructor you'd like to replace with a Creation Method.
6. Delete constructors that are no longer being called.

✓ Compile.

7. If your class has no subclasses, declare its remaining constructor(s) private. If it has subclasses, declare its remaining constructor(s) protected.

✓ Compile and test.

Example

1. I'll use the example shown in the code sketch. We start with a simple Loan class, which has copious constructors to represent some form of a Term Loan, Revolver or RCTL (a Revolver and Term Loan combination).

```
public class Loan ...
    public Loan(double notional, int customerRating, Date maturity) {
        this(notional, 0.00, customerRating, maturity, null);
    }
    public Loan(double notional, int customerRating, Date maturity, Date expiry) {
        this(notional, 0.00, customerRating, maturity, expiry);
    }
    public Loan(double notional, double outstanding, int customerRating, Date maturity,
        Date expiry) {
        this(null, notional, outstanding, customerRating, maturity, expiry);
    }
    public Loan(CapitalStrategy capitalStrategy, double notional, int customerRating,
        Date maturity, Date expiry) {
        this(capitalStrategy, notional, 0.00, customerRating, maturity, expiry);
    }
    public Loan(CapitalStrategy capitalStrategy, double notional, double outstanding,
        int customerRating, Date maturity, Date expiry) {
        this.notional = notional;
        this.outstanding = outstanding;
        this.customerRating = customerRating;
        this.maturity = maturity;
        this.expiry = expiry;
        this.capitalStrategy = capitalStrategy;

        if (capitalStrategy == null) {
            if (expiry == null)
                this.capitalStrategy = new TermCapitalStrategy();
            else if (maturity == null)
                this.capitalStrategy = new RevolverCapitalStrategy();
            else
                this.capitalStrategy = new RCTLCapitalStrategy();
        }
    }
}
```

This class represents different types of loans that behave in similar ways and that share the same instance variables. The class has five constructors, the last of which is the catch-all constructor. If you look at these constructors, it isn't easy to know which ones create Term Loans, which ones create Revolvers, and which ones create RCTLs. I happen to know that an RCTL needs both an expiry date and a maturity date; so to create one, I must call a constructor that lets me pass in both dates. But did you know that? Do you think the next programmer who reads this code will know it?

What else is embedded as implicit knowledge in the above constructors? Plenty. If you call the first constructor, which takes three parameters, you'll get back a Term Loan. But if you want a Revolver, you'll need to call one of the constructors that take two dates, and supply null for the maturity date. Hmmm, I wonder if all users of this code will know this? Or will they just have to learn by encountering some ugly bugs?

2. The next task is to identify the catch-all constructor for the Loan class. This is easy – it is the constructor that takes the most parameters:

```
public Loan(CapitalStrategy capitalStrategy, double notional, double outstanding,
```



```
int customerRating, Date maturity, Date expiry) {
    this.notional = notional;
    this.outstanding = outstanding;
    this.customerRating = customerRating;
    this.maturity = maturity;
    this.expiry = expiry;
    this.capitalStrategy = capitalStrategy;

    if (capitalStrategy == null) {
        if (expiry == null)
            this.capitalStrategy = new TermCapitalStrategy();
        else if (maturity == null)
            this.capitalStrategy = new RevolverCapitalStrategy();
        else
            this.capitalStrategy = new RCTLCapitalStrategy();
    }
}
```

3. Next, I identify a constructor that clients call to create a kind of instance:

```
public Loan(double notional, int customerRating, Date maturity) {
    this(notional, 0.00, customerRating, maturity, null);
}
```

This constructor is called to produce a Term Loan with a default TermCapitalStrategy. In order to produce a Creation Method for this kind of instance, I write a test first:

```
public void testTermLoanCreation() {
    Loan term1 = Loan.createTermLoan(NOTIONAL, CUSTOMER_RATING, MATURITY_DATE);
    assertTrue("type = term loan", term1.toString().indexOf("term loan") > -1);
}
```

This test doesn't compile, run or pass until I add the following public static method to Loan:

```
public static Loan createTermLoan(double notional, int customerRating, Date maturity) {
    return new Loan(null, notional, 0.00, customerRating, maturity, null);
}
```

I make this method call Loan's catch-all constructor since doing so may allow me to delete, at a later step, the constructor I started with.

4. Now, I find all client calls to the constructor identified in the previous step. Since that constructor only creates Term Loans with a default TermCapitalStrategy, it is safe to replace all of the constructor calls with calls to the new Creation Method. So code that looked like:

```
Loan termLoan = new Loan(notional, customerRating, maturity);
```

is changed to:

```
Loan termLoan = Loan.createTermLoan(notional, customerRating, maturity);
```

5. Repeating steps 3 and 4 yields the following set of Loan Creation Methods:

```
public static Loan createTermLoan(double notional, int customerRating, Date maturity) {
    return new Loan(null, notional, 0.00, customerRating, maturity, null);
}
public static Loan createTermLoan(CapitalStrategy capitalStrategy, double notional,
    double outstanding, int customerRating, Date maturity) {
    return new Loan(capitalStrategy, notional, outstanding, customerRating, maturity,
        null);
}
public static Loan createRevolver(double notional, double outstanding,
    int customerRating, Date expiry) {
    return new Loan(null, notional, outstanding, customerRating, null, expiry);
}
public static Loan createRevolver(CapitalStrategy capitalStrategy, double notional,
    double outstanding, int customerRating, Date expiry) {
```

```
        return new Loan(capitalStrategy, notional, outstanding, customerRating, null, expiry);
    }
    public static Loan createRCTL(double notional, double outstanding, int customerRating,
        Date maturity, Date expiry) {
        return new Loan(null, notional, outstanding, customerRating, maturity, expiry);
    }
    public static Loan createRCTL(CapitalStrategy capitalStrategy, double notional,
        double outstanding, int customerRating, Date maturity, Date expiry) {
        return new Loan(capitalStrategy, notional, outstanding, customerRating, maturity,
            expiry);
    }
}
```

6. The compiler is now my friend, as I attempt to delete constructors that are no longer being called. I'm able to delete all but the catch-all constructor, which is being called by all of the new Creation Methods.

7. The catch-all constructor can now be safely declared private:

```
private Loan(CapitalStrategy capitalStrategy, double notional, double outstanding,
    int customerRating, Date maturity, Date expiry)
```

The compiler agrees with my changes and I'm done.

It's now quite clear how to obtain the different kinds of Loan instances. The ambiguities have been revealed and the implicit knowledge has been made explicit. What's left to do? Well, the Creation Methods still do take a fairly large number of parameters, so I may consider applying *Introduce Parameter Object* (295) [F].

Parameterized Creation Methods

As you consider implementing this refactoring, you may calculate in your head that you'd need something on the order of 50 Creation Methods to account for every object configuration supported by your class. Since writing 50 methods doesn't sound like much fun, you may decide not to apply this refactoring. Keep in mind that there are other ways to handle this situation. First, you need not produce a Creation Method for every object configuration: you can write Creation Methods for the most popular configurations and leave some public constructors around to handle the rest of the cases. In addition, it often makes sense to use parameters to cut down on the number of Creation Methods.

Encapsulate Classes with Creation Methods

Clients directly instantiate classes that reside
in one package and implement a common interface

*Make the class constructors non-public and let clients
create instances of them using superclass Creation Methods*



Motivation

A client's ability to directly instantiate classes is useful so long as the client needs to know about the very existence of those classes. But what if the client doesn't need that knowledge? What if the classes live in one package, implement one interface and those conditions aren't likely to change? In that case, the classes in the package could be hidden from clients outside the package using public, superclass Creation Methods, each of which would return an instance that satisfied some common interface.

There are several motivations for doing this. First, it provides a way to rigorously apply the mantra, *separate interface from implementation* [GoF], by ensuring that clients interact with classes via their common interface. Second, it provides a way to reduce the *conceptual weight* [Bloch] of a package by hiding classes that don't need to be publicly visible outside their package (i.e. you don't need to know these classes exist). And third, it simplifies the construction of available *kinds* of instances by making the set available through intention-revealing Creation Methods.

Despite these good things, some folks have reservations about applying this refactoring. I address and respond to their concerns below:

1. They don't like giving a superclass knowledge of its subclasses, since it causes a dependency cycle - i.e. you have to add new Creation Method to a superclass just because you create a new subclass or add/modify a subclass constructor. When I point out that this refactoring happens within the context of one package with subclasses that implement one interface, they usually quiet down.
2. They don't like mixing Creation Methods with implementation methods on a superclass. I don't have a problem doing this, unless the Creation Methods just make it too hard to see what the superclass does, in which case I apply *Extract Creation Class* (34).
3. They don't like this refactoring when code is handed off as object code, since programmers who must use the object code won't be able to add or modify the non-public classes or the Creation Methods. I'm more sympathetic to this reservation. If extensibility within a package is necessary and users don't have source code, I would not encapsulate the classes, but would provide a Creation Class for common instances.

The sketch at the start of this refactoring gives you a glimpse of some object to relational database mapping code. Before the refactoring was applied, programmers (including myself) occasionally instantiated the wrong subclass or the right subclass with incorrect arguments (for example, we called a constructor that took a primitive Java `int` when we really needed to call the constructor that took Java's `Integer` object). The refactoring reduced bug creation by encapsulated the knowledge about the subclasses and producing a single place to get a variety of well-named subclass instances.

Communication	Duplication	Simplicity
When you expect client code to communicate with classes via one interface, your code needs to communicate this. Public constructors don't help, since they allow clients to couple themselves to class types. Communicate your intentions by protecting class constructors, producing instances via superclass Creation Methods and making the return type for the instances a common interface or abstract class type.	Duplication isn't an issue with this refactoring.	Making classes publicly visible when you want clients to interact with them via one interface isn't simple: it invites programmers to instantiate and couple themselves to class types and it communicates that it is ok to extend the public interface of an individual class. Simplify by making it impossible to instantiate these classes and by offering instances via superclass Creation Methods.

Forces

- Your classes share a common public interface.

This is essential because after the refactoring, all client code will interact with class instances via their common interface.

- Your classes reside in the same package.

Mechanics

1. Write an intention-revealing Creation Method on the superclass for a *kind* of instance that a class's constructor produces. Make the method's return type be the common interface type and make the method's body be a call to the class's constructor.
2. For the kind of instance chosen, replace all calls to the class's constructor with calls to the superclass Creation Method.
 - ✓ Compile and test.
3. Repeat steps 1 and 2 for any other kinds of instances that may be created by the class's constructor.
4. Declare the class's constructor to be non-public (i.e. protected or package-protected).
 - ✓ Compile.
5. Repeat the above steps until every constructor on the class is non-public and all available class instances may be obtained via superclass Creation Methods.

Example

1. We begin with a small hierarchy of classes that reside in a package called `descriptors`. These classes assist in the object-to-relation database mapping of database attributes to instance variables:

```
package descriptors;

public abstract class AttributeDescriptor {
    protected AttributeDescriptor(...)

public class BooleanDescriptor extends AttributeDescriptor {
    public BooleanDescriptor(...) {
        super(...);
    }

public class DefaultDescriptor extends AttributeDescriptor {
    public DefaultDescriptor(...) {
        super(...);
    }

public class ReferenceDescriptor extends AttributeDescriptor {
    public ReferenceDescriptor(...) {
        super(...);
    }
}
```

The abstract `AttributeDescriptor` constructor is protected, and the constructors for the three subclasses are public. Let's focus on the `DefaultDescriptor` subclass. The first step is to identify a kind of instance that can be created by the `DefaultDescriptor` constructor. To do that, I look at some client code:

```
protected List createAttributeDescriptors() {
    List result = new ArrayList();
    result.add(new DefaultDescriptor("remoteId", getClass(), Integer.TYPE));
    result.add(new DefaultDescriptor("createdDate", getClass(), Date.class));
    result.add(new DefaultDescriptor("lastChangedDate", getClass(), Date.class));
    result.add(new ReferenceDescriptor("createdBy", getClass(), User.class,
        RemoteUser.class));
    result.add(new ReferenceDescriptor("lastChangedBy", getClass(), User.class,
        RemoteUser.class));
    result.add(new DefaultDescriptor("optimisticLockVersion", getClass(), Integer.TYPE));
    return result;
}
```

Here I see that `DefaultDescriptor` is being used to represent mappings for Integers and Dates. It may also be used to map other types, but I must focus on one kind of instance at a time. So I decide to write a Creation Method to produce attribute descriptors for Integers:

```
public abstract class AttributeDescriptor {
    public static AttributeDescriptor forInteger(...) {
        return new DefaultDescriptor(...);
    }
}
```

I make the return type for the Creation Method an `AttributeDescriptor` because I want clients to interact with all `AttributeDescriptor` subclasses via the `AttributeDescriptor` interface and because I want to hide the very existence of `AttributeDescriptor` subclasses from anyone outside the descriptors package.

If you do test-first programming, you would begin this refactoring by writing a test to obtain the `AttributeDescriptor` instance you want from the superclass Creation Method.

2. Now client calls to create an Integer version of a `DefaultDescriptor` must be replaced with calls to the superclass Creation Method:

```
protected List createAttributeDescriptors() {
    List result = new ArrayList();
    result.add(AttributeDescriptor.forInteger("remoteId", getClass()));
}
```

```
result.add(new DefaultDescriptor("createdDate", getClass(), Date.class));
result.add(new DefaultDescriptor("lastChangedDate", getClass(), Date.class));
result.add(new ReferenceDescriptor("createdBy", getClass(), User.class,
    RemoteUser.class));
result.add(new ReferenceDescriptor("lastChangedBy", getClass(), User.class,
    RemoteUser.class));
result.add(AttributeDescriptor.forInteger("optimisticLockVersion", getClass()));
return result;
}
```

I compile and test that the new code works.

3. Now I continue to write Creation Methods for the remaining kinds of instances that the DefaultDescriptor constructor can create. This leads to 2 more Creation Methods:

```
public abstract class AttributeDescriptor {
    public static AttributeDescriptor forInteger(...) {
        return new DefaultDescriptor(...);
    }
    public static AttributeDescriptor forDate(...) {
        return new DefaultDescriptor(...);
    }
    public static AttributeDescriptor forString(...) {
        return new DefaultDescriptor(...);
    }
}
```

4. I now declare the DefaultDescriptor constructor protected:

```
public class DefaultDescriptor extends AttributeDescriptor {
    protected DefaultDescriptor(...) {
        super(...);
    }
}
```

I compile and everything goes according to plan.

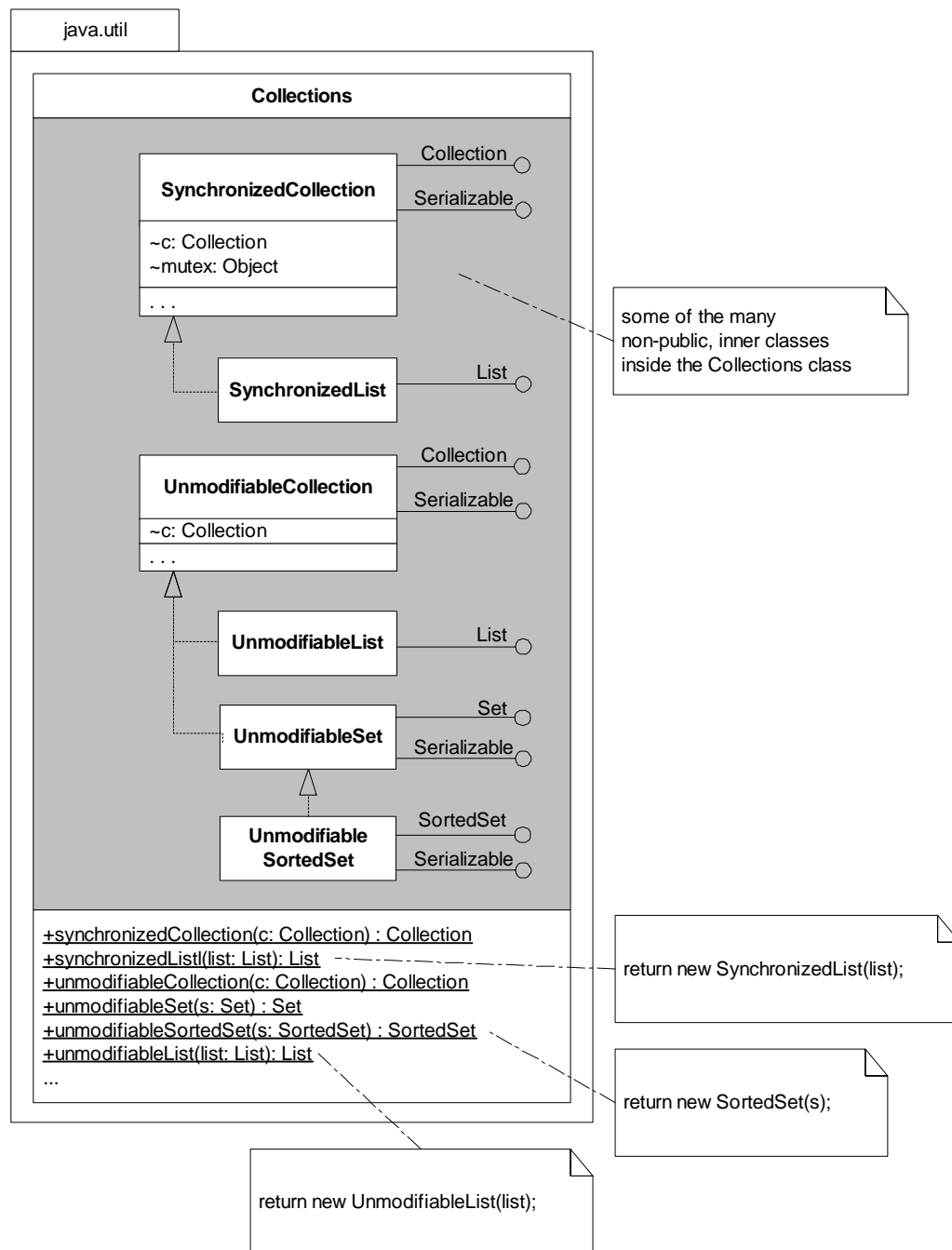
5. Now I repeat the above steps for the other AttributeDescriptor subclasses. When I'm done, the new code:

- gives access to AttributeDescriptor subclasses via their superclass
- ensures that clients obtain subclass instances via the AttributeDescriptor interface
- prevents clients from directly instantiating AttributeDescriptor subclasses
- communicates to other programmers that AttributeDescriptor subclasses are not meant to be public – the convention is to offer up access to them via the superclass and a common interface.

Encapsulating Inner Classes

The JDK's `java.util.Collections` class is a remarkable example of what encapsulating classes with Creation Methods is all about. The class's author, Joshua Bloch, needed to give programmers a way to make Collections, Lists, Sets and Maps unmodifiable and/or synchronized. He wisely chose to implement this behavior using the Decorator pattern. However, instead of creating public, `java.util` Decorator classes (for handling synchronization and unmodifiability) and then expecting programmers to decorate their own collections, he defined the Decorators in the Collections class as non-public inner classes and then gave Collections a set of Creation Methods from which programmers could obtain the kinds of decorated collections they needed.

Below is a sketch of a few of the inner classes and Creation Methods that are specified by the Collections class:



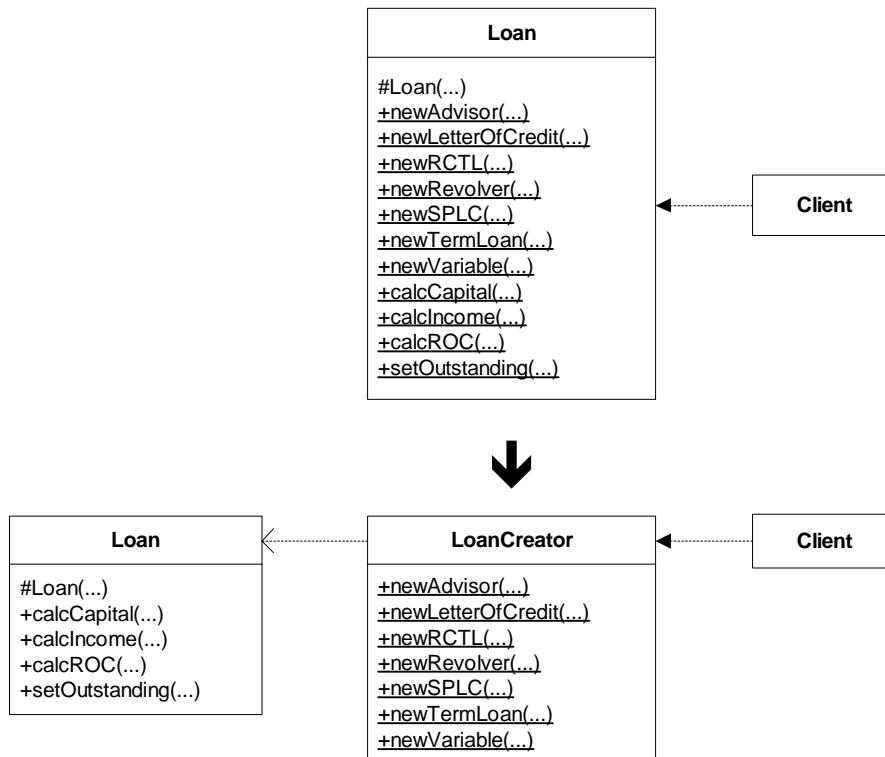
Notice that `java.util.Collections` even contains small hierarchies of inner classes, all of which are non-public. Each inner class has a corresponding method that receives a collection, decorates it and then returns the decorated instance, using a commonly defined interface type (such as `List` or `Set`). This solution reduced the number of classes programmers needed to know about, while providing the necessary functionality.

`java.util.Collections` is an example of a Creation Class (see *Extract Creation Class* (34)).

Extract Creation Class

Too many Creation Methods on a class obscure it's primary responsibility

Move the Creation Methods for a related set of classes to one Creation Class



Motivation

This refactoring is essentially *Extract Class* [F], only it's done on a class's Creation Methods. There's nothing wrong with a few Creation Methods on a class, but as the number of them grows, a class's own primary responsibilities – it's main purpose in life – may begin to feel obscured or overshadowed by creational logic. When that happens, it's better to restore the class's identity by moving its Creation Methods to a Creation Class.

Creation Classes and Abstract Factories [GoF] are similar in that they create families of objects, but they are quite different, as the following table illustrates:

	Creation Class	Abstract Factory
Substitutable at runtime	No	Yes
Instantiates a family of products	Yes	Yes
Supports creation of new products easily	Yes	No
Separates interface from implementation	May or may not	Yes
Is implemented with static methods	Usually	No
Is implemented as a Singleton	No	Often

In general, Creation Classes are good if you have one and only one class for creating a family of products, you don't need to substitute for another object, you can safely go with a Creation Class. anothereven though they are booin that they often create a related set of objects, but are most unlike Abstract Factories in that you don't substitute one Creation Class for another at runtime, because you're not concerned with swapping out one family of products for another. Creation Classes are usually implemented as classes that contain static methods, each of which instantiates and returns an object instance.

Communication	Duplication	Simplicity
When object creation begins to dominate the public interface of the class, the class no longer strongly communicates its main purpose. Communicate the act of object creation by creating a special class just to create object instances.	Duplication is not an issue with respect to this refactoring.	When creational responsibilities mix too much with a class's main responsibilities, the class isn't simple. Simplify it by extracting the creational code into a Creation Class.

Mechanics

1. Identify a class (which we'll call "A") that is overrun with Creation Methods.
 2. Create a class that will become your Creation Class. Name it after it's purpose in life, which will be to create various objects from a set of related classes.
 3. Move all Creational Methods from A to your new class, making sure that all protection privledges are accounted for.
 4. Change all callers to obtain object references from your new Creation Class.
- ✓ Compile and test.

Example

Though I use different example code from Martin Fowler, I do tend to repeat it as I am intrinsically lazy. So if you don't mind, we'll work with the same brainless Loan example, outlined in the code sketch above. Assume that there is test code for the example code below –I didn't include it the text since this refactoring is fairly trivial.

1. We begin with a Loan class that has lots of code for handling the responsibilities of a Loan and being a creator of Loan objects:

```
public class Loan {
    private double notional;
    private double outstanding;
    private int rating;
    private Date start;
    private CapitalStrategy capitalStrategy;
    private Date expiry;
    private Date maturity;
    // . . . more instances variables not shown

    protected Loan(double notional, Date start, Date expiry,
        Date maturity, int riskRating, CapitalStrategy strategy) {
        this.notional = notional;
        this.start = start;
        this.expiry = expiry;
        this.maturity = maturity;
        this.rating = riskRating;
        this.capitalStrategy = strategy;
    }

    public double calcCapital() {
        return capitalStrategy.calc(this);
    }

    public void setOutstanding(double newOutstanding) {
        outstanding = newOutstanding;
    }

    // ... more methods for dealing with the primary responsibilities of a Loan, not shown

    public static Loan newAdvisor(double notional, Date start,
        Date maturity, int rating)
        return new Loan(notional, start, null, maturity, rating, new TermLoanCapital());
    }

    public static Loan newLetterOfCredit(double notional, Date start,
        Date maturity, int rating) {
        return new Loan(notional, start, null, maturity, rating, new TermLoanCapital());
    }

    public static Loan newRCTL(double notional, Date start,
        Date expiry, Date maturity, int rating) {
        return new Loan(notional, start, expiry, maturity, rating, new RCTLCapital());
    }

    public static Loan newRevolver(double notional, Date start,
        Date expiry, int rating) {
        return new Loan(notional, start, expiry, null, rating, new RevolverCapital());
    }

    public static Loan newSPLC(double notional, Date start,
        Date maturity, int rating) {
        return new Loan(notional, start, null, maturity, rating, new TermLoanCapital());
    }

    public static Loan newTermLoan(double notional, Date start,
        Date maturity, int rating) {
        return new Loan(notional, start, null, maturity, rating, new TermLoanCapital());
    }

    public static Loan newVariableLoan(double notional, Date start,
        Date expiry, Date maturity, int rating) {
        return new Loan(notional, start, expiry, maturity, rating, new RCTLCapital());
    }
}
```

```
}
```

2. Next, I create a class called `LoanCreator`, since it's sole purpose in life is to be a place where clients can obtain `Loan` instances:

```
public class LoanCreator {  
}
```

3. Now I move all of the Creation Methods from `Loan` to `LoanCreator`, placing `LoanCreator` in the same package as `Loan` (and it's Capital strategies) so it has the protection level it needs to instantiate Loans:

```
public class LoanCreator {  
    public static Loan newAdvisor(double notional, Date start,  
                                  Date maturity, int rating)  
    {  
        return new Loan(notional, start, null, maturity, rating, new TermLoanCapital());  
    }  
    public static Loan newLetterOfCredit(double notional, Date start,  
                                          Date maturity, int rating) {  
        return new Loan(notional, start, null, maturity, rating, new TermLoanCapital());  
    }  
    public static Loan newRCTL(double notional, Date start,  
                               Date expiry, Date maturity, int rating) {  
        return new Loan(notional, start, expiry, maturity, rating, new RCTLCapital());  
    }  
    public static Loan newRevolver(double notional, Date start,  
                                   Date expiry, int rating) {  
        return new Loan(notional, start, expiry, null, rating, new RevolverCapital());  
    }  
    public static Loan newSPLC(double notional, Date start,  
                               Date maturity, int rating) {  
        return new Loan(notional, start, null, maturity, rating, new TermLoanCapital());  
    }  
    public static Loan newTermLoan(double notional, Date start,  
                                   Date maturity, int rating) {  
        return new Loan(notional, start, null, maturity, rating, new TermLoanCapital());  
    }  
    public static Loan newVariableLoan(double notional, Date start,  
                                       Date expiry, Date maturity, int rating) {  
        return new Loan(notional, start, expiry, maturity, rating, new RCTLCapital());  
    }  
}
```

4. To finish, I simply change calls of the form:

```
Loan termLoan = Loan.newTermLoan(...)
```

to

```
Loan termLoan = LoanCreator.newTermLoan(...)
```

And I compile and test.

Move Object Composition to Creation Method

Client code is responsible for wrapping objects together
to obtain one instance with the desired behavior

*Move the object composition responsibility
to an intention-revealing Creation Method*

Motivation

Prerequisites

Mechanics

Example

Replace Multiple Instances with Singleton

Your code creates multiple instances of an object that consumes too much memory or takes too long to instantiate

Replace the multiple instances with a Singleton

Example about Zip Code, City, State Object

Singleton is one of my least favorite patterns.

Replace Singleton with Object Reference

A class is a Singleton but has no business being a Singleton

*Replace the Singleton with a plain old, non-global instance
and pass that instance to objects that need it.*

```
public void someMethod() ...  
    Profile.getInstance().getUserLevel()
```



```
public void someMethod(Profile profile) ...  
    profile.getUserLevel()
```

Motivation

Replace Singleton with Registry

By J. B. Rainsberger

Motivation

You have a package or library that lives within an application and relies on global objects (singletons) provided by a part of the application. Your package is therefore coupled with the current application, but you would like your package to be used somewhere else.

You prefer not to apply Replace Singleton with Object Reference, because it will cause an unknown ripple effect throughout the application. This ripple effect is not something you can afford to handle at the present moment, so you are looking for a refactoring to help you get part of the way towards fixing the overall design issues.

After applying Replace Singleton with Registry, your package has access to the same data it had before, but that access is made local to the package in the form of a Registry. It then suffices to change the application so that it registers its data with the package's Registry. You can then use the package in other applications, as long as the application places the data your package needs in the prescribed, well-known location.

What is a Registry?

Briefly, a Registry is a namespace for objects. Clients can store information within a Registry so that other clients can retrieve that data without binding these clients to each other. We usually implement a Registry as a singleton, so the Registry is a well-known, global location for objects that allows providers and consumers of the data to operate independently of one another.

Forces

You have an application with (usually many) singletons that individual packages use to perform their work. Usually this is configuration information or widely-used resources like databases and external servers.

You would like to use one of your packages in a different application, or simply improve its design to make it application-independent.

You can refactor the application to register its singleton instances with well-known objects within the package. If you cannot do this, consider creating a simple application facade [insert reference] to help during the refactoring.

Mechanics

Identify the application-level objects your package needs to operate. Create a class called `PackageConfiguration` that aggregates all these objects.

Make `PackageConfiguration` a singleton, but add `setX()` methods for each object your package wants the application to register.

Within your application, as each object becomes available, call the corresponding `setX()` method on the `PackageConfiguration` object to "register" the object.

Within your package, replace each reference to the application's globally-accessible objects with the corresponding `getX()` method on the `PackageConfiguration` to retrieve the registered object.

When you have finished, you will have inserted a Registry, acting as a mediator [insert reference] between the application and your package. By registering application-level objects with the Registry, your application is one step closer to being decoupled from your package. It is possible now to have multiple packages retrieve objects from the Registry in order to perform their work. By retrieving application-level objects from the Registry, your package is one step closer to being decoupled from the application. It is possible now to have any application (although only one per virtual machine) register its global objects with the Registry without the package knowing the source of the objects.

Why did this happen?

You may be wondering how this abuse of singletons would arise in the first place. Put simply, the singleton is an easy way to make data available "from the application down" -- that is, storing data within the application and making it available to the components that need them. In many cases, the application itself only requires an attribute for each of these objects; however, in order to make the various components more "independent", programmers often create singletons in the hopes of pulling information from the application, rather than having the application push that information to the components.

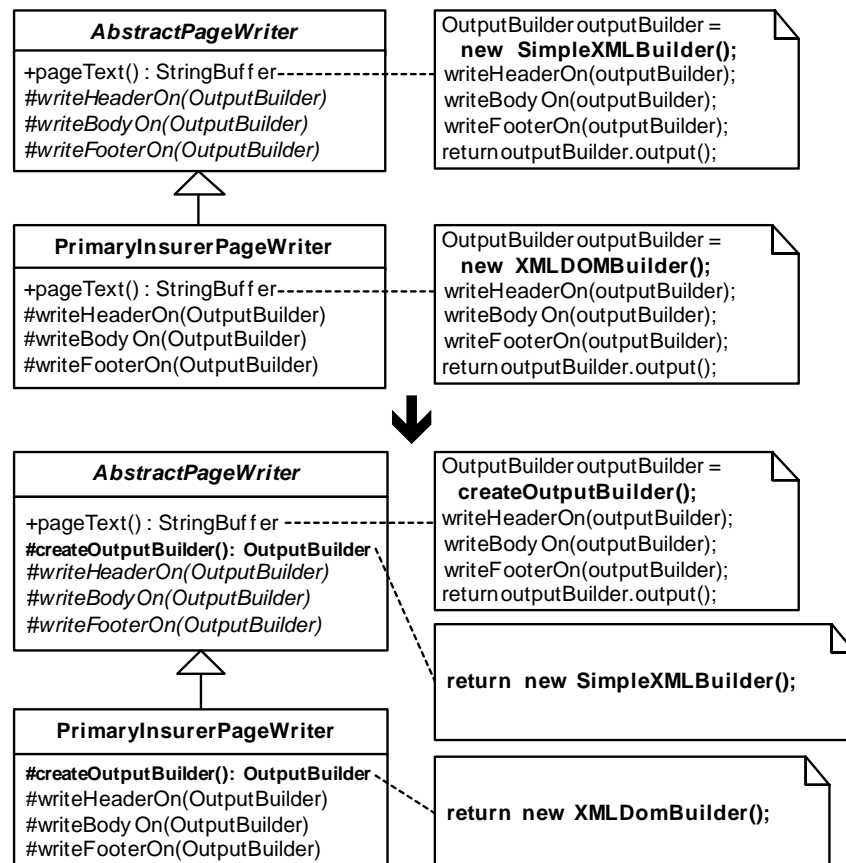
Unfortunately, as often happens, the component programmer requires something to be configured at the application level, and not at the component level. The programmer may, under the constraints of time and patience, give in to the temptation of simply "grabbing the data from the application," rather than providing a means for the application to configure the component.

[...]

Introduce Polymorphic Creation with Factory Method

Classes in a hierarchy implement a method similarly,
except for an object creation step

*Make a single superclass version of the method that
calls a Factory Method to handle the instantiation*



Motivation

What is a Factory Method [GoF]? It is a polymorphic method for creating and returning a Product. The method is declared in a superclass or interface. A superclass may implement the method and a subclass may override it, in order to make local decisions about the creation, including whether to instantiate a subclass of Product and/or how to initialize an instance.

Why would you refactor to a Factory Method? One motivation involves duplicate code: you find a method either in a superclass and overridden by a subclass or in several subclasses and this method is implemented nearly identically, except for an object creation step. You see how you could replace all versions of this method with a single superclass Template Method [GoF], provided that it could issue the object creation call, while letting the superclass and/or subclasses do the actual object creation work. No pattern is better suited to that task than Factory Method.

In his refactoring, *Form Template Method* (345) [F], Martin Fowler observes that, “inheritance is a powerful tool for eliminating duplicate behavior.” Inheritance is also what enables us to implement a Factory Method’s polymorphic object creation, since subclasses may control the class of object that gets instantiated. Template Methods often call Factory Methods [GoF, page 330], and many programmers refactor to both patterns to reduce duplication in class hierarchies.

Communication	Duplication	Simplicity
A well-chosen name for a Factory Method communicates intention better than a direct constructor call. Factory Methods also serve to communicate that the instances they return all implement a common interface.	Duplication of a method often results from a need to create an object instance in different ways. Remove the duplication by making a single method that obtains the instance it needs via a call to a Factory Method.	It’s usually simpler to read code that issues a call to a Factory Method than it is to read code that performs the actual instantiation. However, for those who aren’t comfortable with polymorphism, Factory Methods can seem to be more complex than direct instantiation calls.

Mechanics

Choose from the following two sets of mechanics:

A. When a method is duplicated because a superclass and subclass instantiate a type of object differently, refactor as follows:

1. In the superclass method, apply *Extract Method (110)* [F] on the object instantiation code to produce a Factory Method.

Make sure the return type for the Factory Method is a generic type, not the type of the concrete product being instantiated.

2. On each subclass that overrides the superclass method to create custom objects, extract the instantiation logic (using *Extract Method (110)* [F]) to produce a Factory Method with the same signature as the superclass Factory Method
3. Remove subclass versions of the method that are no longer needed.

If you don't expect subclasses to ever override this method, declare it as final.

✓ Compile and test.

B. When a method is duplicated across several subclasses because they instantiate a type of object differently, refactor as follows:

1. Create a Factory Method on the superclass. Declare it abstract if it does not make sense to have a default implementation, otherwise make it instantiate and return a default instance.
2. On each subclass that duplicates the method to do custom object creation, extract the instantiation logic (using *Extract Method (110)* [F]) to produce a Factory Method with the same signature as the superclass Factory Method.
3. Apply *Form Template Method (345)* [F], compile and test.

When you finish this step, the once duplicated method will now be a Template Method on the superclass, and this Template Method will call your new Factory Method. If you don't expect subclasses to ever override this method, declare it as final.

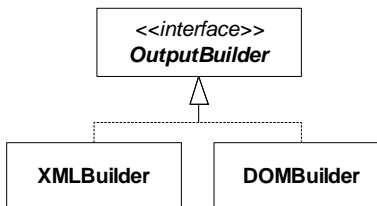
✓ Compile and test.

Example

The code sketch at the start of this refactoring comes from a system that outputs page text data (usually in the form of XML) for use in rendering HTML web pages. Subclasses of the `AbstractPageWriter` class decide what page text data to produce. `AbstractPageWriter` has a Template Method [GoF] called `pageText()` that looks like this:

```
public abstract class AbstractPageWriter...
    public String pageText() {
        OutputBuilder outputBuilder = new XMLBuilder();
        writeHeaderOn(outputBuilder);
        writeBodyOn(outputBuilder);
        writeFooterOn(outputBuilder);
        return outputBuilder.toString();
    }
    protected abstract void writeBodyOn(OutputBuilder builder);
    protected abstract void writeFooterOn(OutputBuilder builder);
    protected abstract void writeHeaderOn(OutputBuilder builder);
```

By default, `pageText()` creates an `OutputBuilder` of type `XMLBuilder` and passes it to three methods, after which it returns the `OutputBuilder`'s output. Subclasses override the three methods to customize what they output. Before we look at an example subclass, let's look at `OutputBuilders`:



`XMLBuilder` is a class that can build simple XML documents. It is usually sufficient for producing output in a system. On some occasions, however, code that builds output needs something a little more sophisticated, such as a `DOMBuilder`, which gives access to the Document Object Model.

A subclass of `AbstractPageWriter`, called `PrimaryInsurerPageWriter`, needed a `DOMBuilder`, so a programmer overrode the `pageText()` method as follows:

```
public class PrimaryInsurerPageWriter extends AbstractPageWriter...
    public String pageText() {
        OutputBuilder outputBuilder = new DOMBuilder();
        writeHeaderOn(outputBuilder);
        writeBodyOn(outputBuilder);
        writeFooterOn(outputBuilder);
        return outputBuilder.toString();
    }
}
```

As you can see, this is nearly a replica of the superclass `pageText()` method, the only difference being what kind of `OutputBuilder` is instantiated. Such duplication is a “breeding ground for bugs,” as Martin Fowler likes to call it. The duplication can be removed by refactoring to a Factory Method [GoF], as the steps below will show. Note: the refactoring mechanics labeled as “A” will be used in this example.

1. On the superclass, `AbstractPageWriter`, we apply *Extract Method (110)* [F] to produce a Factory Method [GoF], called `createOutputBuilder()`:

```
public abstract class AbstractPageWriter...
{
    public String pageText() {
        OutputBuilder outputBuilder = createOutputBuilder();
        writeHeaderOn(outputBuilder);
        writeBodyOn(outputBuilder);
        writeFooterOn(outputBuilder);
        return outputBuilder.toString();
    }
    protected OutputBuilder createOutputBuilder() {
        return new XMLBuilder();
    }
}
```

2. We perform a similar step on the subclass, PrimaryInsurerPageWriter:

```
public class PrimaryInsurerPageWriter extends AbstractPageWriter...
{
    public String pageText() {
        OutputBuilder outputBuilder = createOutputBuilder();
        writeHeaderOn(outputBuilder);
        writeBodyOn(outputBuilder);
        writeFooterOn(outputBuilder);
        return outputBuilder.toString();
    }
    protected OutputBuilder createOutputBuilder() {
        return new DOMBuilder();
    }
}
```

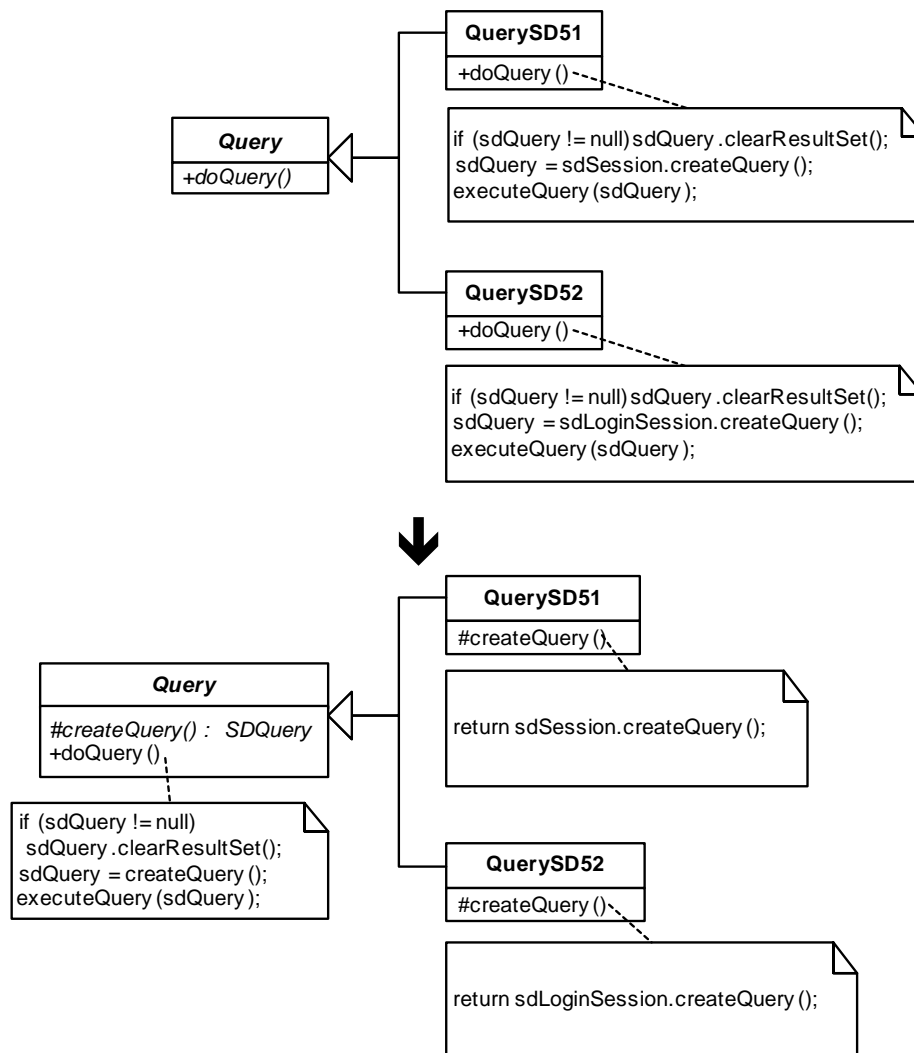
3. Now the pageText() method from PrimaryInsurerPageWriter can be deleted:

```
public class PrimaryInsurerPageWriter extends AbstractPageWriter...
{
    public String pageText() {
        OutputBuilder outputBuilder = createOutputBuilder();
        writeHeaderOn(outputBuilder);
        writeBodyOn(outputBuilder);
        writeFooterOn(outputBuilder);
        return outputBuilder.toString();
}
}
```

We compile and run tests, such as testPrimaryInsurerPageOutput(), to confirm that everything still works:

```
public void testPrimaryInsurerPageOutput() {
    String primaryInsurerOutput = getPrimaryInsurerPageWriter().pageText();
    assertTrue(primaryInsurerOutput.indexOf(KIM_NAME) > -1);
    assertTrue(primaryInsurerOutput.indexOf(KIM_ADDRESS) > -1);
    assertTrue(primaryInsurerOutput.indexOf(KIM_OCCUPATION) > -1);
    ...
}
```

Example 2: Duplication Across Subclasses



This example is similar to the previous one, only this time we begin with duplication in two subclasses. We can remove this duplication by introducing a Factory Method and a Template Method. I'll use the mechanics labeled as "B" to demonstrate how this refactoring works.

1. We start with some classes that perform database queries:

```
abstract class Query...
    public abstract void doQuery() throws QueryException;

class QuerySD51 extends Query ...
    public void doQuery() throws QueryException {
        if (sdQuery != null) sdQuery.clearResultSet();
        sdQuery = sdSession.createQuery(SDQuery.OPEN_FOR_QUERY);
        executeQuery(sdQuery);
    }

class QuerySD52 extends Query ...
    public void doQuery() throws QueryException {
        if (sdQuery != null) sdQuery.clearResultSet();
```



```
        sdQuery = sdLoginSession.createQuery(SDQuery.OPEN_FOR_QUERY);
        executeQuery(sdQuery);
    }
```

I add a Factory Method to the superclass, Query, and declare it abstract so that subclasses must implement it:

```
abstract class Query...
    protected abstract SDQuery createQuery() throws QueryException;
```

2. Now I'll create a Factory Method in each subclass by extracting the instantiation logic from the subclass implementations of doQuery():

```
class QuerySD51 extends Query ...
    protected SDQuery createQuery() {
        return sdSession.createQuery(SDQuery.OPEN_FOR_QUERY);
    }
    public void doQuery() throws QueryException {
        if (sdQuery != null) sdQuery.clearResultSet();
        sdQuery = createQuery();
        executeQuery(sdQuery);
    }

class QuerySD52 extends Query ...
    protected SDQuery createQuery() {
        return sdLoginSession.createQuery(SDQuery.OPEN_FOR_QUERY);
    }
    public void doQuery() throws QueryException {
        if (sdQuery != null) sdQuery.clearResultSet();
        sdQuery = createQuery();
        executeQuery(sdQuery);
    }
}
```

3. Finally, I apply *Form Template Method (345)* [F], to produce a single, superclass doQuery() method:

```
abstract class Query...
    protected abstract SDQuery createQuery() throws QueryException;
    public void doQuery() throws QueryException {
        if (sdQuery != null) sdQuery.clearResultSet();
        sdQuery = createQuery();
        executeQuery(sdQuery);
    }

class QuerySD51 extends Query ...
    protected SDQuery createQuery() {
        return sdSession.createQuery(SDQuery.OPEN_FOR_QUERY);
    }

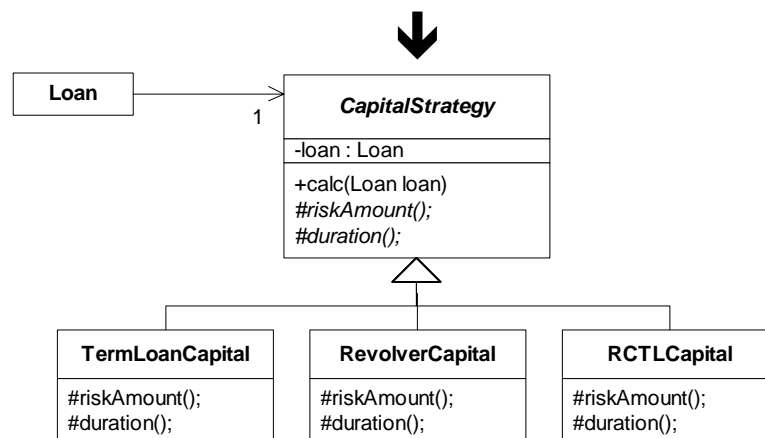
class QuerySD52 extends Query ...
    protected SDQuery createQuery() {
        return sdLoginSession.createQuery(SDQuery.OPEN_FOR_QUERY);
    }
}
```

Replace Conditional Calculations with Strategy

You use a lot of conditional logic in a calculation

Delegate the calculation to a Strategy object

```
public class Loan ...
public double calcCapital() {
    return riskAmount() * duration() * RiskFactor.forRiskRating(rating);
}
private double riskAmount() {
    if (unusedPercentage != 1.00)
        return outstanding + calcUnusedRiskAmount();
    else return outstanding;
}
private double calcUnusedRiskAmount() {
    return (notional - outstanding) * unusedPercentage;
}
private double duration() {
    if (expiry == null)
        return ((maturity.getTime() - start.getTime())/MILLIS_PER_DAY)/365;
    else if (maturity == null)
        return ((expiry.getTime() - start.getTime())/MILLIS_PER_DAY)/365;
    else {
        long millisToExpiry = expiry.getTime() - start.getTime();
        long millisFromExpiryToMaturity = maturity.getTime() - expiry.getTime();
        double revolverDuration = (millisToExpiry/MILLIS_PER_DAY)/365;
        double termDuration = (millisFromExpiryToMaturity/MILLIS_PER_DAY)/365;
        return revolverDuration + termDuration;
    }
}
private void setUnusedPercentage() {
    if (expiry != null && maturity != null) {
        if (rating > 4) unusedPercentage = 0.95;
        else unusedPercentage = 0.50;
    } else if (maturity != null) {
        unusedPercentage = 1.00;
    } else if (expiry != null) {
        if (rating > 4) unusedPercentage = 0.75;
        else unusedPercentage = 0.25;
    }
}
```



Motivation

A lot of conditional logic can obscure any calculation, even a simple one. When that happens, your calculation can be misunderstood by others and harder to maintain, debug and extend. Strategy is a pattern that deals well with calculations. A context object obtains a Strategy object and then delegates a calculation (or calculations) to that Strategy. This lightens the context class by moving the conditional calculation logic to a small collection of independent calculation objects (*strategies*), each of which can handle one of the various ways of doing the calculation.

Does this sound like a pattern you'd refactor to a lot? It may, but in my experience, I don't refactor to Strategy that often. I certainly have refactored to it, but I find that a lot of calculation logic I either write or come across isn't sufficiently complicated to justify using Strategy. In addition, when there is enough conditional logic to merit using the pattern, I have to consider whether a Template Method would be a better choice. But using a Template Method assumes that you can place the skeleton of your calculation in a base class, and have subclasses supply some or all of the calculation details. That may or may not be possible given your situation. For example, if you already have subclasses and the various ways of calculating something won't easily fit into those subclasses, you may not be able to *Form Template Method* [F]. Or, you may find that by placing calculations in separate subclasses, you limit your ability to swap one calculation type for another at runtime, since it would mean changing the type of object a client is working with, rather than simply substituting one Strategy object for another.

Once you do decide to refactor to Strategy, you have to consider how the calculation embedded in each strategy class will get access to the variables it needs to do its calculation. To accomplish that, I usually pass the context class as a reference to the Strategy object, and make whatever variables are needed by each Strategy accessible via public methods on the context class.

The final thing to consider is how your context class will obtain its Strategy. Whenever possible, I like to shield client code from having to worry about both instantiating a Strategy instance and passing it to a context's constructor. Creation Methods can help with this: just define one or more methods that return a context instance, properly outfitted with the appropriate Strategy instance.

Communication	Duplication	Simplicity
Copious conditional logic obscures the steps of a calculation. Communicate the steps clearly by separating each calculation variety into its own Strategy class. Then clarify which variety of calculation your object uses by writing code to pass the appropriate Strategy to the object for its use in performing the calculation.	Conditional calculation code can often contain duplicate conditional statements that are used to calculate various variables in an algorithm. Replace all of the conditional logic by encapsulating each variety of the calculation in its own Strategy class.	Classes that contain a lot of conditional logic are never simple. But if a class contains lots of conditional logic for calculating something in a variety of ways, it may also be more complex than it needs to be, as it knows too much. Simplify the class by extracting each variety of the calculation into its own Strategy class and then delegate to one of these classes to obtain a calculation.

Mechanics

1. On a class (which we'll call "A") identify a calculation method, or helper methods to such a method, that contain a lot of conditional logic. This class will be known as your context class as it will be the context for a Strategy object.
2. Create a concrete class and name it based on the behavior performed by the chosen calculation method. This will be your Strategy.

You can append the word "Strategy" to the class name if you find it helps communicate the purpose of this new type.

3. Apply *Move Method* [F] to move the primary calculation method and any helper methods to your Strategy class. If the code you move needs to obtain information from A, pass A as a parameter to the primary calculation method or as a parameter to the Strategy class's constructor and make sure the information on A is publicly available.

You can alternatively pass the necessary information from A to the Strategy, without passing a reference of A to the Strategy. This will result in less coupling between A and your Strategy, but may require you to pass a lot of information. See *Design Patterns* [GoF] for an in-depth discussion about communication between the context, A, and the Strategy.

4. Create a field (which we'll call "S") in A for the Strategy and instantiate it.
5. Update the primary calculation method in A to delegate the calculation to S.
6. Compile and test
7. On the Strategy class, apply *Replace Conditional with Polymorphism* [F] on the primary calculation method and any helper methods you moved from A. It is best to do this step slowly, by focusing on extracting one subclass at a time, then performing steps 8 and 9 below and then repeating this step. When finished, you will have substantially reduced the conditional logic in your Strategy class and you will have defined concrete Strategy classes for each variety of the calculation you started with.

Consider applying *Form Template Method* [F] for your Strategy's primary calculation method. You may also make your original Strategy class abstract.

8. Add code to A to either use its internal logic to set the value of S or to allow an external client to pass in a value for S.

If you go with the latter approach, let clients pass in a value for S via constructor calls if clients won't need to change S's value at runtime. Otherwise, supply a setter method to let clients set the value of S at runtime. For convenience, you can also do both. If clients will be able to pass in a value of S to A, you'll need to update the code for every existing client of A.

9. Compile and test.

Example

The example in the code sketch above deals with calculating capital for bank loans. It shows a fair amount of conditional logic that's used in performing this calculation, although it is even less conditional logic than was contained in the original code, which had to handle capital calculations for 7 different loan profiles.

In the example, the context class is called `Loan`. We'll be seeing how `Loan`'s method for calculating capital can be *strategized*, i.e. delegated to a Strategy object. As you study the example, you may wonder why `Loan` wasn't just subclassed to support the three different styles of capital calculations. That *was* an option, however, because the application that uses `Loan` needed to change a `Loan`'s capital calculation at runtime, without changing the class type of the `Loan`, it was better to use the Strategy pattern.

1. We'll begin by looking at the `Loan` class's `calcCapital()` method and its helper methods (note: I show a few tests for `calcCapital()` in step 6 below):

```
public class Loan ...
    private double notional;
    private double outstanding;
    private int rating;
    private double unusedPercentage;
    private Date start;
    private Date expiry;
    private Date maturity;
    private static final int MILLIS_PER_DAY = 86400000;

    public double calcCapital() {
        return riskAmount() * duration() * RiskFactor.forRiskRating(rating);
    }
    private double calcUnusedRiskAmount() {
        return (notional - outstanding) * unusedPercentage;
    }
    private double duration() {
        if (expiry == null)
            return ((maturity.getTime() - start.getTime()) / MILLIS_PER_DAY) / 365;
        else if (maturity == null)
            return ((expiry.getTime() - start.getTime()) / MILLIS_PER_DAY) / 365;
        else {
            long millisToExpiry = expiry.getTime() - start.getTime();
            long millisFromExpiryToMaturity = maturity.getTime() - expiry.getTime();
            double revolverDuration = (millisToExpiry / MILLIS_PER_DAY) / 365;
            double termDuration = (millisFromExpiryToMaturity / MILLIS_PER_DAY) / 365;
            return revolverDuration + termDuration;
        }
    }
    private double riskAmount() {
        if (unusedPercentage != 1.00)
            return outstanding + calcUnusedRiskAmount();
        else
            return outstanding;
    }
    public void setOutstanding(double newOutstanding) {
        outstanding = newOutstanding;
    }
    private void setUnusedPercentage() {
        if (expiry != null && maturity != null) {
            if (rating > 4)
                unusedPercentage = 0.95;
            else
                unusedPercentage = 0.50;
        } else if (maturity != null) {
            unusedPercentage = 1.00;
        } else if (expiry != null) {
            if (rating > 4)
```

```
        unusedPercentage = 0.75;
    else
        unusedPercentage = 0.25;
    }
}
```

2. The Strategy I'd like to define will handle the `calcCapital()` calculation. So I create a class called `CapitalStrategy`.

```
public class CapitalStrategy {
}
```

3. Now I'm up to the hardest step: I need to move methods from `Loan` to `CapitalStrategy`. I begin with the `calcCapital()` method. In this case, I don't want to move this method, but rather, copy it to `CapitalStrategy`:

```
public class CapitalStrategy {
    public double calc() {
        return riskAmount() * duration() * RiskFactor.forRiskRating(rating);
    }
}
```

That code won't even compile, because `CapitalStrategy` doesn't contain the methods it is calling. No problem. I pass `calc()` a `Loan` parameter and update the code as follows:

```
public double calc(Loan loan) {
    return loan.riskAmount() * loan.duration() * RiskFactor.forRiskRating(loan.rating);
}
```

That gets us closer, but the compiler still complains that the methods and variable I'm accessing on `Loan` aren't visible (i.e. they are private, not public). I change the visibility to public and finally the compiler is happy. Later, I'll be moving some of these public methods/fields to `CapitalStrategy` or making them accessible via `Loan` getter methods.

Now I focus on moving each piece of the calculation from `Loan` to `CapitalStrategy`. The method, `riskAmount()` (which is now public) is first on my radar screen.

```
public double riskAmount() {
    if (unusedPercentage != 1.00)
        return outstanding + calcUnusedRiskAmount();
    else
        return outstanding;
}
```

This method relies on other fields and methods within `Loan`. I study the code and see that the field, `outstanding`, is used extensively in the `Loan` class, but the field, `unusedPercentage`, along with the methods, `setUnusedPercentage()` and `calcUnusedRiskAmount()` are only there to help the `calcCapital()` method. So I decide to move all of this code, with the exception of the field, `outstanding`, to `CapitalStrategy`:

```
public class CapitalStrategy {
    private Loan loan;
    public double calc(Loan loan) {
        this.loan = loan;
        return riskAmount() * loan.duration() * RiskFactor.forRiskRating(loan.rating);
    }
    private double calcUnusedPercentage() {
        if (loan.expiry != null && loan.maturity != null) {
            if (loan.rating > 4)
                return 0.95;
        }
    }
}
```

```

        else
            return 0.50;
    } else if (loan.maturity != null) {
        return 1.00;
    } else if (loan.expiry != null) {
        if (loan.rating > 4)
            return 0.75;
        else
            return 0.25;
    }
    return 0.0;
}
private double calcUnusedRiskAmount() {
    return (loan.notional - loan.outstanding) * calcUnusedPercentage();
}
public double riskAmount() {
    if (calcUnusedPercentage() != 1.00)
        return loan.outstanding + calcUnusedRiskAmount();
    else
        return loan.outstanding;
}
}

```

To make this compile, I need to make more fields on Loan public:

```

public class Loan ...
    public double notional;
    public double outstanding;
    public int rating;
    private double unusedPercentage; //replaced with calculation method on CapitalStrategy
    public Date start;
    public Date expiry;
    public Date maturity;

```

By now I'm not happy having all these public fields. So I make getter methods for them and update the CapitalStrategy code accordingly. After this, all I do is move the duration() calculation over to CapitalStrategy and this step of the refactoring is done. CapitalStrategy now looks like this:

```

public class CapitalStrategy {
    private Loan loan;
    private static final int MILLIS_PER_DAY = 86400000;
    public double calc(Loan loan) {
        this.loan = loan;
        return riskAmount() * duration() * RiskFactor.forRiskRating(loan.getRating());
    }
    private double calcUnusedPercentage() {
        if (loan.getExpiry() != null && loan.getMaturity() != null) {
            if (loan.getRating() > 4) return 0.95;
            else return 0.50;
        } else if (loan.getMaturity() != null) {
            return 1.00;
        } else if (loan.getExpiry() != null) {
            if (loan.getRating() > 4) return 0.75;
            else return 0.25;
        }
        return 0.0;
    }
    private double calcUnusedRiskAmount() {
        return (loan.getNotional() - loan.getOutstanding()) * calcUnusedPercentage();
    }
    public double duration() {
        if (loan.getExpiry() == null)
            return (
                (loan.getMaturity().getTime() - loan.getStart().getTime()) / MILLIS_PER_DAY
                / 365;
            )
        else if (loan.getMaturity() == null)

```

```

        return (
            (loan.getExpiry().getTime() - loan.getStart().getTime()) / MILLIS_PER_DAY)
            / 365;
        else {
            long millisToExpiry = loan.getExpiry().getTime() - loan.getStart().getTime();
            long millisFromExpiryToMaturity =
                loan.getMaturity().getTime() - loan.getExpiry().getTime();
            double revolverDuration = (millisToExpiry / MILLIS_PER_DAY) / 365;
            double termDuration = (millisFromExpiryToMaturity / MILLIS_PER_DAY) / 365;
            return revolverDuration + termDuration;
        }
    }
    public double riskAmount() {
        if (calcUnusedPercentage() != 1.00)
            return loan.getOutstanding() + calcUnusedRiskAmount();
        else
            return loan.getOutstanding();
    }
}

```

4. Now I need to make a field in the Loan class for the CapitalStrategy class:

```

public class Loan...
    private CapitalStrategy capitalStrategy = new CapitalStrategy();

```

5. And I'm finally ready to have Loan delegate its calculation of capital to CapitalStrategy's calc() method:

```

public double calcCapital() {
    return capitalStrategy.calc(this);
}

```

6. I can now compile and run my tests. Here are a few of the tests that ensure whether the capital calculation works for various types of loan profiles:

```

public void testTermLoanCapital() {
    Loan termLoan = Loan.newTermLoan(10000.00, startOfLoan(), maturity(), RISK_RATING);
    termLoan.setOutstanding(10000.00);
    assertEquals("Capital for Term Loan", 37500.00, termLoan.calcCapital(), penny);
}
public void testRevolverROC() {
    Loan revolver = Loan.newRevolver(10000.00, startOfLoan(), expiry(), RISK_RATING);
    revolver.setOutstanding(2000.00);
    assertEquals("Capital for Revolver", 6000.00, revolver.calcCapital(), penny);
}
public void testRevolverTermROC() {
    Loan rctl = Loan.newRCTL(10000.00, startOfLoan(), expiry(), maturity(), RISK_RATING);
    rctl.setOutstanding(5000.00);
    assertEquals("Capital for RCTL", 28125.00, rctl.calcCapital(), penny);
}

```

These tests, and similar ones, all run successfully.

7. At this point I've moved a lot of code out of the Loan class and into the CapitalStrategy class, which now encapsulates the bulky conditional calculation logic. I want to tame this logic by decomposing CapitalStrategy into several subclasses, one for each way we calculate capital. I do this by applying *Replace Conditional with Polymorphism* [F].

First, I identify a total of three different ways of doing the capital calculation, each of which corresponds to a specific Loan profile: Term loan, Revolver or RCTL (a combination of a

Revolver, which converts to a Term Loan on an expiry date). I decide to start by creating a subclass of `CapitalStrategy` that is capable of calculating capital for a Term Loan:

```
public class TermLoanCapital extends CapitalStrategy {  
}
```

Now, I find the specific calculation code that applies to a Term Loan and push it down into the new subclass:

```
public class TermLoanCapital extends CapitalStrategy {  
    protected double duration() {  
        return (  
            (loan.getMaturity().getTime() - loan.getStart().getTime()) / MILLIS_PER_DAY  
            / 365;  
        )  
    }  
    protected double riskAmount() {  
        return loan.getOutstanding();  
    }  
}
```

I now push on to steps 8 and 9 of the refactoring, after which I'll circle back to define, configure and test two more concrete Strategy classes: `RevolverCapital` and `RCTLCapital`.

8. Now I need to configure the `Loan` class with the `TermLoanCapital` strategy when it is applicable, so that I can test whether it works. To do this, I make the following modifications:

```
public class Loan...  
    private CapitalStrategy capitalStrategy;  
  
    protected Loan(double notional, Date start, Date expiry,  
        Date maturity, int riskRating, CapitalStrategy strategy) {  
        this.notional = notional;  
        this.start = start;  
        this.expiry = expiry;  
        this.maturity = maturity;  
        this.rating = riskRating;  
        this.capitalStrategy = strategy;  
    }  
    public static Loan newRCTL(double notional, Date start, Date expiry,  
        Date maturity, int rating) {  
        return new Loan(notional, start, expiry, maturity, rating, new CapitalStrategy());  
    }  
    public static Loan newRevolver(double notional, Date start, Date expiry,  
        int rating) {  
        return new Loan(notional, start, expiry, null, rating, new CapitalStrategy());  
    }  
    public static Loan newTermLoan(double notional, Date start, Date maturity,  
        int rating) {  
        return new Loan(notional, start, null, maturity, rating, new TermLoanCapital());  
    }  
}
```

9. I compile and test and all goes well. Now I circle back to step 7, to define the additional concrete Strategy classes, configure the `Loan` class to work with them and test everything. When I'm done, almost all of the original conditional calculation logic is gone and I have three Strategies for calculating capital:

```
public class Loan...  
    public static Loan newRCTL(double notional, Date start, Date expiry,  
        Date maturity, int rating) {  
        return new Loan(notional, start, expiry, maturity, rating, new RCTLCapital());  
    }  
    public static Loan newRevolver(double notional, Date start, Date expiry,  
        int rating) {
```

```

        return new Loan(notional, start, expiry, null, rating, new RevolverCapital());
    }
    public static Loan newTermLoan(double notional, Date start, Date maturity,
        int rating) {
        return new Loan(notional, start, null, maturity, rating, new TermLoanCapital());
    }
}

public abstract class CapitalStrategy {
    protected Loan loan;
    protected static final int MILLIS_PER_DAY = 86400000;
    public double calc(Loan loan) {
        this.loan = loan;
        return riskAmount() * duration() * RiskFactor.forRiskRating(loan.getRating());
    }
    protected abstract double duration();
    protected abstract double riskAmount();
}

public class TermLoanCapital extends CapitalStrategy {
    protected double duration() {
        return (
            (loan.getMaturity().getTime() - loan.getStart().getTime()) / MILLIS_PER_DAY)
            / 365;
    }
    protected double riskAmount() {
        return loan.getOutstanding();
    }
}

public class RevolverCapital extends CapitalStrategy {
    private double calcUnusedPercentage() {
        if (loan.getRating() > 4) return 0.75;
        else return 0.25;
    }
    private double calcUnusedRiskAmount() {
        return (loan.getNotional() - loan.getOutstanding()) * calcUnusedPercentage();
    }
    protected double duration() {
        return (
            (loan.getExpiry().getTime() - loan.getStart().getTime()) / MILLIS_PER_DAY)
            / 365;
    }
    protected double riskAmount() {
        return loan.getOutstanding() + calcUnusedRiskAmount();
    }
}

public class RCTLCapital extends CapitalStrategy {
    private double calcUnusedPercentage() {
        if (loan.getRating() > 4) return 0.95;
        else return 0.50;
    }
    private double calcUnusedRiskAmount() {
        return (loan.getNotional() - loan.getOutstanding()) * calcUnusedPercentage();
    }
    protected double duration() {
        long millisToExpiry = loan.getExpiry().getTime() - loan.getStart().getTime();
        long millisFromExpiryToMaturity =
            loan.getMaturity().getTime() - loan.getExpiry().getTime();
        double revolverDuration = (millisToExpiry / MILLIS_PER_DAY) / 365;
        double termDuration = (millisFromExpiryToMaturity / MILLIS_PER_DAY) / 365;
        return revolverDuration + termDuration;
    }
    protected double riskAmount() {
        return loan.getOutstanding() + calcUnusedRiskAmount();
    }
}

```

Thinking I'm now done, I inspect the results of the refactoring. I wonder, "Is there anything left to simplify or communicate better?" "Is there any duplication to remove?" The duration calculations for the three strategies execute a similar formula: find the difference in time between two dates, divide them by the number of milliseconds in a day, and divide that by 365. That formula is being duplicated! To remove the duplication, I apply *Pull Up Method* [F]:

```
public abstract class CapitalStrategy...
    private static final int DAYS_PER_YEAR = 365;
    protected double calcDuration(Date start, Date end) {
        return ((end.getTime() - start.getTime()) / MILLIS_PER_DAY) / DAYS_PER_YEAR;
    }

public class TermLoanCapital extends CapitalStrategy...
    protected double duration() {
        return calcDuration(loan.getStart(), loan.getMaturity());
    }

public class RevolverCapital extends CapitalStrategy {
    protected double duration() {
        return calcDuration(loan.getStart(), loan.getExpiry());
    }
}

public class RCTLCapital extends CapitalStrategy...
    protected double duration() {
        double revolverDuration = calcDuration(loan.getStart(), loan.getExpiry());
        double termDuration = calcDuration(loan.getExpiry(), loan.getMaturity());
        return revolverDuration + termDuration;
    }
}
```

I compile, run the tests and everything is good. Now, for the moment, I'm done.

Replace Implicit Tree with Composite

You implicitly form a tree structure, using a primitive representation, such as a String

Replace your primitive tree representation with a Composite

```
String orders = "<orders>";
orders += "<order number='123'>";
orders += "<item number='x1786'>";
orders += "carDoor";
orders += "</item>";
orders += "</order>";
orders += "</orders>";
```



```
TagNode orders = new TagNode("orders");
    TagNode order = new TagNode("order");
    order.addAttribute("number", "123");
    orders.add(order);
    TagNode item = new TagNode("item");
    item.addAttribute("number", "x1786");
    item.addValue("carDoor");
    order.add(item);
String xml = orders.toString();
```

Motivation

One problem with implicit tree construction is the tight coupling between the code that builds the tree and how the tree is represented. Consider the example above, in which an XML document is built using a String. The nodes on the built XML tree and the way that they are formatted are combined in one place. While that may seem simple, it actually makes it harder to change the tree's representation and forces every programmer to remember every tree representation rule: like using single quotes for attributes or closing all open tags. I've seen programmers fight many bugs that originated in primitive tree formatting mistakes.

A Composite encapsulates how a tree is represented. This means that a client only needs to tell a Composite what to add to a tree and where to add it. When a client needs a representation of the tree, it can ask the Composite to render it. This simpler arrangement leads to less error-prone code.

But this doesn't mean that you should always avoid using primitive tree construction. What if your system doesn't create many trees? In that case, why go to the trouble of creating a Composite when some primitive tree construction code would do? If you later find that you or others are creating more trees, you can refactor to a solution that simplifies the tree construction perhaps by decoupling the tree-building code from the tree-representation code.

The choice may also involve your development speed. On a recent project, I was tasked with generating an HTML page from XML data using an XSLT processor. For this task, I needed to generate an XML tree that would be used in the XSLT transformation. I knew I could use a Composite to build that tree, but I instead choose to build it with a String. Why? Because I was more interested in going fast and facing every technical hurdle involved in doing the XSLT transformation than I was in producing refined XML tree construction code. When I completed

the XSLT transformation, I went back to refactor the primitive tree construction code to use a Composite, since that code was going to be emulated in many areas of the system.

Communication	Duplication	Simplicity
The best tree-construction code communicates the structure of a tree without overwhelming readers with unnecessary tree-construction details. Primitive tree-construction code exposes too many details. Trees composed using Composite communicate better by hiding tedious and repetitive tree-construction tasks.	Code that manually builds a tree often repeats the same set of steps: format a node, add the node to the tree and balance the node with a corresponding node or some such thing. Composite-constructed trees minimize duplication by encapsulating repetitive instructions, like formatting nodes and tree-construction mechanics.	It's easier to make mistakes building trees manually than it is to build trees using Composite. Manually-constructed trees must ensure that child nodes are added correctly – for example, a tag in an XML tree must have a corresponding end tag. By knowing how to construct themselves, Composite-constructed trees are simpler.

Mechanics

1. Identify the primitive tree-construction code you'd like to refactor.
2. Identify node types for your new *Composite*. Keep it simple: test-first design one or more concrete node types and don't worry about creating an abstract node type (you may not need one). Create a method to validate the contents of your budding *Composite*.
3. Give your nodes the ability to have children. Do not give nodes the ability to remove children if your application only adds nodes and never removes them.
 - ✓ Compile and test.
4. If needed, give clients a way to set attributes on nodes.
 - ✓ Compile and test.
5. Replace the original tree-construction code with calls to your new Composite.
 - ✓ Compile and test.

Example

1. We'll begin with the XML example from the code sketch above:

```
String orders = "<orders>";
orders += "<order number='123'>";
orders += "<item number='x1786'>";
orders += "carDoor";
orders += "</item>";
orders += "</order>";
orders += "</orders>";
```

2. In this case, every node in the tree has an open tag (“<orders>”) and close tag (“</orders>”). While some of the nodes have attributes and values, I identify just one node type that we need to produce a Composite version of this tree. I test-first design a node type called TagNode, give this class a way to set its name and create a toString() method to return the resulting XML:

```
public void testOneNodeTree() {
    String expectedResult =
        "<orders>" +
        "</orders>";
    TagNode orders = new TagNode("orders");
    assertXMLequals("xml comparison", expectedResult, orders.toString());
}

public class TagNode {
    private String tagName;

    public TagNode(String name) {
        tagName = name;
    }
    public String toString() {
        String result = new String();
        result += "<" + tagName + ">";
        result += "</" + tagName + ">";
        return result;
    }
}
```

3. Next, I give TagNode the ability to have children.

```
public void testAddingChildrenToTree() {
    String expectedResult =
        "<orders>" +
        "    <order>" +
        "        <item>" +
        "            </item>" +
        "        </order>" +
        "</orders>";
    TagNode orders = new TagNode("orders");
    TagNode order = new TagNode("order");
    TagNode item = new TagNode("item");
    orders.add(order);
    order.add(item);
    assertXMLequals("adding children", expectedResult, orders.toString());
}

public class TagNode {
    private String tagName;
    private List children = new ArrayList();
    public TagNode(String name) {
        tagName = name;
    }
    public void add(TagNode childNode) {
        children.add(childNode);
    }
    public String toString() {
        String result = new String();
        result += "<" + tagName + ">";
        Iterator it = children.iterator();
        while (it.hasNext()) {
            TagNode node = (TagNode)it.next();
            result += node.toString();
        }
        result += "</" + tagName + ">";
        return result;
    }
}
```

4. Now the Composite must be extended to support XML attributes or values or both. Again, I do this by letting my test code drive the development process:

```
public void testTreeWithAttributesAndValues() {
    String expectedResult =
        "<orders>" +
```

```

        "<order>" +
        "<item number='12660' quantity='1'>" +
        "Dog House" +
        "</item>" +
        "<item number='54678' quantity='1'>" +
        "Bird Feeder" +
        "</item>" +
        "</order>" +
        "</orders>";
    TagNode orders = new TagNode("orders");
    TagNode order = new TagNode("order");
    TagNode item1 = new TagNode("item");
    item1.addAttribute("number", "12660");
    item1.addAttribute("quantity", "1");
    item1.setValue("Dog House");
    TagNode item2 = new TagNode("item");
    item2.addAttribute("number", "54678");
    item2.addAttribute("quantity", "1");
    item2.setValue("Bird Feeder");
    orders.add(order);
    order.add(item1);
    order.add(item2);
    assertXMLequals("attributes&values", expectedResult, orders.toString());
}

public class TagNode {
    private String tagName;
    private String tagValue = "";
    private String attributes = "";
    private List children = new ArrayList();
    public TagNode(String name) {
        tagName = name;
    }
    public void add(TagNode childNode) {
        children.add(childNode);
    }
    public void addAttribute(String name, String value) {
        attributes += (" " + name + "=" + value + "");
    }
    public void addValue(String value) {
        tagValue = value;
    }
    public String toString() {
        String result = new String();
        result += "<" + tagName + attributes + ">";
        Iterator it = children.iterator();
        while (it.hasNext()) {
            TagNode node = (TagNode)it.next();
            result += node.toString();
        }
        if (!tagValue.equals(""))
            result += tagValue;
        result += "</" + tagName + ">";
        return result;
    }
}

```

5. In the final step, I replace the original primitive tree-construction code with the Composite code, compile and test:

```

TagNode orders = new TagNode("orders");
TagNode order = new TagNode("order");
order.addAttribute("number", "123");
orders.add(order);
    TagNode item = new TagNode("item");
    item.addAttribute("number", "x1786");
    item.addValue("carDoor");
    order.add(item);

```

Encapsulate Composite with Builder

Your Composite code exposes too many details, forcing clients to create, format, add and remove nodes and handle validation logic

*Encapsulate the Composite with a simpler,
more intention-revealing Builder*

```
TagNode orders = new TagNode("orders");
TagNode order = new TagNode("order");
order.addAttribute("number", "123");
orders.add(order);
TagNode item = new TagNode("item");
item.addAttribute("number", "x1786");
item.addValue("carDoor");
order.add(item);
String xml = orders.toString();
```



```
XMLBuilder orders = new XMLBuilder("orders");
orders.addBelow("order");
orders.addAttribute("number", "123");
orders.addBelow("item");
orders.addAttribute("number", "x1786");
orders.addValue("carDoor");
String xml = orders.toString();
```

Motivation

I'm always interested in simplifying client code: I want it to read as clearly as English. So when it comes to creating really simple tree-construction code, I like the Builder pattern even better than the Composite pattern. Builders give clients a clean and easy-to-use interface while hiding details about how the nodes of a Composite are hooked together and what accompanying steps must take place during construction.

If you study a typical piece of client code that creates some tree structure, you'll often find node creation and setup logic mixed together with tree creation and validation logic. A Builder-based alternative can simplify such code by taking on the burden of node creation and tree validation logic and let client code concentrate on what is important: building the tree. The result of refactoring to Builder is often simpler, more intention-revealing client code.

I use Builders a lot with XML. XML documents represent trees, so they work well with both the Composite and Builder patterns. But Composite-only solutions for creating an XML tree expose too many details. XML Builders, by contrast, offer a nice way to have your cake and eat it too: clients talk to a simple XML Builder interface, while the XML Builder itself relies on a Composite for representing the XML tree. The example below will show you how this is done. In addition, I've included an extended example which shows how an XML Builder was updated to implement and encapsulate performance logic used in rendering a Composite of XML nodes to a string.

Communication	Duplication	Simplicity
Client code that creates a tree needs to communicate the essence of the activity: what is added to the tree, and where it is added. A Composite solution doesn't communicate this clearly because it exposes too many details. By handling the tree-construction details, Builders enable client code to communicate clearly.	Composite-based tree-construction code is filled with calls to create new nodes and add them to trees. Builder code removes this duplication by handling node creation and simplifying how nodes are added to a tree.	With a Composite, a client must know <i>what</i> , <i>where</i> and <i>how</i> to add items to a tree. With a Builder, a client needs to know only <i>what</i> and <i>where</i> to add to the tree; the Builder takes care of the rest. Builders often simplify client code by handling the mechanics of tree construction.

Mechanics

1. Identify the *Composite* that you would like to encapsulate.
2. Create a new *Builder* class:
 - Give the new class a private instance variable for the encapsulated *Composite*.
 - Initialize the *Composite* in a constructor.
 - Create a method to return the results of doing a build.
3. Create intention-revealing methods on your *Builder* for every type of node that gets added to your *Composite*. These methods will add new nodes to an inner *Composite* and keep track of the state of the tree.

You may create additional methods to let users set attributes on nodes, or you can let users add new nodes and set attributes on them using one convenient method call.

4. Replace the tree-construction *Composite* calls with calls to the *Builder*.

✓ Compile and test..

Example

1. We'll begin with the Composite code that was shown in the code sketch above. As I study this code, I realize that it contains more detail than it needs to:

```

TagNode orders = new TagNode("orders");
TagNode order = new TagNode("order");
order.addAttribute("number", "123");
orders.add(order);
    TagNode item = new TagNode("item");
    item.addAttribute("number", "x1786");
    item.addValue("carDoor");
    order.add(item);

```

2. I define an XMLBuilder class, encapsulate the original Composite, initialize it and write a toString() method to obtain the results of a build. I do this all from test code, which helps me confirm that my new class produces correct XML.

```

public void testOneElementTree() {
    String expected =
        "<orders>" +
        "</orders>";
    XMLBuilder builder = new XMLBuilder("orders");

```

```
    assertXMLequals("one element tree", expected, builder.toString());
}
```

Now, my Builder looks like this:

```
public class XMLBuilder {
    private TagNode root;
    public XMLBuilder(String rootName) {
        root = new TagNode(rootName);
    }
    public String toString() {
        return root.toString();
    }
}
```

3. Next, I create methods for every type of node that gets added to the Composite. In this case it's trivial: there are only TagNodes. But I still have to consider the different ways in which clients will add nodes to the inner Composite. I begin with the case of adding nodes as children of parent nodes:

```
public void testAddBelow() {
    String expected =
        "<orders>" +
        "  <order>" +
        "    <item>" +
        "  </item>" +
        "  </order>" +
        "</orders>";
    XMLBuilder builder = new XMLBuilder("orders");
    builder.addBelow("order");
    builder.addBelow("item");
    assertXMLequals("adding below", expected, builder.toString());
}
```

This leads to the creation of the addBelow() method, along with a few changes to the XMLBuilder class:

```
public class XMLBuilder {
    private TagNode root;
    private TagNode current;
    public XMLBuilder(String rootName) {
        root = new TagNode(rootName);
        current = root;
    }
    public void addBelow(String child) {
        TagNode childNode = new TagNode(child);
        current.add(childNode);
        current = childNode;
    }
    public String toString() {
        return root.toString();
    }
}
```

Next I must enable the XMLBuilder to add a node at the same level as an existing node (i.e., not as a child, but as a sibling). This leads to more test and XMLBuilder code:

```
public void testAddBeside() {
    String expected =
        "<orders>" +
        "  <order>" +
        "    <item>" +
        "  </item>" +
        "  <item>" +
        "  </item>" +
        "</orders>" +

```

```

    "</orders>";
    XMLBuilder builder = new XMLBuilder("orders");
    builder.addBelow("order");
    builder.addBelow("item");
    builder.addBeside("item");
    assertXMLEquals("adding beside", expected, builder.toString());
}

public class XMLBuilder {
    private TagNode root;
    private TagNode current;
    private TagNode parent;
    public XMLBuilder(String rootName) {
        root = new TagNode(rootName);
        current = root;
        parent = root;
    }
    public void addBelow(String child) {
        TagNode childNode = new TagNode(child);
        current.add(childNode);
        parent = current;
        current = childNode;
    }
    public void addBeside(String sibling) {
        TagNode siblingNode = new TagNode(sibling);
        parent.add(siblingNode);
        current = siblingNode;
    }
    public String toString() {
        return root.toString();
    }
}

```

I continue on this approach until I have a working Builder that satisfies all of my tests. In some cases, adding new behavior to the XMLBuilder is trivial, since it merely requires delegating calls to the inner Composite. For example, here is how XML attributes are implemented:

```

public void testAddBelowWithAttribute() {
    String expected =
        "<orders>" +
        "  <order number='12345' quantity='2'>" +
        "    </order>" +
        "</orders>";
    builder = createBuilder("orders");
    builder.addBelow("order");
    builder.addAttribute("number", "12345");
    builder.addAttribute("quantity", "2");
    assertXMLEquals("built xml", expected, builder.toString());
}

public class XMLBuilder . . .
    public void addAttribute(String name, String value) {
        current.addAttribute(name, value);
    }
}

```

4. Now it is time to replace the original client code that used the Composite with the XMLBuilder. I do this one line at a time, removing some lines and rewriting others. The final code makes no references to the now encapsulated Composite, TagNode.

```

XMLBuilder orders = new XMLBuilder("orders");
orders.addBelow("order");
orders.addAttribute("number", "123");
orders.addBelow("item");
orders.addAttribute("number", "x1786");
orders.addValue("carDoor");

```

Notice how the calls to the `XMLBuilder` are generic – the methods and data passed to them reveal nothing about the underlying structure of the tree. Should we need to work with a variety of Builders, we won't have to change very much client code.

Extended Example

I could not resist telling you about a performance improvement that was made to the above-mentioned `XMLBuilder` class, since it reveals the elegance and simplicity of the Builder pattern. Some of my colleagues at a company called Evant had done some profiling of our system and they'd found that a `StringBuffer` used by the `XMLBuilder`'s encapsulated composite (`TagNode`) was causing performance problems. This `StringBuffer` is used as a Collecting Parameter – it is created and then passed to every node in a composite of `TagNodes` in order to produce the results returned from calling `TagNode`'s `toString()`. To see how this works, see the example in *Move Accumulation to Collecting Parameter (94)*.

The `StringBuffer` that was being used in this operation was not instantiated with any particular size, which means that as more and more XML is added to the `StringBuffer`, it must automatically grow when it can no longer hold all its data. That's fine, since the `StringBuffer` class was written to be able to automatically grow. But there is a performance penalty one pays when you allow a `StringBuffer` to automatically grow: i.e. when it has to grow, it has work to do to transparently increase its size. That performance penalty in the Evant system was not acceptable and so the team needed to make an improvement.

The solution was to know what size the `StringBuffer` needed to be before instantiating it, and then to instantiate it with the proper size so that it would not need to grow. How could we compute this size? Easy. As each node gets added to an XML tree via an `XMLBuilder`, the builder increments a buffer size based on the size of the strings in the node. Then the final computed buffer size could be used when instantiating the `StringBuffer`. Let's see how this was implemented.

As usual, we start by writing a test. The test below will build an XML tree by making calls to an `XMLBuilder`, then it will obtain the size of the resulting XML string returned by the builder and finally, it will compare the size of the string with the computed buffer size for use by a `StringBuffer`:

```
public void testToStringBufferSize() {
    String expected =
        "<orders>" +
            "<order number='123'>" +
            "</order>" +
        "</orders>";
    builder = createBuilder("orders");
    builder.addBelow("order");
    builder.addAttribute("number", "123");

    int stringSize = builder.toString().length();
    int computedSize = ((XMLBuilder)builder).bufferSize();
    assertEquals("buffer size", stringSize, computedSize);
}
```

To pass this test and others like it, the following `XMLBuilder` attributes and methods were added or updated:

```
public class XMLBuilder {
    private int outputBufferSize;
    private static int TAG_CHARS_SIZE = 5;
    private static int ATTRIBUTE_CHARS_SIZE = 4;

    public void addAttribute(String name, String value) {
```

```

        // logic for adding an attribute to a tag
        incrementBufferSizeByAttributeLength(name, value);
    }
    public void addBelow(String child) {
        // logic for adding a Tag below another Tag
        incrementBufferSizeByTagLength(child);
    }
    public void addBeside(String sibling) {
        // logic for adding a Tag beside another Tag
        incrementBufferSizeByTagLength(sibling);
    }
    public void addBesideParent(String uncle) {
        // logic for adding a Tag beside the current Tag's parent
        incrementBufferSizeByTagLength(uncle);
    }
    public void addValue(String value) {
        // logic for adding a value to a node
        incrementBufferSizeByValueLength(value);
    }
    public int bufferSize() {
        return outputBufferSize;
    }
    private void incrementBufferSizeByAttributeLength(String name, String value) {
        outputBufferSize += (name.length() + value.length() + ATTRIBUTE_CHARS_SIZE);
    }
    private void incrementBufferSizeByTagLength(String tag) {
        int sizeOfOpenAndCloseTags = tag.length() * 2;
        outputBufferSize += (sizeOfOpenAndCloseTags + TAG_CHARS_SIZE);
    }
    private void incrementBufferSizeByValueLength(String value) {
        outputBufferSize += value.length();
    }
    protected void init(String rootName) {
        // logic for initializing the builder and root node
        outputBufferSize = 0;
        incrementBufferSizeByTagLength(rootName);
    }
}

```

The changes made to the XMLBuilder are transparent to the users of the builder, as it encapsulates this new performance logic. The only additional change must be made to the XMLBuilder's toString() method, so that it can instantiate a StringBuffer of the correct size, and pass it on to the root TagNode, which will accumulate the contents of the XML tree. To make that happen, the toString() method was changed from

```

public String toString() {
    return root.toString();
}

```

to:

```

public String toString() {
    return root.toStringHelper(new StringBuffer(outputBufferSize));
}

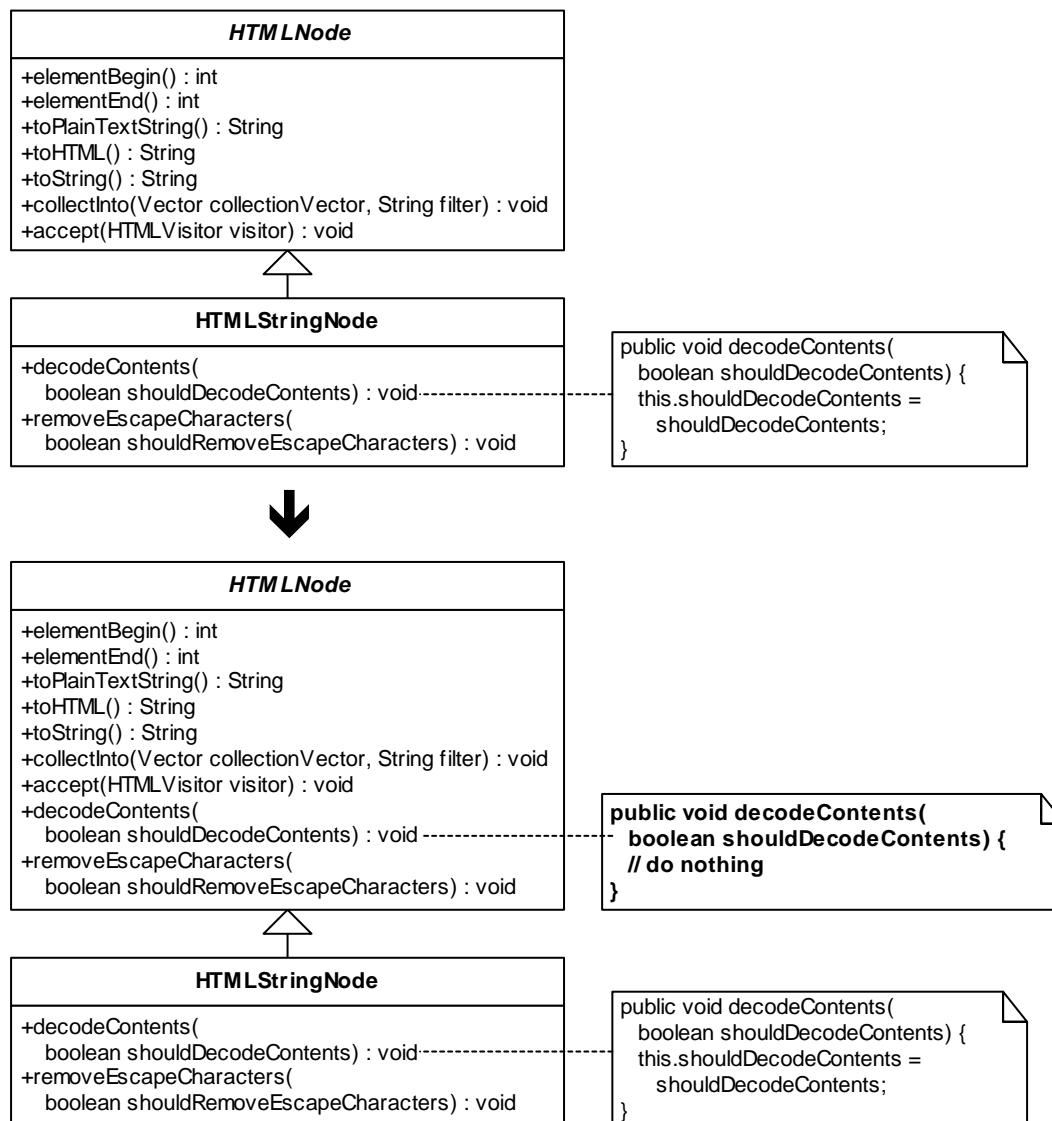
```

And that was it. The tests passed and the XMLBuilder was now significantly faster.

Form Superset Interface

You need a superclass to have the same interface as a subclass

*Copy the public, subclass-specific methods to the superclass,
altering each copied method to perform null behavior*



Motivation

This refactoring is useful on the road to forming what *Design Patterns* calls a *transparent enclosure*. Proxies and Decorators are typical transparent enclosures: they implement the same interface as the instances they enclose. In Java, an `UnmodifiableList` is a transparent enclosure, since it implements the `List` interface and is instantiated with an object that implements `List`. The object that an `UnmodifiableList` gets instantiated with is said to be the object it *encloses*. `UnmodifiableList` encloses a `List` instance *transparently* because it itself implements `List`, which is the same interface as the object it encloses. So, if a method expects a client to pass it a `List` argument, one could pass it an `ArrayList` (which implements `List`) or an `UnmodifiableList` that encloses an `ArrayList`, and the client code wouldn't know the difference.

This refactoring is necessary when you're on your way to creating a transparent enclosure for a subclass but lack the appropriate interface to do the job. The interface you need must have all of the public methods of the subclass, including those public methods that are inherited from a superclass. By first creating a superset interface in the superclass, you can then *Extract Interface* [F] to obtain the interface needed to produce the transparent enclosure.

This refactoring is critical to refactorings like *Move Embellishment to Decorator* (73) and *Move Protection to Proxy* (86).

Mechanics

1. Find a *missing method*: a public method on the subclass that isn't declared on the superclass.
2. Add a copy of the *missing method* to the superclass, modifying its body to perform null behavior.

✓ Compile.
3. Repeat steps 1 and 2 until the superclass and subclass have the same interface.

Test that all code related to the superclass works as expected.

Example

I need to write a transparent enclosure for a class called `HTMLStringNode`, which implements the following public methods:

```
public class HTMLStringNode extends HTMLNode...
    public String toPlainTextString() {
    public String toHTML()...
    public String toString()...
    public void collectInto(Vector collectionVector, String filter)...
    public void accept(HTMLVisitor visitor)...
    public void decodeContents(boolean shouldDecodeContents)...
    public void removeEscapeCharacters(boolean shouldRemoveEscapeCharacters)...
```

1. `HTMLStringNode` inherits most of its public methods from `HTMLNode`, except for

```
    public void decodeContents(boolean shouldDecodeContents)...
    public void removeEscapeCharacters(boolean shouldRemoveEscapeCharacters)...
```

I'll work with the `decodeContents(...)` method first. Here's its implementation in `HTMLStringNode`:

```
public class HTMLStringNode extends HTMLNode...
    public void decodeContents(boolean shouldDecodeContents) {
        this.shouldDecodeContents = shouldDecodeContents;
    }
```

2. I copy the `decodeContents(...)` method to `HTMLNode`, making its method body do nothing:

```
public class HTMLNode...
    public void decodeContents(boolean shouldDecodeContents) {
    }
```

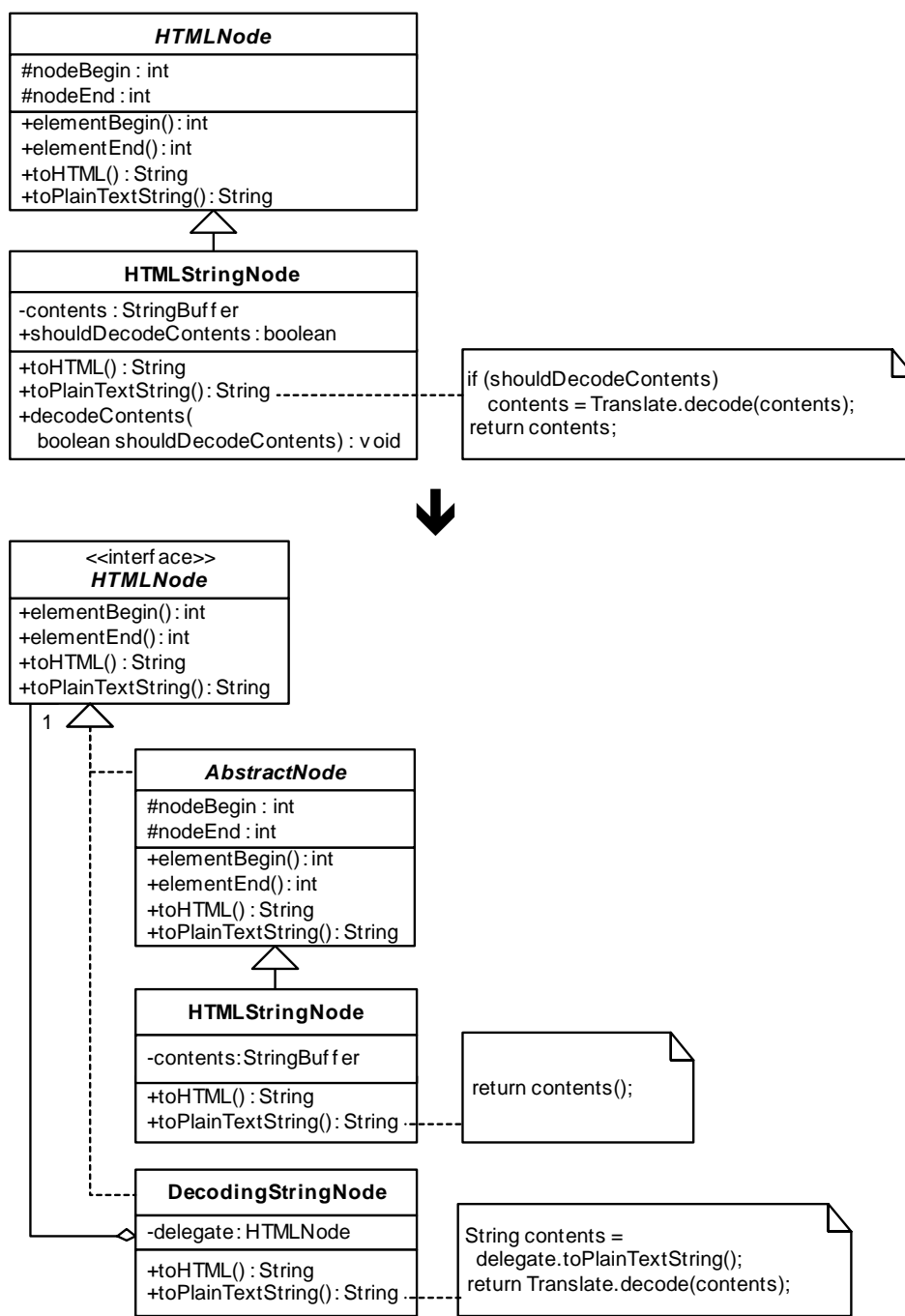
3. I repeat step 2 for the method, `removeEscapeCharacters(...)` and then test that all code in the system still works.

I'm now at a point where I can *Extract Interface* [F] from `HTMLNode` in order to produce an interface that may be used to create a transparent enclosure of an `HTMLStringNode`. For a complete example of creating a transparent enclosure, see *Move Embellishment to Decorator* (73).

Move Embellishment to Decorator

A class contains code that provides an embellishment to its core responsibility

Move the embellishment to a Decorator

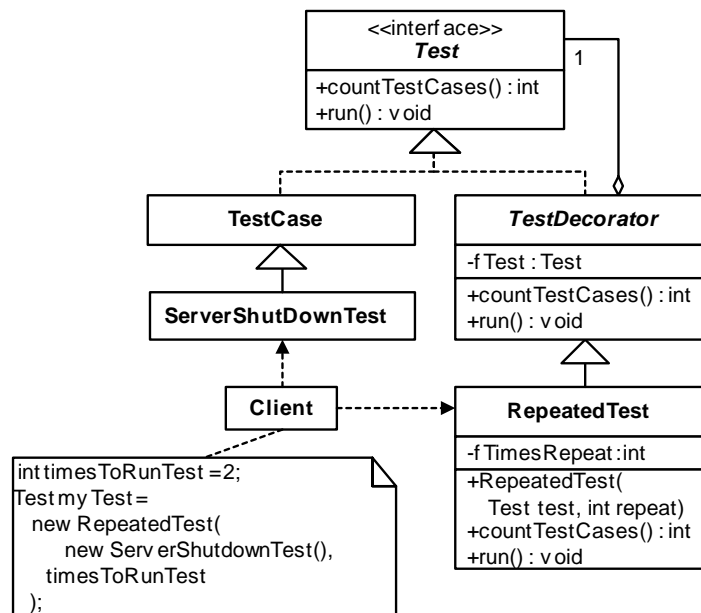


Motivation

When new features in a system are needed, programmers often add new code to old classes. Such new code often embellishes the core responsibility or primary behavior of an existing class. The trouble with these embellishments is that they usually complicate their host classes: typical embellishments add new fields, new methods and new logic, all of which is special-case behavior that only needs to be executed some of the time.

The Decorator pattern offers a good remedy: place each embellishment in its own class and let that class wrap the type of object it needs to embellish so that clients may wrap the embellishment around objects at runtime, when special-case behavior is needed.

The testing framework, JUnit, provides a good example. The framework makes it easy to write and run tests. Each test is an object of type `TestCase` and there's an easy way to tell the framework to run all of your `TestCase` objects. But if you want to run one test multiple times, there's no embellishment within `TestCase` to do so. For that "extended feature" you need to decorate a `TestCase` object with a `RepeatedTest` Decorator:



Decorator is not a pattern you'd refactor to if you wanted to move an embellishment out of a class with 82 public methods. Why? Because a Decorator must be what *Design Patterns* calls a *transparent enclosure*: it must implement *all* of the public methods of the objects it decorates (and that would require a lot of useless code for a class with 82 public methods). Transparent enclosures wrap other classes in a way that is transparent to the classes being wrapped – i.e. a decoratee doesn't know it has been decorated. And since a Decorator and its decoratee share the same interface, Decorator(s) are transparent to client code that uses them (provided the client doesn't rely on object identity logic, like `if (node instanceof HTMLStringNode)...`).

[finish up this section]

Communication	Duplication	Simplicity
Some code just doesn't have to be run very often. But if you lump that code in with code that is <i>always</i> necessary, you don't communicate what is and is not important. Decorators provide a way to communicate to others which specific embellishments are available for another object's core responsibility.	As logic gets more complicated, you often see code that tries to accommodate many combinations of behavior. This can lead to a lot of duplicate code. Decorators offer a better way to handle diverse combinations of behavior without duplicating code.	Code that mixes together the essential with the optional isn't as simple as code that contains solely what is essential. On the other hand, Decorators aren't always simple to use when you have to worry about the order in which you add them.

Mechanics

1. Identify an *embellished class*: a class that contains an embellishment to its core responsibility.

Not every class with an embellishment to its core responsibility will be a good candidate for being “decorated.” First ensure that the set of public methods which your Decorator will need to implement isn’t too large. Since a Decorator forms a “transparent enclosure” around the object it decorates, the set of public methods it must implement has to be the same as the set of public methods declared and inherited by the *embellished class*. If your *embellished class* declares and inherits lots of public methods, either reduce that number (by deleting, moving or changing the visibility of methods) or don’t proceed with this refactoring at all.

2. Identify or create an *enclosure type*: a class or interface that declares all of the public methods declared in and inherited by the *embellished class* ().

If you already have an *enclosure type*, it’s likely to be the superclass of the *embellished class* or an interface implemented by the *embellished class*. If you don’t already have an *enclosure type*, create one by applying *Form Superset Interface* (70) followed by *Extract Interface* [F].

3. Create a *decorator*: a class that implements or extends the *enclosure type*. Name this class based on the embellishment it adds to the *embellished class* (e.g. `RepeatedTest` is a *decorator* that add the ability to run a `Test` a specific number of times).

If this is the first decorator you’re creating, make it concrete, not abstract. Define an abstract decoratos only when you have multiple concrete Decorators that must share part of their implementation. [see GoF, page 179, item 2]

4. Add a non-public *delegate* field on the *decorator*, declaring the field’s type to be the *enclosure type*. Add a constructor to the *decorator* that accepts a parameter, typed as the *enclosure type*, and sets the *delegate* equals to the value of the parameter.

You can pass additional parameters to your *decorator*’s constructor to be used to perform that *decorator*’s embellishment (e.g. the “repeat” parameter in `public RepeatedTest(Test test, int repeat)`).

5. For each method defined by the *decorator*, forward each method call to the *delegate*.
 - ✓ Compile and test that a decorated instance of the *embellished class* works just like an *embellished class* instance that has been directed to perform its embellishment.
6. You'll now add the embellishment logic in the *embellished class* to the *decorator* (being sure to leave the embellishment logic in the *embellished class*). This step may be implemented by applying *Extract Method* [F] to get the embellishment into a single method, copying that method to the *decorator* and calling the copied method from the *decorator* method(s) that must add the embellishment before or after their *delegate* call.
 - ✓ Compile and test by creating a decorated instance of the *embellished class* and verifying that any updated decorator methods work just like an *embellished class* instance that has been directed to perform its embellishment.
7. Update all client code which requires the embellishment to obtain it by means of a decorated instance of the *embellished class*. It is best to make the type of the decorated instance be the *enclosure type*.
 - ✓ Compile and test.
8. Remove the embellishment from the *embellished class*, including any associated configuration logic. By "configuration logic" I mean a public method that turns the embellishment on or off. If such a method exists, remove it from the *embellished class*, the *enclosure type*, the *decorator* and any client code that calls it.
 - ✓ Compile and test that the removed code didn't break anything.

If you repeat these mechanics to create additional decorators, factor out common decorator logic into an abstract decorator class and test that objects decorated with multiple decorators work correctly. It can be difficult to support multiple decorators, since one decorator can interfere with another (e.g. a decorator that encrypts data can interfere with a decorator that filters out certain words if the encryption happens prior to the filtering). It is best to have decorators be so independent of each other that they can be added to objects in any combination. In practice, that may not be possible, in which case you can write Creation Methods to give access to objects decorated in various ways.

Example

The HTML parser (<http://sourceforge.net/projects/htmlparser>) is an open-source tool that allows programs to see the contents of HTML files as specific HTML objects. When the parser encounters tag data or strings sandwiched between tag data, it translates what it finds into the appropriate HTML objects, like `HTMLTag`, `HTMLStringNode`, `HTMLEndTag`, `HTMLImageTag` and so forth. The parser is frequently used to

- translate the contents of one HTML file to another
- report information about a piece of HTML
- verify the contents of HTML

The Decorator refactoring example we'll look at concerns the parser's `HTMLStringNode` class. Instances of this class are created at runtime when the parser finds chunks of text sandwiched between tags. For example, consider this HTML:

```
<BODY>
  This text will be recognized as an HTMLStringNode
</BODY>
```

Given the above, the parser creates the following objects at runtime:

```
HTMLTag (for the <BODY> tag)
  HTMLStringNode (for the String, "This text will be recognized as an HTMLStringNode")
HTMLEndTag (for the </BODY> tag)
```

There are a few ways to examine the contents of an HTML object: one can obtain the object's plain text representation using `toPlainTextString()` or one can obtain the object's HTML using `toHTML()`. `HTMLStringNode` returns the same value whether you call `toPlainTextString()` or `toHTML()`.

A common embellishment to `HTMLStringNode` involves decoding "numeric or character entity references" found in `HTMLStringNodes`. Typical character reference decodings include:

```
&amp; decoded to &
&lt; decoded to <
&rt; decoded to >
```

The parser's `Translate` class contains the method, `decode(String dataToDecode)`, which can decode a comprehensive set of numeric and character entity references. Such decoding is an embellishment that is often applied to `HTMLStringNodes`, after they've been found by the parser. For example, consider the following code which iterates through a collection of `HTMLNodes`, decoding the nodes that are instances of `HTMLStringNode`:

```
public void testDecodingAmpersand() throws Exception {
    String WORKSHOP_TITLE =
        "<H1>The Testing &amp; Refactoring Workshop</H1>";

    String DECODED_WORKSHOP_TITLE =
        "<H1>The Testing & Refactoring Workshop</H1>";

    StringBuffer decodedContent = new StringBuffer();
    HTMLParser parser = HTMLParser.createParser(WORKSHOP_TITLE);
    HTMLEnumeration nodes = parser.elements();
    while (nodes.hasMoreNodes()) {
```

```

        HTMLNode node = nodes.nextNode();
        if (node instanceof HTMLStringNode) {
            HTMLStringNode stringNode = (HTMLStringNode)node;
            decodedContent.append(
                Translate.decode(stringNode.toPlainTextString()); // decoding step
            }
        if (node instanceof HTMLTag)
            decodedContent.append(node.toHTML());
    }
    assertEquals("decoded content",
        DECODED_WORKSHOP_TITLE,
        decodedContent.toString()
    );
}

```

While studying code that made use of the parser, it was clear that clients needed to decode character and numeric references in HTMLStringNodes some of the time. Yet these clients always had to perform the decoding themselves, using the same process of iterating nodes, finding nodes that were HTMLStringNodes and decoding them. I decided it was better for the parser to deal with the decoding of HTMLStringNodes when asked to do so by a client.

I thought of several ways to go about this refactoring, and then settled on a dead simple approach: add the decoding embellishment directly to the HTMLStringNode and observe how that code looks afterwards. Truth be told, I knew this implementation wouldn't live long, but I wanted to see how far I could push the code before it was begging for a Decorator implementation. Using test-driven development, I added the decoding embellishment to HTMLStringNode. This work involved changing the HTMLParser class, changing the StringParser class (which actually instantiates HTMLStringNodes) and changing HTMLStringNode.

The HTMLParser was changed to include a flag for toggling HTMLStringNode decoding on or off:

```

public class HTMLParser...
    private boolean shouldDecodeStringNodes = false;

    public void setStringNodeDecoding(boolean shouldDecodeStringNodes) {
        this.shouldDecodeStringNodes = shouldDecodeStringNodes;
    }

    public boolean shouldDecodeStringNodes() {
        return shouldDecodeStringNodes;
    }

```

The StringParser was changed to consult the HTMLParser's shouldDecodeStringNodes flag when configuring newly instantiated HTMLStringNodes:

```

public class StringParser...
    public HTMLNode find(HTMLReader reader, String input, int position)...
        HTMLStringNode stringNode =
            new HTMLStringNode(textBuffer, textBegin, textEnd);
        if (reader.getParser().shouldDecodeStringNodes())
            stringNode.decodeContents(true);
        return stringNode;

```

Finally, the HTMLStringNode class was changed to support the decoding embellishment:

```

public class HTMLStringNode...
    protected StringBuffer textBuffer;
    private boolean shouldDecodeContents = false;

    public void decodeContents(boolean shouldDecodeContents) {
        this.shouldDecodeContents = shouldDecodeContents;
    }

```

```

    }

    public String toPlainTextString() {
        return nodeContents();
    }

    public String toHTML() {
        return nodeContents();
    }

    private String nodeContents() {
        String result = textBuffer.toString();
        if (shouldDecodeContents)
            result = Translate.decode(result);
        return result;
    }

```

With such code in place, client code could now obtain decoded `HTMLStringNodes` by simply telling the parser to do `HTMLStringNode` decoding:

```

public void testDecodingAmpersand() throws Exception {
    String WORKSHOP_TITLE =
        "<H1>The Testing & Refactoring Workshop</H1>";

    String DECODED_WORKSHOP_TITLE =
        "<H1>The Testing & Refactoring Workshop</H1>";

    StringBuffer decodedContent = new StringBuffer();
    HTMLParser parser = HTMLParser.createParser(WORKSHOP_TITLE);
    parser.setStringNodeDecoding(true);
    HTMLEnumeration nodes = parser.elements();

    while (nodes.hasMoreNodes())
        decodedContent.append(nodes.nextNode().toHTML());

    assertEquals("decoded content",
        DECODED_WORKSHOP_TITLE,
        decodedContent.toString()
    );
}

```

This decoding feature in the parser didn't bloat the code unduly. However, once you support one embellishment, it's often useful to support others. When looking over more client code that used the parser (including my own), I found that it was also common to remove escape characters (like `\n` for newline, `\t` for tabs) from `HTMLStringNodes`. So I decided to add an escape-character-removal feature to the parser as well, which meant adding another `HTMLParser` flag (`shouldRemoveEscapeCharacters`), updating the `StringParser` to configure `HTMLStringNodes` to remove escape characters and adding the following escape character removal code to `HTMLStringNode`:

```

public class HTMLStringNode...
    private boolean shouldRemoveEscapeCharacters = false;

    public String toPlainTextString() {
        return nodeContents();
    }

    public String toHTML() {
        return nodeContents();
    }

    private String nodeContents() {
        String result = textBuffer.toString();
        if (shouldDecodeContents)
            result = Translate.decode(result);

```

```

    if (shouldRemoveEscapeCharacters)
        result = HTMLParserUtils.removeEscapeCharacters(result);
    return result;
}

```

While these embellishments were immediately useful, I knew that the implementation was not ideal, since every new embellishment would require special-case logic in `HTMLStringNode`, a new toggle switch in `HTMLParser` and conditional configuration logic in the `StringParser`, which instantiates the `HTMLStringNodes`. This was the right time to simplify this design by refactoring to Decorator. Each `HTMLStringNode` embellishment could be moved into its own `HTMLStringNode` Decorator and the various Decorators could be configured at runtime to provide whatever embellished behavior was necessary. The example below shows how my initial `HTMLStringNode` decoding code was refactoring to be performed using a `DecodingStringNode` decorator.

1. Step one involves identifying the *embellished class*, which in this case is `HTMLStringNode`. Before I proceed, I want to check whether Decorator will make a good fit here. I do this by checking if `HTMLStringNode` is primitive enough (i.e. it doesn't hold a lot of state and doesn't inherit and declare many public methods) to make decoration viable. Since a decorator will have to implement the superset of public methods declared by `HTMLStringNode` and its superclass, `HTMLNode`, I study both classes and find that, combined, they declare 10 public methods and share 2 fields. I'm not worried about fields, but the 10 public methods is worrisome: since it can be awkward to implement Decorator when there is a wide interface to support.

So I check if I can move some of `HTMLNode`'s methods to another class, delete unnecessary methods or make some methods non-public (after confirming that they didn't need to be public in the first place). In this case I find that I can move two special-case print methods to a utility class and delete one method that had been the result of someone's experiment (and which had not been cleaned up from the code, as yet). So that got the count down to 7 public methods, which seemed manageable. I decided to proceed with the refactoring.

2. I now need to create an *enclosure type*, a class or interface that declares all public method of `HTMLStringNode` and whatever public methods it inherits. A good *enclosure type* won't contain fields (i.e. state). `HTMLStringNode`'s superclass, `HTMLNode`, is not a good enclosure type because it contains two primitive `int` fields, `nodeBegin` and `nodeEnd`. Why does it matter if a class contains fields? Decorators add behavior to the objects they decorate, but they don't need duplicate copies of the fields in the objects they decorate. In this case, since `HTMLStringNode` already inherits `nodeBegin` and `nodeEnd` from `HTMLNode`, Decorators of `HTMLStringNode` don't also need to inherit those fields.

So I rule out `HTMLNode` as the *enclosure type*. This means that I need to create a Java interface for the *enclosure type*. And since this interface must contain all the public methods from both `HTMLStringNode` and `HTMLNode`, I apply *Form Superset Interface* (70). That refactoring alters the `HTMLNode` class to include the following public methods (bold font is used to show methods that originated from `HTMLStringNode`):

```

public abstract class HTMLNode...
    public int elementBegin()...
    public int elementEnd()...
    public String toPlainTextString() {
    public String toHTML()...
    public String toString()...
    public void collectInto(Vector collectionVector, String filter)...
    public void accept(HTMLVisitor visitor)...
    public void decodeContents(boolean shouldDecodeContents)...
    public void removeEscapeCharacters(boolean shouldRemoveEscapeCharacters)...

```


I apply *Extract Interface* [F] on `HTMLNode` to produce the *enclosure type*. Since I would like the *enclosure type* to be called `HTMLNode`, I rename the original `HTMLNode` to `AbstractNode` and rename the *enclosure type* to `HTMLNode`.

3. I can now create a *decorator*. This class will implement the *enclosure type*, `HTMLNode`, so that it can transparently enclose an `HTMLStringNode`. I decide to call the *decorator*, `DecodingStringNode`, based on the embellishment it will provide to `HTMLStringNodes`. Here's the class:

```
public class DecodingStringNode implements HTMLNode {
    public String toPlainTextString() {
        return null;
    }
    public String toHTML() {
        return null;
    }
    public void collectInto(Vector collectionVector, String filter) {
    }
    public int elementBegin() {
        return 0;
    }
    public int elementEnd() {
        return 0;
    }
    public void accept(HTMLVisitor visitor) {
    }
    public void decodeContents(boolean shouldDecodeContents) {
    }
    public void removeEscapeCharacters(boolean shouldRemoveEscapeCharacters) {
    }
}
```

4. I must now give the `DecodingStringNode` a field for storing the object it will decorate, and let this field be set via a constructor:

```
public class DecodingStringNode implements HTMLNode...
    private HTMLNode delegate;

    public DecodingStringNode(HTMLNode delegate) {
        this.delegate = delegate;
    }
}
```

The constructor only takes one argument. In some cases, decorators will need more information passed into them in order to implement their embellishment. For example, an `EscapeCharacterRemover` might need a list of escape character to use when remove escape characters from `HTMLStringNodes`.

5. Now I make the *decorator* call its *delegate* for each of its public methods:

```
public class DecodingStringNode implements HTMLNode {
    private HTMLNode delegate;

    public DecodingStringNode(HTMLNode delegate) {
        this.delegate = delegate;
    }
    public String toPlainTextString() {
        return delegate.toPlainTextString();
    }
    public String toHTML() {
        return delegate.toHTML();
    }
}
```

```

    public void collectInto(Vector collectionVector, String filter) {
        delegate.collectInto(collectionVector, filter);
    }
    public int elementBegin() {
        return delegate.elementBegin();
    }
    public int elementEnd() {
        return delegate.elementEnd();
    }
    public void accept(HTMLVisitor visitor) {
        delegate.accept(visitor);
    }
    public void decodeContents(boolean shouldDecodeContents) {
        delegate.decodeContents(shouldDecodeContents);
    }
    public void removeEscapeCharacters(boolean shouldRemoveEscapeCharacters) {
        delegate.removeEscapeCharacters(shouldRemoveEscapeCharacters);
    }
}

```

That was easy. Now I'll test the changes by decorating an `HTMLStringNode` instance with a `DecodingStringNode` and verify that all public methods on the decorated class behave just like an undecorated `HTMLStringNode` instance:

```

public class DecodingStringNodeTest...
    public void testDecoratorTransparency() throws Exception {
        String ENCODED_ACTIVITY =
            "This activity & the previous one both involve...";

        String DECODED_ACTIVITY =
            "This activity & the previous one both involve...";

        HTMLParser parser = HTMLParser.createParser(ENCODED_ACTIVITY);
        parser.setStringNodeDecoding(true);
        HTMLNode stringNode = parser.elements().nextNode();
        DecodingStringNode decodedStringNode =
            new DecodingStringNode(stringNode);

        assertEquals(
            "decoded plain text",
            stringNode.toPlainTextString(),
            decodedStringNode.toPlainTextString()
        );

        assertEquals(
            "decoded html",
            stringNode.toHTML(),
            decodedStringNode.toHTML()
        );

        assertEquals("begin",
            stringNode.elementBegin(),
            decodedStringNode.elementBegin()
        );

        assertEquals("end",
            stringNode.elementEnd(),
            decodedStringNode.elementEnd()
        );

        // ...and so on for all public HTMLNode methods
    }
}

```

That works fine, so I move on.

6. Now comes the fun part. I have to add the embellishment logic in `HTMLStringNode` to the *decorator*, `DecodingStringNode`, and verify that a decorated instance of the

HTMLStringNode works just like an HTMLStringNode instance that has been directed to perform its embellishment. I do the “adding” part of this step in three moves:

a) I use *Extract Method* [F] to extract the embellishment code in HTMLStringNode into its own method:

```
public class HTMLStringNode extends AbstractNode...
    public String toPlainTextString() {
        return nodeContents();
    }

    public String toHTML() {
        return nodeContents();
    }

    private String nodeContents() {
        String result = textBuffer.toString();
        if (shouldDecodeContents)
            result = decode(result);           // call to extracted method
        if (shouldRemoveEscapeCharacters)
            result = HTMLParserUtils.removeEscapeCharacters(result);
        return result;
    }

    private String decode(String result) {      // extracted method
        return Translate.decode(result);
    }
```

b) I copy the newly extracted method to DecodingStringNode:

```
public class DecodingStringNode implements HTMLNode...
    private String decode(String result) {
        return Translate.decode(result);
    }
```

c) I call DecodingStringNode’s decode method from the two methods that must add the embellishment after their delegate call: toPlainTextString() and toHTML():

```
public class DecodingStringNode implements HTMLNode...
    public String toPlainTextString() {
        return decode(delegate.toPlainTextString());
    }

    public String toHTML() {
        return decode(delegate.toHTML());
    }
```

To test that this works, I simply run the same test as I did at the end of the previous step, which is testing that a decorated HTMLStringNode behaves just like an undecorated instance. The test passes, so I move on.

7. I must now update any code that requires the decoding embellishment to obtain it by using the DecodingStringNode. StringParser requires such an update. It has code that looks like this:

```
public HTMLNode find(HTMLReader reader, String input,
                    int position, boolean ignoreStateMode)...
    ...
    HTMLStringNode stringNode = new HTMLStringNode(textBuffer, textBegin, textEnd);
    if (reader.getParser().shouldDecodeStringNodes())
        stringNode.decodeContents(true);
    if (reader.getParser().shouldRemoveEscapeCharacters())
```

```
stringNode.removeEscapeCharacters(true);
return stringNode;
```

The first thing I do here is change the type of the variable, `stringNode` to that of the *enclosure type*, `HTMLNode` and then I modify the decoding-related conditional fragment to decorate a `stringNode` with a `DecodingStringNode`:

```
public HTMLNode find(HTMLReader reader,String input,
                    int position, boolean ignoreStateMode)...
...
HTMLNode stringNode = new HTMLStringNode(textBuffer,textBegin,textEnd);
if (reader.getParser().shouldDecodeStringNodes()) {
    stringNode = new DecodingStringNode(stringNode);
    stringNode.decodeContents(true);
}
if (reader.getParser().shouldRemoveEscapeCharacters())
    stringNode.removeEscapeCharacters(true);
return stringNode;
```

I compile and successfully test that the above changes work. Notice that the call to `stringNode.decodeContents(true)` is still there. I'll be getting rid of that in the next step. Meanwhile, I continue updating other code that can now now use the `DecodingStringNode`.

8. The final step involves removing decoding logic and any associated configuration logic from `HTMLStringNode`. It is always such a joy to remove code. Here's what I remove:

```
public class HTMLStringNode extends AbstractNode...
private boolean shouldDecodeContents = false;
private boolean shouldRemoveEscapeCharacters = false;

public String toPlainTextString() {
    return nodeContents();
}

public String toHTML() {
    return nodeContents();
}

private String nodeContents() {
    String result = textBuffer.toString();
if (shouldDecodeContents)
    result = decode(result);
    if (shouldRemoveEscapeCharacters)
        result = HTMLParserUtils.removeEscapeCharacters(result);
    return result;
}

private String decode(String result){
    return Translate.decode(result);
}

public void decodeContents(boolean shouldDecodeContents){
    this.shouldDecodeContents = shouldDecodeContents;
}

public void removeEscapeCharacters(boolean shouldRemoveEscapeCharacters) {
    this.shouldRemoveEscapeCharacters = shouldRemoveEscapeCharacters;
}
```

Deleting all of that code works fine, but I'm not done having fun deleting code. The method, `decodeContents()`, exists on `HTMLStringNode`'s superclass, `AbstractNode`, on the interface, `HTMLNode`, and on the decorator, `DecodingStringNode`, which implements

HTMLNode (the enclosure type). I can now that method from all of the above classes, along with the HTMLNode interface:

```
public abstract class AbstractNode implements Serializable, HTMLNode...
    public void decodeContents(boolean shouldDecodeContents) {
    }

public interface HTMLNode...
    void decodeContents(boolean shouldDecodeContents);

public class DecodingStringNode implements HTMLNode...
    public void decodeContents(boolean shouldDecodeContents) {
        delegate.decodeContents(shouldDecodeContents);
    }
```

The compiler is mostly happy with the above deletions, except for client code that still calls the method, `decodeContents()`. The `StringParser` code from the step 7 made such a call, which I can now safely delete:

```
public HTMLNode find(HTMLReader reader, String input,
                    int position, boolean ignoreStateMode)...
    ...
    HTMLNode stringNode = new HTMLStringNode(textBuffer, textBegin, textEnd);
    if (reader.getParser().shouldDecodeStringNodes()) {
        stringNode = new DecodingStringNode(stringNode);
        stringNode.decodeContents(true); // this is no longer required
    }
    if (reader.getParser().shouldRemoveEscapeCharacters())
        stringNode.removeEscapeCharacters(true);
    return stringNode;
```

After removing all traces of client calls to `decodeContents()`, the compiler is happy, my tests run and I'm ready for a nice cold drink!

Move Protection to Proxy

A class contains protection code,
which isn't needed much of the time

Move the protection code to a Proxy class

Decorated Collections

[Todo: Write up the story of the move from the Java 1.0 synchronized Vector and Hashtable classes to the Java 1.1 unsynchronized collections classes that use Collections.synchronizedMap() to obtain a synchronization decorator].

Vector

```
public synchronized void addElement(Object obj) {  
    modCount++;  
    ensureCapacityHelper(elementCount + 1);  
    elementData[elementCount++] = obj;  
}
```

```
static class SynchronizedCollection implements Collection, Serializable {  
    Collection c; // Backing Collection  
    Object mutex; // Object on which to synchronize
```

```
SynchronizedCollection(Collection c) {  
    this.c = c; mutex = this;  
}
```

```
public boolean add(Object o) {  
    synchronized(mutex) {return c.add(o);}  
}  
public boolean remove(Object o) {  
    synchronized(mutex) {return c.remove(o);}  
}
```

Collections...

```
public static List synchronizedList(List list) {  
    return new SynchronizedList(list);  
}
```

```
static class SynchronizedList extends SynchronizedCollection  
    implements List {
```

```
    private List list;
```

```
SynchronizedList(List list) {  
    super(list);  
    this.list = list;  
}
```

```
SynchronizedList(List list, Object mutex) {  
    super(list, mutex);  
    this.list = list;
```

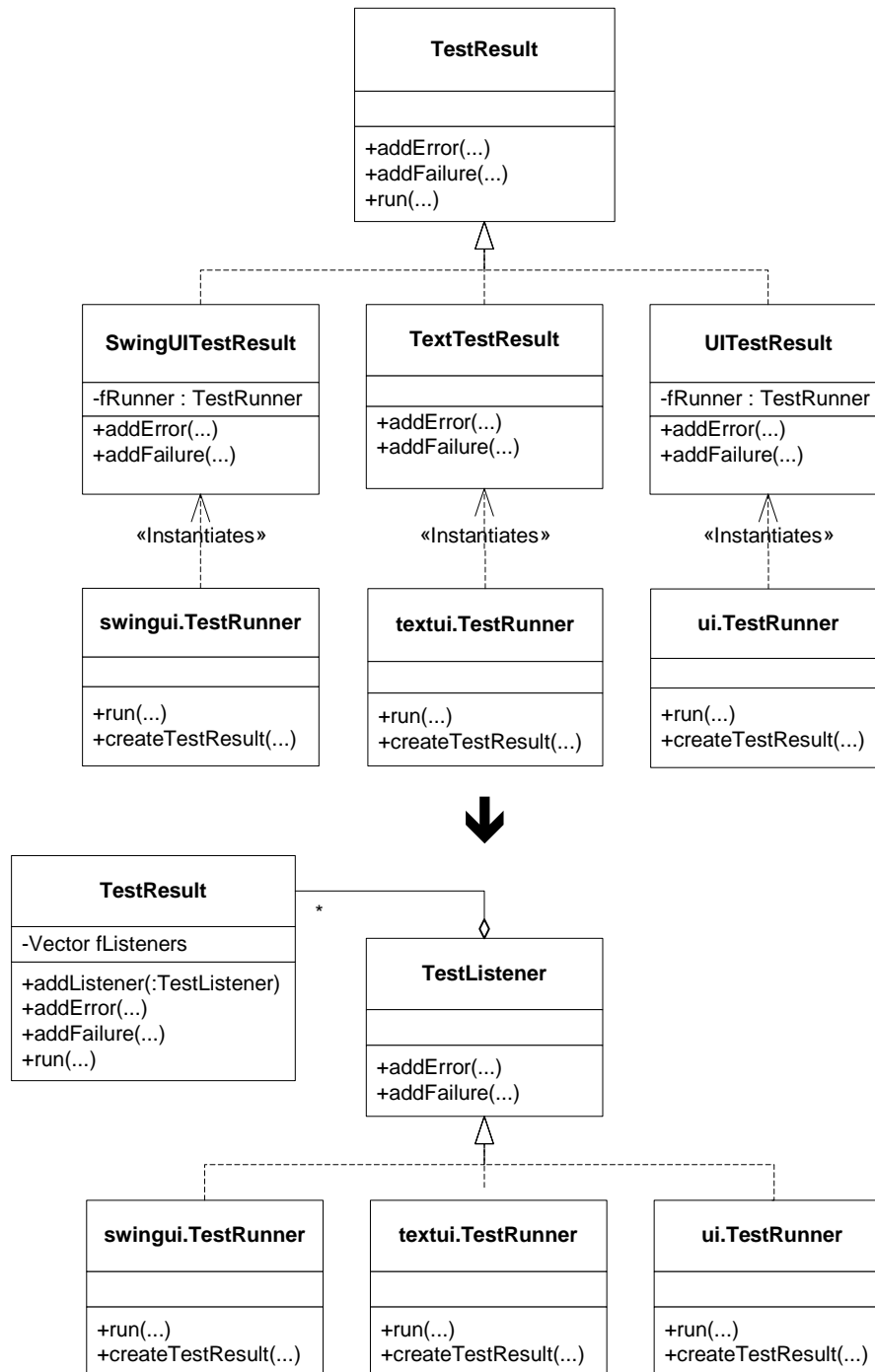
```
    }

    public void add(int index, Object element) {
        synchronized(mutex) {list.add(index, element);}
    }
    public Object remove(int index) {
        synchronized(mutex) {return list.remove(index);}
    }
}
```

Replace Hard-Coded Notifications with Observer

Your class or numerous subclasses perform custom object notifications at designated times

Replace your custom notification code with the Observer pattern



Motivation

The Observer pattern is popular. Many programmers know it well and use it often. But the trick is to learn when you actually *need* to use Observer and when you don't.

Consider under what circumstances the authors of Design Patterns suggest using Observer (see *Design Patterns*, page 294):

- When an abstraction has two aspects, one dependent on the other. Encapsulating these aspects in separate objects lets you vary and reuse them independently.
- When a change to one object requires changing others, and you don't know how many objects need to be changed.
- When an object should be able to notify other objects without making assumptions about who these objects are. In other words, you don't want these objects tightly coupled.

Now, what happens when you *do* know the object you want to update and it isn't necessarily to have *loose coupling* with that object? For example, class A needs to update objects of type B, based on some event. Since this is a notification responsibility, you may want to implement a solution using the Observer pattern (or Java's Listeners -- essentially the same idea). But do you really need to go that far? Could Observer be too heavyweight a solution given this example? What if you simply wrote code in class A that would notify B objects at appropriate times?

Certainly that could work just fine, until objects of type C also need to be notified about A's events. You could then experiment with your code. See if adding more hard-coded notification logic in class A overcomplicates the class. If it doesn't, you've solved your immediate need without writing much new code.

Eventually, class A's notification responsibilities may grow. As the responsibilities grow, you must observe your own interactions with the code. Ask yourself questions like:

- Am I finding duplicate notification code?
- Am I creating relatively dumb subclasses just to satisfy new notification needs?
- Is my notification logic becoming too complex?
- Is it awkward to pass in object references to class A just for the purpose of notification?

The answers to these questions may lead you to refactor to Observer. Doing so should lead to simpler, smaller and easier-to-read code. Just remember that once you do decide to refactor to Observer, try to do so in the simplest way possible. For example, if your observers will never need to stop getting notifications, do *not* write the `removeObserver()` code on your Subject class - it would only be wasted code that no one uses.

Communication	Duplication	Simplicity
Hard-coded object notifications enable runtime collaborations, but the code doesn't communicate this very well: objects get passed into constructors, and notifications happen in random methods. Compare this to a class that implements the Observer pattern – both <i>who</i> can observe its events and <i>when</i> they get notified is clearly communicated in the class declaration.	If you are compelled to write special code for every class that must be notified at runtime, you can easily produce more code than you need, perhaps resulting in parallel or near-parallel class hierarchies. For only a few notifications, this is no big deal. But as you add more and more special notification code, duplication and code bloat take over.	A few runtime object notifications can be easily handled with simple custom code. But when the number of notifications increases, lots of special code will be written or more and more subclasses will be produced to obtain the necessary behavior. At that point, your code can be simplified by using the Observer pattern.

Mechanics

1. Identify a *Subject*: a class that accepts an object reference and contains hard-coded notification instructions that couple it directly to the object reference type.
2. Define an *Observer*: an interface that consists of the set of public methods called by the *Subject* on the referenced object.
3. Add to the *Subject* an Observers list and a way for clients to add to that list via a public `addObserver(Observer observer)` method. Add a corresponding `removeObserver(Observer observer)` method *only* if observers need to be removed at runtime.
4. Replace code in the *Subject* that accepted an object reference and directly notified that reference with code that iterates over *Subject*'s Observer list, updating each *Observer* instance.
5. For any class that needs to get notified by *Subject*, make it implement the *Observer* interface.
6. Replace code that passed in an object reference to the *Subject* with code that registers that object reference as an *Observer* of the *Subject*. You'll use *Subject*'s `addObserver(Observer observer)` method for this purpose.

✓ Compile and test.

Example

The code sketch above is from Kent Beck and Erich Gamma's JUnit Testing Framework. Prior to JUnit 3.x, the authors defined specific `TestResult` subclasses (like `UITestResult`, `SwingTestResult` and `TextTestResult`) that were responsible for gathering up test information and reporting it to `TestRunners`. Each `TestResult` subclass was coupled to a specific `TestRunner`, such as an `AWT TestRunner`, `Swing TestRunner` or `Text-based TestRunner`. At runtime, after creating a `TestResult` subclass, a `TestRunner` would pass itself in as a reference to that `TestResult`, and then wait to be notified by the `TestResult`. Each `TestResult` subclass was hard-coded this way to talk with a specific `TestRunner`, and that is where our refactoring begins.

In JUnit 3.1, Kent and Erich refactored the `TestResult`/`TestRunner` code to use the Observer pattern. This enabled them to eliminate *all* of the special `TestResult` subclasses (`UITestResult`, `SwingTestResult` and `TextTestResult`) and simplify each of the concrete `TestRunners`.

Our example will look at this real-world refactoring of the JUnit framework. I've deliberately simplified some of the JUnit code in order to concentrate on the refactoring, not the inner workings of JUnit. However, if you want to study the JUnit code (which I highly recommend), you can download it at <http://www.junit.org>.

1. Our first task is to find a Subject. In this case, the `UITestResult` class will be our Subject, but later our Subject will become the `TestResult` class. What is the reason for this? Well, as a subclass of `TestResult`, `UITestResult` doesn't add much new behavior: it exists only because it has the ability to talk directly to an AWT `TestRunner` class. Our refactoring will seek to eliminate `UITestResult` and move its behavior up to the `TestResult` class.

Let's look at the code for all three classes, minus some details you don't need to worry about. I highlight in bold the coupling between `UITestResult` and its AWT `TestRunner`:

```
package junit.framework;
public class TestResult extends Object {
    protected Vector fFailures;

    public TestResult() {
        fFailures= new Vector(10);
    }
    public synchronized void addFailure(Test test, Throwable t) {
        fFailures.addElement(new TestFailure(test, t));
    }
    public synchronized Enumeration failures() {
        return fFailures.elements();
    }
    protected void run(TestCase test) {
        startTest(test);
        try {
            test.runBare();
        }
        catch (AssertionFailedError e) {
            addFailure(test, e);
        }
        endTest(test);
    }
}

package junit.ui;
class UITestResult extends TestResult {
    private TestRunner fRunner;
    UITestResult(TestRunner runner) {
        fRunner= runner;
    }
    public synchronized void addFailure(Test test, Throwable t) {
        super.addFailure(test, t);
        fRunner.addFailure(this, test, t);
    }
    ...
}

package junit.ui;
public class TestRunner extends Frame {
    private TestResult fTestResult;
    ...
    protected TestResult createTestResult(TestRunner runner) {
        return new UITestResult(TestRunner.this);
    }
    synchronized public void runSuite() {
        ...
        fTestResult = createTestResult(TestRunner.this);
        testSuite.run(fTestResult);
    }
}
```

```

    }
    public void addFailure(TestResult result, Test test, Throwable t) {
        fNumberOfFailures.setText(Integer.toString(result.testFailures()));
        appendFailure("Failure", test, t);
    }
}

```

2. Our next task is to define an Observer interface. Kent and Erich call this a `TestListener`:

```

package junit.framework;
public interface TestListener {
    public void addError(Test test, Throwable t);
    public void addFailure(Test test, Throwable t);
    public void endTest(Test test);
    public void startTest(Test test);
}

```

3. We must now add a list of Observers to our Subject and provide clients (that implement the Observer interface) a way to add themselves to this list. We do this work on the `TestResult` class rather than the `UITestResult` class, which we hope to eliminate:

```

public class TestResult extends Object {
    protected Vector fFailures;
    protected Vector fListeners;
    public TestResult() {
        fFailures= new Vector();
        fListeners= new Vector();
    }
    public synchronized void addListener(TestListener listener) {
        fListeners.addElement(listener);
    }
}

```

4. Now we need to make our Subject update its Observers when an event happens. This involves refactoring `TestResult` methods like `addFailure()`, `addError()` and so on. For simplicity, we will examine only how `addFailure()` is refactored. Here's what the original method looked like on `UITestResult`:

```

class UITestResult. . .
    public synchronized void addFailure(Test test, Throwable t) {
        super.addFailure(test, t);
        fRunner.addFailure(this, test, t);
    }
}

```

Rather than refactor `UITestResult`'s `addFailure()` method, we focus on the same method in `TestResult`, the superclass. `TestResult`'s `addFailure` method will continue to do what it used to do, but it will now iterate through its registered Observers, calling each one's `addFailure()` method. In this context, since Observers are usually `TestRunners`, this code will inform each registered `TestRunner` that a failure has been added. When that happens, the `TestRunners` have a chance to do things like update a GUI to reflect just how many test failures have occurred. Here's what `TestResult`'s refactored `addFailure()` method looks like:

```

class TestResult. . .
    public synchronized void addFailure(Test test, AssertionError t) {
        fFailures.addElement(new TestFailure(test, t));
        for (Enumeration e= cloneListeners().elements(); e.hasMoreElements(); ) {
            ((TestListener)e.nextElement()).addFailure(test, t);
        }
    }
}

```

5. Now, in order for the AWT `TestRunner` to register itself as an Observer of a `TestResult`, we must make the `ui.TestRunner` class implement the `TestListener` interface:

```
package junit.ui;
public class TestRunner extends Object implements TestListener . . .
```

6. The final step is to register the Observer with the Subject of choice. In this case, we'll look at the code that registers the `ui.TestRunner` with a `TestResult` instance:

```
package junit.ui;
public class TestRunner extends Object implements TestListener {
    private Vector fFailedTests;
    private TestResult fTestResult;

    protected TestResult createTestResult() {
        return new TestResult();
    }
    synchronized public void runSuite() {
        . . .
        fTestResult = createTestResult();
        fTestResult.addListener(TestRunner.this);
        . . .
    }
}
```

Finally, we can now compile and test that our refactored `ui.TestRunner` and `TestResult` work together the way we expect. In the real world, Kent and Erich refactored all of the `TestResult` subclasses and `TestRunners` to use the Observer pattern.

Move Accumulation to Collecting Parameter

You have a single bulky method
that accumulates information to a variable

*Accumulate results to a Collecting Parameter
that gets passed to extracted methods.*

```
class TagNode. . .
public String toString() {
    String result = new String();
    result += "<" + tagName + " " + attributes + ">";
    Iterator it = children.iterator();
    while (it.hasNext()) {
        TagNode node = (TagNode)it.next();
        result += node.toString();
    }
    if (!tagValue.equals(""))
        result += tagValue;
    result += "</" + tagName + ">";
    return result;
}
```



```
class TagNode. . .
public String toString() {
    return toStringHelper(new StringBuffer(""));
}
private String toStringHelper(StringBuffer result) {
    writeOpenTagTo(result);
    writeChildrenTo(result);
    writeEndTagTo(result);
    return result.toString();
}
```

Motivation

Kent Beck defined the Collecting Parameter pattern in his classic book, *Smalltalk Best Practice Patterns*. A Collecting Parameter is an object that you pass to methods in order to collect information from those methods. A good reason to use this pattern is when you want to decompose a bulky method into smaller methods (using *Extract Method* [F]), and you need to accumulate information from each of the extracted methods. Instead of making each of the extracted methods return a result, which you later combine into a final result, you can incrementally accumulate your result by passing a collecting parameter to each of the extract methods, which in turn, write their results to the collecting parameter.

Collecting Parameter works nicely with the Composite pattern, since you can use a Collecting Parameter to recursively accumulate information from a Composite structure. Kent Beck and Erich Gamma combined these two patterns in their JUnit testing framework to enable a single `TestResult` object to gather test result information from every test in a hierarchical structure of test case objects.

I recently combined Collecting Parameter with Composite when I refactored a class's `toString()` method (see the code sketch above). My initial goal was to replace a lot of slow `String` concatenation code with faster `StringBuffer` code, but when I realized that a simple

replacement would generate lots of `StringBuffer` instances (because the code is recursive), I retreated from this approach. Then my programming partner at the time, Don Roberts, seized the keyboard, saying “I’ve got it, I’ve got it” and then quickly refactored the code to use a single `StringBuffer` as a Collecting Parameter. The resulting code (partially shown in the code sketch) had a far simpler design, communicated better with the reader and, thanks to the `StringBuffer`, was more efficient.

Communication	Duplication	Simplicity
Bulky methods don’t communicate well. Communicate what you are accumulating by placing each step into intention-revealing methods that write results to a parameter.	You don’t often reduce duplicate code using this refactoring. The only exception would be if you have different types of Collecting Parameters that can be passed into the same methods.	Extract Method is at the heart of this refactoring. You use it to reduce a bulky method into a simpler method that delegates to intention-revealing methods.

Mechanics

1. Identify a chunk of code that accumulates information into a variable (we’ll call that variable “result”). Result will become your Collecting Parameter. If result’s type won’t let you iteratively gather data across methods, change result’s type. For example, Java’s `String` won’t let us accumulate results across methods, so we use a `StringBuffer`.
2. Find an information accumulation step and extract it into a private method (using *Extract Method* [F]). Make the method’s return type be void and pass it result. Inside the method, write information to result.
3. Repeat steps 2 for every accumulation step, until the original code has been replaced with calls to extracted methods that accept and write to result.

✓ Compile and test.

Example

In this example, we will see how to refactor Composite-based code to use a Collecting Parameter. We’ll start with a composite that can model an XML tree (see *Replace Implicit Tree with Composite* (60) for a complete example of this XML composite code).

The composite is modeled with a single class, called `TagNode`, which has a `toString()` method. The `toString()` method recursively walks the nodes in the XML tree, and produces a final `String` representation of what it finds. It does a fair amount of work in 11 lines of code. We will refactor `toString()` to make it simpler and easier to understand.

1. The following `toString()` method recursively accumulates information from every tag in a composite structure and stores results in a variable called “result”:

```
class TagNode. . .
public String toString() {
    String result = new String();
    result += "<" + tagName + " " + attributes + ">";
    Iterator it = children.iterator();
    while (it.hasNext()) {
        TagNode node = (TagNode)it.next();
        result += node.toString();
    }
    if (!tagValue.equals(""))
        result += tagValue;
}
```

```
    result += "</" + tagName + ">";
    return result;
}
```

I change result's type to be a StringBuffer in order to support this refactoring:

```
StringBuffer result = new StringBuffer("");
```

2. I identify the first information accumulation step: code that concatenates an xml open tag along with any attributes to the result variable. I Extract Method on this code as follows:

```
result += "<" + tagName + " " + attributes + ">";
```

is extracted to:

```
private void writeOpenTagTo(StringBuffer result) {
    result.append("<");
    result.append(name);
    result.append(attributes.toString());
    result.append(">");
}
```

The original code now looks like this:

```
StringBuffer result = new StringBuffer("");
writeOpenTagTo(result);
...
```

3. Next, I want to continue to extract methods from toString(). I focus on the code that adds child XML nodes to the result. This code contains a recursive step (which I highlight below in bold):

```
class TagNode. . .
    public String toString(). . .
        Iterator it = children.iterator();
        while (it.hasNext()) {
            TagNode node = (TagNode)it.next();
            result += node.toString();
        }
        if (!tagValue.equals(""))
            result += tagValue;
        . . .
}
```

Since this code makes a recursive call, it isn't so easy to extract into a method. The following code will show you why:

```
private void writeChildrenTo(StringBuffer result) {
    Iterator it = children.iterator();
    while (it.hasNext()) {
        TagNode node = (TagNode)it.next();
        node.toString(result); // can't do this because toString() doesn't take arguments.
    }
    . . .
}
```

Since toString() doesn't take a StringBuffer as an argument I can't simply extract the method. I have to find another solution and I decide to solve the problem using a helper method. This method will do the work that toString() used to do, but it will take a StringBuffer as a Collecting Parameter:


```
public String toString() {
    return toStringHelper(new StringBuffer(""));
}

private String toStringHelper(StringBuffer result) {
    writeOpenTagTo(result);
    . . .
    return result.toString();
}
```

With the new `toStringHelper()` method in place, I can go back to my original task: extracting the next accumulation step:

```
private String toStringHelper(StringBuffer result) {
    writeOpenTagTo(result);
    writeChildrenTo(result);
    ...
    return result.toString();
}

private void writeChildrenTo(StringBuffer result) {
    Iterator it = children.iterator();
    while (it.hasNext()) {
        TagNode node = (TagNode)it.next();
        node.toStringHelper(result); // now recursive call will work
    }
    if (!value.equals(""))
        result.append(value);
}
```

As I stare at the `writeChildrenTo()` method, I realize that it is handling two steps: adding children recursively and adding a value to a tag, when one exists. To make these two separate steps stand out, I extract the code for handling a value into its own method:

```
private void writeValueTo(StringBuffer result) {
    if (!value.equals(""))
        result.append(value);
}
```

To finish the refactoring, I extract one more method that writes an XML close tag. Here's what the final code looks like:

```
public class TagNode . . .
{
    public String toString() {
        return toStringHelper(new StringBuffer(""));
    }
    private String toStringHelper(StringBuffer result) {
        writeOpenTagTo(result);
        writeChildrenTo(result);
        writeValueTo(result);
        writeEndTagTo(result);
        return result.toString();
    }
    private void writeOpenTagTo(StringBuffer result) {
        result.append("<");
        result.append(name);
        result.append(attributes.toString());
        result.append(">");
    }
    private void writeChildrenTo(StringBuffer result) {
        Iterator it = children.iterator();
        while (it.hasNext()) {
            TagNode node = (TagNode)it.next();
            node.toStringHelper(result);
        }
    }
}
```

```
    }
    private void writeValueTo(StringBuffer result) {
        if (!value.equals(""))
            result.append(value);
    }
    private void writeEndTagTo(StringBuffer result) {
        result.append("</");
        result.append(name);
        result.append(">");
    }
}
```

Or so I thought that was the final code. An astute reader of the above code pointed out that when the `writeChildrenTo()` method recursively calls `toStringHelper()`, it is returned a `String`, which it promptly ignores. In other words, the only time that the return result of `toStringHelper()` is used is when it is called from the `toString()` method. This means that the code can be made more efficient as follows:

```
public String toString() {
    StringBuffer result = new StringBuffer("");
    toStringHelper(result);
    return result.toString();
}
public void toStringHelper(StringBuffer result) {
    writeOpenTagTo(result);
    writeChildrenTo(result);
    writeValueTo(result);
    writeEndTagTo(result);
}
```

I compile, run my tests and everything is good.

JUnit's Collecting Parameter

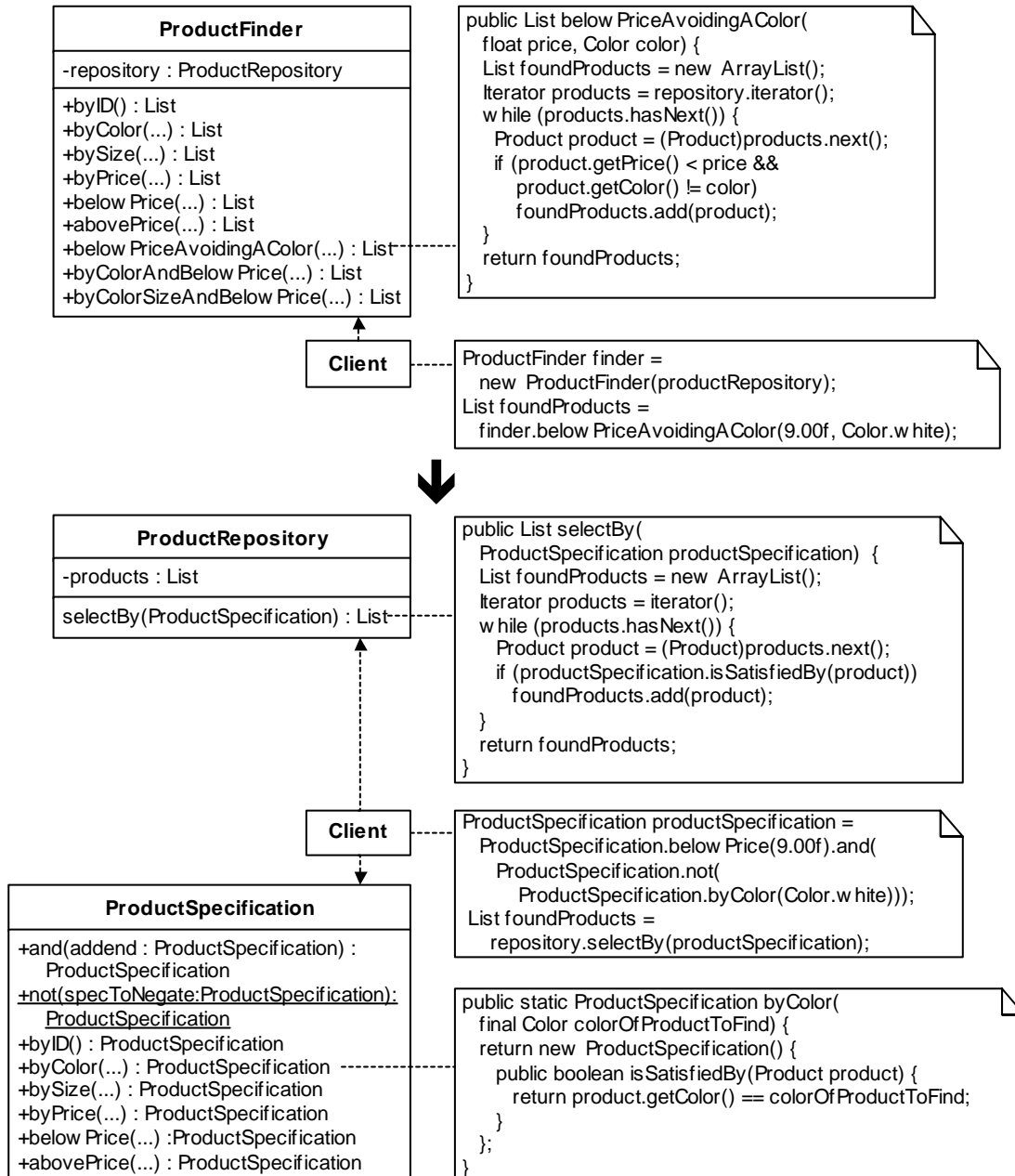
To get a better understanding of the Collecting Parameter pattern, let's have a look at another example, which comes from the unit testing framework, JUnit. In JUnit, every test is an object. Test objects get put into suites, which may be put into more suites, which results in a composite of tests. To report on how each test performs (did it pass, fail or generate errors?), some object needs to accumulate and report results as each test in the Composite is executed. `TestResult` is that object and it serves the role of Collecting Parameter.

[add uml and more description]

Replace Conditional Searches with Specification

You use conditional logic to search for objects in a repository

Perform the search using a Specification



Motivation

The Specification pattern was originally written by Eric Evans and Martin Fowler in a PLoP '97 paper [Evans&Fowler] and was later rewritten by Eric in his excellent book, *Domain-Driven Design* [Evans]. In essence, a Specification is an object that contains a rule about another object, which I'll call an entity. When you want to select certain entities from an entity pool, you construct a Specification object (or a composite Specification) to find your target entities.

A common motivation for refactoring to Specification is to tame a combinatorial explosion problem. For example, say you've written a `ProductFinder` that knows how to find `Products` in a variety of ways:

```
ProductFinder...
    public List byID(String productID)...
    public List byColor(Color colorOfProductToFind)...
    public List byPrice(float priceLimit)...
    public List bySize(int sizeToFind)...
    public List belowPrice(float price)...
```

Over time, the combination of ways to find products grows:

```
ProductFinder...
    public List byColorAndBelowPrice(Color color, float price)...
    public List byColorSizeAndBelowPrice(Color color, int size, float price)...
    public List byColorAndAbovePrice(Color color, float price)...
```

Eventually you have too many finder methods to account for all of the ways you search for products. The Specification pattern was made for such a problem. It suggests that you turn the *elements* of your product selection criteria (i.e. search *criterion*) into objects so that clients may combine these objects however they like to perform specialized product searches. A search itself can then be implemented by a single method that takes one Specification object (which is often a Composite of Specifications), iterates over a collection of products (or whatever objects your care to find) and returns a list of objects that satisfy the Specification.

Now you may be wondering why not just use plain old SQL? There are a few reasons. First, it's common to have a pool of entities in memory, where SQL won't help you much. Second, as Eric explains in his book, Specifications are meant to be independent of things like databases, for they represent pure domain logic. That said, if you need an SQL representation of a Specification you could consider writing code to transform a Specification into SQL. Some colleagues of mine did just that by writing a visitor to visit a Specification and generate the appropriate SQL. We'll explore their refactoring in [ToDo: write it and link it].

Communication	Duplication	Simplicity
If entity selection logic is encapsulated by intention-revealing methods (for example, on a Finder class) the code may communicate sufficiently about what kinds of entity selections are possible. On the other hand, if the entity selection logic is scattered around a system, consolidating that logic via Specification classes will help the code better communicate what entity criterion can be used to perform entity selections.	When a system needs to select entities from an entity pool in a variety of ways, it's possible to have a great deal of duplicate code. Specification helps you tame that duplication by giving clients the ability to easily compose their own Specification objects, which may then be used by a single entity selection method.	If you're only using the Specification pattern for entity selection, but you don't have a combinatorial explosion problem (i.e. a variety of ways to select entities from a pool), using the Specification pattern could be less simple and straightforward than using primitive entity selection logic.

Mechanics

1. Create an abstract *specification* class, named after an *entity type* you're interested in selecting from an *entity pool* and the word "Specification" (e.g. ProductSpecification, InvoiceSpecification, etc.)
2. Declare an `isSatisfiedBy()` method on the *specification* class that returns a boolean and takes one argument: the *entity type* you're interest in selecting from the *entity pool* (e.g. `public boolean isSatisfiedBy(ProductSpecification productSpecification)`).
3. Create a `selectBy(...)` method that takes the *specification* as an argument and returns a collection of *entities* (e.g., `public List selectBy(ProductSpecification spec)`). Place the `selectBy(...)` method on whatever class has access to the *entity pool* and would most naturally provide *entity selection logic* for the *entity type* in question. (A common choice would be a class that acts as an *entity repository*).
4. Find *entity selection logic* (usually iteration code) that relies on one *criterion* (e.g. `product.getPrice() == targetAmount`) and create a *concrete specification* class named after this *criterion* (e.g. ProductPriceSpecification). Implement the *concrete specification's* `isSatisfiedBy(...)` method using the *criterion* logic, altering that logic to evaluate and return a boolean result using the entity argument passed to `isSatisfiedBy(...)`. Since most *criterion* logic usually evaluates data supplied by a client, you'll usually need to create a *concrete specification* constructor that accepts whatever information is needed to perform the evaluation in `isSatisfiedBy(...)`. However, a nice alternative in languages that allow you to create classes on the fly (e.g. anonymous inner classes in Java) is to apply the refactoring, *Encapsulate Classes with Creation Methods* (27) - (see the example for details).

If you have several pieces of *entity selection logic* that rely on *criteria* (e.g. `product.getPrice() < limitPrice && product.getColor() != colorToAvoid`), it often makes sense to create *concrete specifications* for each *criterion* so clients may combine *criterion* however they need to. Handle simple *criterion* first and wait until step 7 to refactor *criteria*.

5. Replace the *criterion* and associated *entity selection logic* with a call to the `selectBy(...)` method, passing it a newly constructed *concrete specification*. If the resulting code is a one-line method, consider applying *Inline Method* (117) [F].

✓ Compile and test.

6. Repeat steps 4 and 5 for additional *entity selection logic* and its associated *criterion*.
7. If you have *entity selection logic* that relies on *criteria*, apply *Replace Implicit Tree with Composite* (60). This involves making *concrete specifications* for the *criteria's criterion* so that clients may combine the *criterion* into a Composite specification using special *composite specification* classes for operations like `and()`, `or()` and `not()`. After you create the requisite *composite specification(s)* for your *criteria*, apply step 5.

✓ Compile and test.

8. Repeat step 7 for all remaining *criteria*.
9. Apply *Encapsulate Classes with Creation Methods* (27) by implementing Creation Methods on the *specification* class for each *concrete specification*.

You can implement this step in Java by turning your concrete specification classes into inner or anonymous inner classes. Using anonymous inner classes removes the necessity to define a constructor for each concrete specification, resulting in significantly less code.

✓ Compile and test.

Example

The code sketch and motivation sections already gave you an introduction to this example, which is inspired from an inventory management system. That system's Finder classes (AccountFinder, InvoiceFinder, ProductFinder and so forth), eventually came to suffer from a combinatorial explosion smell, which necessitated the refactoring to Specification. It's worth noting that this does not reveal a problem with Finder classes: the point is that a time may come when a refactoring to Specification is justified.

Let's begin by studying the tests and code for a ProductFinder, which is in need of this refactoring. We'll start with the test code. Before any test can run, we need a ProductRepository object that's filled with various toy Products and a ProductFinder object that knows about the ProductRepository:

```
public class ProductFinderTests extends TestCase...
    private ProductFinder finder;

    private Product fireTruck =
        new Product("f1234", "Fire Truck",
            Color.red, 8.95f, ProductSize.MEDIUM);

    private Product barbieClassic =
        new Product("b7654", "Barbie Classic",
            Color.yellow, 15.95f, ProductSize.SMALL);

    private Product frisbee =
        new Product("f4321", "Frisbee",
            Color.pink, 9.99f, ProductSize.LARGE);

    private Product baseball =
        new Product("b2343", "Baseball",
            Color.white, 8.95f, ProductSize.NOT_APPLICABLE);

    private Product toyConvertible =
        new Product("p1112", "Toy Porshe Convertible",
            Color.red, 230.00f, ProductSize.NOT_APPLICABLE);

    protected void setUp() {
        finder = new ProductFinder(createProductRepository());
    }

    private ProductRepository createProductRepository() {
        ProductRepository repository = new ProductRepository();
        repository.add(fireTruck);
        repository.add(barbieClassic);
        repository.add(frisbee);
        repository.add(baseball);
        repository.add(toyConvertible);
        return repository;
    }
```

```
}
```

We'll begin by looking at a few simple tests and the implementation code that satisfies them. `testFindByColor` checks if the `ProductFinder.byColor` method correctly finds red toys, while `testFindByPrice` checks if `ProductFinder.byPrice` correctly finds toys at a given price:

```
public class ProductFinderTests extends TestCase...
{
    public void testFindByColor() {
        List foundProducts = finder.byColor(Color.red);
        assertEquals("found 2 red products", 2, foundProducts.size());
        assertTrue("found fireTruck", foundProducts.contains(fireTruck));
        assertTrue(
            "found Toy Porshe Convertible",
            foundProducts.contains(toyConvertible));
    }

    public void testFindByPrice() {
        List foundProducts = finder.byPrice(8.95f);
        assertEquals("found products that cost $8.95", 2, foundProducts.size());
        for (Iterator i = foundProducts.iterator(); i.hasNext(); ) {
            Product p = (Product) i.next();
            assertTrue(p.getPrice() == 8.95f);
        }
    }
}
```

Here's the implementation code that satisfies these tests:

```
public class ProductFinder...
{
    private ProductRepository repository;

    public ProductFinder(ProductRepository repository) {
        this.repository = repository;
    }

    public List byColor(Color colorOfProductToFind) {
        List foundProducts = new ArrayList();
        Iterator products = repository.iterator();
        while (products.hasNext()) {
            Product product = (Product) products.next();
            if (product.getColor().equals(colorOfProductToFind))
                foundProducts.add(product);
        }
        return foundProducts;
    }

    public List byPrice(float priceLimit) {
        List foundProducts = new ArrayList();
        Iterator products = repository.iterator();
        while (products.hasNext()) {
            Product product = (Product) products.next();
            if (product.getPrice() == priceLimit)
                foundProducts.add(product);
        }
        return foundProducts;
    }
}
```

There's plenty of duplicate code in the two methods above. We'll be getting rid of that duplication during this refactoring. Meanwhile, let's look at some more tests and code that are involved in the combinatorial explosion problem. Below, one test is concerned with finding Products by color, size and below a certain price while the other test is concerned with finding Products by color and above a certain price:

```
public class ProductFinderTests extends TestCase...
    public void testFindByColorSizeAndBelowPrice() {
        List foundProducts =
            finder.byColorSizeAndBelowPrice(
                Color.red,
                ProductSize.SMALL,
                10.00f);
        assertEquals(
            "found no small red products below $10.00",
            0,
            foundProducts.size());

        foundProducts =
            finder.byColorSizeAndBelowPrice(
                Color.red,
                ProductSize.MEDIUM,
                10.00f);
        assertEquals(
            "found firetruck when looking for cheap medium red toys",
            fireTruck,
            foundProducts.get(0));
    }

    public void testFindBelowPriceAvoidingAColor() {
        List foundProducts =
            finder.belowPriceAvoidingAColor(9.00f, Color.white);
        assertEquals(
            "found 1 non-white product < $9.00",
            1,
            foundProducts.size());
        assertTrue("found fireTruck", foundProducts.contains(fireTruck));

        foundProducts = finder.belowPriceAvoidingAColor(9.00f, Color.red);
        assertEquals(
            "found 1 non-red product < $9.00",
            1,
            foundProducts.size());
        assertTrue("found baseball", foundProducts.contains(baseball));
    }
}
```

Here's what the implementation code looks like for the above tests:

```
public class ProductFinder...
    public List byColorSizeAndBelowPrice(Color color, int size, float price) {
        List foundProducts = new ArrayList();
        Iterator products = repository.iterator();
        while (products.hasNext()) {
            Product product = (Product) products.next();
            if (product.getColor() == color
                && product.getSize() == size
                && product.getPrice() < price)
                foundProducts.add(product);
        }
        return foundProducts;
    }

    public List belowPriceAvoidingAColor(float price, Color color) {
        List foundProducts = new ArrayList();
        Iterator products = repository.iterator();
        while (products.hasNext()) {
            Product product = (Product) products.next();
            if (product.getPrice() < price && product.getColor() != color)
                foundProducts.add(product);
        }
        return foundProducts;
    }
}
```

Again, you see plenty of duplicate code since each of the specific finder methods iterates over the same repository and selects just those `Products` that match the specified criteria. We're now ready to begin the refactoring.

1. The first step is to create an abstract `ProductSpecification` class:

```
public abstract class ProductSpecification {  
}
```

2. I'll now add an `isSatisfiedBy(...)` method to `ProductSpecification`:

```
public abstract class ProductSpecification {  
    public abstract boolean isSatisfiedBy(Product product);  
}
```

3. Next, I create a `selectBy(...)` method that will accept a `ProductSpecification` and return a `List` containing `Products` that satisfy the supplied specification. I could put this method on the `ProductFinder` class, but upon reflection, I see that it doesn't really belong there, for when I finish this refactoring, there will be no need for a `ProductFinder`. So I decide to put the method on the `ProductRepository` class:

```
public class ProductRepository...  
    public Iterator iterator() {  
        return products.iterator();  
    }  
  
    public List selectBy(ProductSpecification productSpecification) {  
        List foundProducts = new ArrayList();  
        Iterator products = iterator();  
        while (products.hasNext()) {  
            Product product = (Product)products.next();  
            if (productSpecification.isSatisfiedBy(product))  
                foundProducts.add(product);  
        }  
        return foundProducts;  
    }  
}
```

Since `selectBy(...)` iterates over all products in the product repository, it will eventually enable me to remove lots of identical iteration logic in `ProductFinder`'s numerous find methods.

4. I can now create my first concrete `ProductSpecification`: I choose a `Specification` for products of a given color:

```
public class ProductColorSpecification extends ProductSpecification {  
    private Color colorOfProductToFind;  
  
    public ProductColorSpecification(Color colorOfProductToFind) {  
        this.colorOfProductToFind = colorOfProductToFind;  
    }  
  
    public boolean isSatisfiedBy(Product product) {  
        return product.getColor().equals(colorOfProductToFind);  
    }  
}
```

5. The body of the `ProductFinder.byColor(...)` method can now be replaced by a call to the new `ProductFinder.selectBy(...)` method, passing it a newly constructed `ProductColorSpecification`:

```
public class ProductFinder...  
    public List byColor(Color colorOfProductToFind) {  
        return selectBy(new ProductColorSpecification(colorOfProductToFind));  
        List foundProducts = new ArrayList();  
        Iterator products = repository.iterator();  
    }  
}
```

```
while (products.hasNext()) {  
    Product product = (Product) products.next();  
    if (product.getColor().equals(colorOfProductToFind))  
        foundProducts.add(product);  
}  
return foundProducts;  
}
```

Since callers on `byColor(...)` could just as easily perform its one-line method, I'll apply *Inline Method (117)* [F] by deleting `byColor()` and changing its callers to directly execute its code:

```
public class ProductFinder...  
    public List byColor(Color colorOfProductToFind) {  
        return selectBy(new ProductColorSpecification(colorOfProductToFind));  
    }  
  
public class ProductFinderTests extends TestCase...  
    public void testFindByColor()...  
        List foundProducts = finder.byColor(Color.red);  
        List foundProducts = repository.selectBy(  
            new ProductColorSpecification(Color.red));
```

I compile and test to make sure that everything's working.

6. I now repeat steps 4 & 5, creating *concrete specifications* for *entity selection logic* that relies on *critereon*. For the entire code example (not shown), this leads me to create the following *concrete specifications*:

Product Critereon	Concrete Product Specification
id	ProductIdSpecification
size	ProductSizeSpecification
price	ProductPriceSpecification
below price	ProductBelowPriceSpecification
above price	ProductAbovePriceSpecification

At this point I do not create concrete specifications for combinations of critereon (i.e. criteria), such as:

Product Criteria:

color & size & below price
color & below price
color & above price
below price and not some specified color

7. I'm now ready to deal with product *criteria*. In the code I'm refactoring, there are several places where product *criteria* is used to select a given set of products. In one case a client seeks *criteria* products below a given price and not of a particular color:

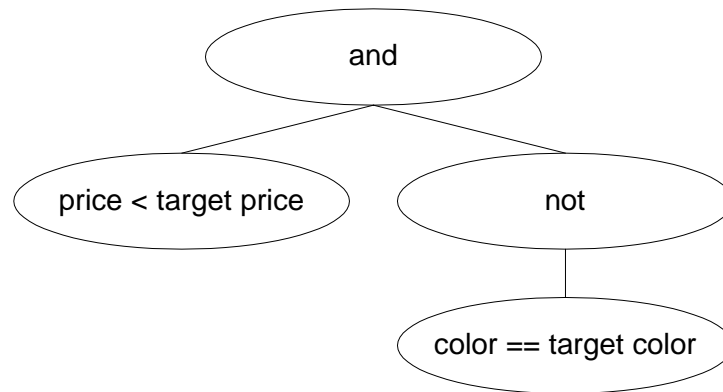
```
public class ProductFinder...  
    public List belowPriceAvoidingAColor(float price, Color color) {  
        List foundProducts = new ArrayList();  
        Iterator products = repository.iterator();  
        while (products.hasNext()) {  
            Product product = (Product) products.next();  
            if (product.getPrice() < price && product.getColor() != color)  
                foundProducts.add(product);  
        }  
    }  
}
```

```

    }
    return foundProducts;
}

```

There are two pieces of criterion in this product criteria: `product.getPrice() < price` and `product.getColor() != color`). `ProductBelowPriceSpecification` is a `ProductSpecification` that helps find products below a given price while `ProductColorSpecification` helps find products of a given color. What we need now is a composite specification that will let us define one specification to combine two concrete specifications, negating the color one. This is a job for the refactoring, *Replace Implicit Tree with Composite* (60). What is the implicit tree in this case? The price and color product *criteria* logic form the following tree structure:



To implement this refactoring, two composite specification classes will need to be created: `AndProductSpecification` and `NotProductSpecification`:

```

public class AndProductSpecification extends ProductSpecification {
    private ProductSpecification augend, addend;

    public AndProductSpecification(
        ProductSpecification augend, ProductSpecification addend) {
        this.augend = augend;
        this.addend = addend;
    }

    public boolean isSatisfiedBy(Product product) {
        return augend.isSatisfiedBy(product) &&
            addend.isSatisfiedBy(product);
    }
}

public class NotProductSpecification extends ProductSpecification {
    private ProductSpecification specToNegate;

    public NotProductSpecification(
        ProductSpecification specToNegate) {
        this.specToNegate = specToNegate;
    }

    public boolean isSatisfiedBy(Product product) {
        return !specToNegate.isSatisfiedBy(product);
    }
}

```

With these classes, I can implement step 5 of the mechanics, which involves changing the implementation of `ProductFinder.belowPriceAvoidingAColor(...)` to use a composition of specifications:

```
public class ProductFinder...
    public List belowPriceAvoidingAColor(float price, Color color) {
        ProductSpecification completeSpec =
            new AndProductSpecification(
                new ProductBelowPricesSpecification(price),
                new NotProductSpecification(
                    new ProductColorSpecification(color)));
        return repository.selectBy(completeSpec);
    }
}
```

After compiling and running tests to confirm that this works, I see that I can continue to follow the advice in step 5 of the mechanics, which suggests that I can inline the `belowPriceAvoidingAColor(...)` method. Here's some client code that gets updated during *Inline Method (117)* [F]:

```
public class ProductFinderTests extends TestCase...
    public void testFindBelowPriceAvoidingAColor() {
        List foundProducts =
        finder.belowPriceAvoidingAColor(9.00f, Color.white);
        ProductSpecification specification =
            new AndProductSpecification(
                new ProductBelowPricesSpecification(9.00f),
                new NotProductSpecification(
                    new ProductColorSpecification(Color.white)));
        List foundProducts = repository.selectBy(specification);
        assertEquals(
            "found 1 non-white product < $9.00",
            1,
            foundProducts.size());
        assertTrue("found fireTruck", foundProducts.contains(fireTruck));

        foundProducts = finder.belowPriceAvoidingAColor(9.00f, Color.red);
        specification =
            new AndProductSpecification(
                new ProductBelowPricesSpecification(9.00f),
                new NotProductSpecification(
                    new ProductColorSpecification(Color.red)));
        foundProducts = repository.selectBy(specification);
        assertEquals(
            "found 1 non-red product < $9.00",
            1,
            foundProducts.size());
        assertTrue("found baseball", foundProducts.contains(baseball));
    }
}
```

8. I repeat step 7 for all remaining product *criteria*. When I'm done, the `ProductFinder` class no longer has any methods, which means I can safely delete it.

9. The final piece of work to complete is an implementation of the refactoring, *Encapsulate Classes with Creation Methods* (27). The classes I want to encapsulate are the public concrete specifications, all of which subclass `ProductSpecification` and implement its public method, `isSatisfiedBy()`. I take advantage of Java's anonymous inner class capability to implement this refactoring. For example, `ProductColorSpecification` currently looks like this:

```
public class ProductColorSpecification extends ProductSpecification {
    private Color colorOfProductToFind;
```

```
public ProductColorSpecification(Color colorOfProductToFind) {
    this.colorOfProductToFind = colorOfProductToFind;
}

public boolean isSatisfiedBy(Product product) {
    return product.getColor().equals(colorOfProductToFind);
}
}
```

I can encapsulate this entire class like so:

```
public abstract class ProductSpecification...
    public static ProductSpecification byColor(final Color color) {
        return new ProductSpecification() {
            public boolean isSatisfiedBy(Product product) {
                return product.getColor().equals(color);
            }
        };
    }
}
```

Notice how there is no longer a need for a public constructor since I can pass the data needed by the concrete specification as a parameter to the *Creation Method*.

Client code that once directly instantiated *ProductColorSpecifications* can now be updated like so:

```
public void testFindByColor()...
    List foundProducts = repository.selectBy(
        new ProductColorSpecification(Color.red));
    ProductSpecification.byColor(Color.red);
```

This same refactoring is used to encapsulate the two *composite specifications* from step 7, and() and not():

```
public abstract class ProductSpecification...
    public ProductSpecification and(
        final ProductSpecification addend) {
        return new AndProductSpecification(this, addend);
    }

    private class AndProductSpecification extends ProductSpecification {
        private ProductSpecification augend, addend;

        public AndProductSpecification(
            ProductSpecification augend, ProductSpecification addend) {
            this.augend = augend;
            this.addend = addend;
        }

        public boolean isSatisfiedBy(Product product) {
            return augend.isSatisfiedBy(product) &&
                addend.isSatisfiedBy(product);
        }
    }

    public static ProductSpecification not(final ProductSpecification specToNegate) {
        return new ProductSpecification() {
            public boolean isSatisfiedBy(Product product) {
                return !specToNegate.isSatisfiedBy(product);
            }
        };
    }
}
```

```
    }  
    };  
}
```

Encapsulating those classes allows me to update caller code like so:

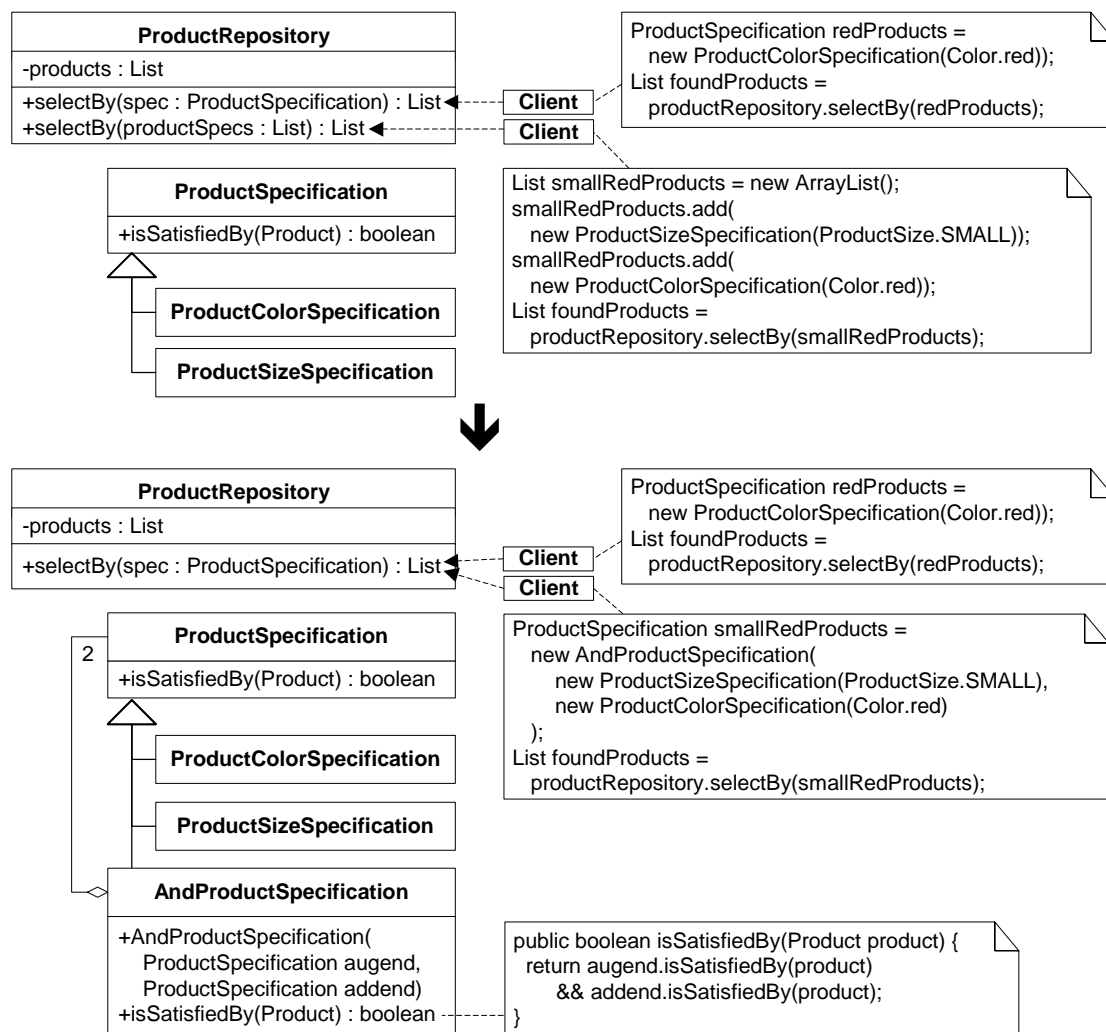
```
public class ProductFinderTests extends TestCase...  
    public void testFindBelowPriceAvoidingAColor()...  
        ProductSpecification specification =  
        new AndProductSpecification(  
            new ProductBelowPriceSpecification(9.00f),  
            new NotProductSpecification(  
                ProductSpecification.byColor(Color.white));  
  
        ProductSpecification specification =  
            ProductSpecification.belowPrice(9.00f).and(  
                ProductSpecification.not(  
                    ProductSpecification.byColor(Color.white)));
```

And that marks the end of this refactoring.

Replace One/Many Distinctions with Composite

A class handles single and multiple objects differently

Handle them identically using a Composite



Motivation

A good reason to refactor to *Composite* [GoF] is to get rid of code that distinguishes between single objects and collections of those objects. There is often simply no good reason to make the distinction. [more to write...]

Communication	Duplication	Simplicity
What behavior a class offers is important to communicate. How the class performs that behavior is less important. By having separate public methods to handle one or many objects, your code communicates too much about how it provides its behavior. Communicate this clearly by using the Composite pattern to handle one/many distinctions uniformly.	One of the primary reasons to refactor to Composite is to remove duplication. Identify separate chunks of code that execute similar behavior on one/many objects and consolidate this code by treating the objects uniformly.	Having separate code to process single objects and collections of objects isn't simple – it is a symptom of code that is needlessly complex. Simplify this code by using a Composite to handle the objects uniformly.

Mechanics

1. Identify code that distinguishes between one and many objects and make a composite class to treat the many as one.
2. Find code that calls code identified in step 1 and that assembles a collection of many objects. Replace that code with code that assembles a composite of the elements.
✓ Compile and test.
3. Repeat step 2 for all callers that distinguish between one and many objects.
4. Repeat all steps until you can delete the code that handles the many objects.
5. Optionally, encapsulate your composite class by applying *Encapsulate Classes with Creation Methods* (23).
✓ Compile and test.

Example

This example deals with `ProductSpecifications` and how they are used to obtain a desired set of `Products` from a `ProductRepository`. The example also illustrates the Specification pattern (see *Replace Conditional Searches with Specification (99)*).

Let's begin by studying some test code for the `ProductRepository`. Before any test can run, a `ProductRepository` (called `repository`) containing toy `Products` must be created:

```
public class ProductRepositoryTest extends TestCase...
    private ProductRepository repository;

    private Product fireTruck =
        new Product("f1234", "Fire Truck",
            Color.red, 8.95f, ProductSize.MEDIUM);

    private Product barbieClassic =
        new Product("b7654", "Barbie Classic",
            Color.yellow, 15.95f, ProductSize.SMALL);

    private Product frisbee =
        new Product("f4321", "Frisbee",
            Color.pink, 9.99f, ProductSize.LARGE);

    private Product baseball =
        new Product("b2343", "Baseball",
            Color.white, 8.95f, ProductSize.NOT_APPLICABLE);

    private Product toyConvertible =
        new Product("p1112", "Toy Porshe Convertible",
            Color.red, 230.00f, ProductSize.NOT_APPLICABLE);

    protected void setUp() {
        repository = new ProductRepository();
        repository.add(fireTruck);
        repository.add(barbieClassic);
        repository.add(frisbee);
        repository.add(baseball);
        repository.add(toyConvertible);
    }
```

The first test we'll study looks for `Products` of a certain color. In the code below, notice the call to `repository.selectBy(...)`, which takes a `ProductSpecification` argument. That argument is obtained via a *Creation Method*, `ProductSpecification.byColor(...)`, which returns a `ProductSpecification` instance:

```
public class ProductRepositoryTest extends TestCase...
    public void testFindByColor() {
        List foundProducts =
            repository.selectBy(ProductSpecification.byColor(Color.red));
        assertEquals("found 2 red products", 2, foundProducts.size());
        assertTrue("found fireTruck", foundProducts.contains(fireTruck));
        assertTrue(
            "found Toy Porshe Convertible",
            foundProducts.contains(toyConvertible));
    }
```

The `repository.selectBy(...)` method being called looks like this:

```
public class ProductRepository...
    private List products = new ArrayList();

    public Iterator iterator() {
        return products.iterator();
    }
```

```

    }

    public List selectBy(ProductSpecification productSpecification) {
        List foundProducts = new ArrayList();
        Iterator products = iterator();
        while (products.hasNext()) {
            Product product = (Product)products.next();
            if (productSpecification.isSatisfiedBy(product))
                foundProducts.add(product);
        }
        return foundProducts;
    }
}

```

Let's now look at another test, which calls a different `repository.selectBy(...)` method. This test assembles a `List` of `ProductSpecifications` in order to select specific kinds of products from the repository:

```

public class ProductRepositoryTest extends TestCase...
    public void testFindByColorSizeAndBelowPrice() {
        List specifications = new ArrayList();
        specifications.add(ProductSpecification.byColor(Color.red));
        specifications.add(ProductSpecification.bySize(ProductSize.SMALL));
        specifications.add(ProductSpecification.byBelowPrice(10.00f));
        List foundProducts = repository.selectBy(specifications);
        assertEquals(
            "found no small red products below $10.00",
            0,
            foundProducts.size());
        specifications.clear();
        specifications.add(ProductSpecification.byColor(Color.red));
        specifications.add(ProductSpecification.bySize(ProductSize.MEDIUM));
        specifications.add(ProductSpecification.byBelowPrice(10.00f));
        foundProducts = repository.selectBy(specifications);
        assertEquals(
            "found firetruck when looking for cheap medium red toys",
            fireTruck,
            foundProducts.get(0));
    }
}

```

The `List`-based `repository.selectBy(...)` method looks like this:

```

public class ProductRepository {
    public List selectBy(List productSpecs) {
        List foundProducts = new ArrayList();
        Iterator products = iterator();
        while (products.hasNext()) {
            Product product = (Product)products.next();
            Iterator specs = productSpecs.iterator();
            boolean satisfiesAllSpecs = true;
            while (specs.hasNext()) {
                ProductSpecification productSpec =
                    ((ProductSpecification) specs.next());
                satisfiesAllSpecs &=
                    productSpec.isSatisfiedBy(product);
            }
            if (satisfiesAllSpecs)
                foundProducts.add(product);
        }
        return foundProducts;
    }
}

```

If you compare the two `repository.selectBy(...)` methods you'll notice duplicate code. Much of that duplication can be removed by having one method use the implementation of the other:

```
public class ProductRepository...
    public List selectBy(ProductSpecification productSpecification) {
        ProductSpecification[] specs = { productSpecification };
        return selectBy(Arrays.asList(specs));
    }

    public List selectBy(List productSpecs)...
        // same implementation as before
```

That certainly helps, but it still leaves two `selectBy(...)` methods. Do we need both? If we *always* called the `selectBy(...)` method that takes the `List` of `ProductSpecifications` we could apply *Inline Method (117)* [F] on the `selectBy(ProductSpecification productSpec)` method. That refactoring would entail updating client code like so:

```
public class ProductRepositoryTest extends TestCase...
    public void testFindByColor()...
        List foundProducts =
        repository.selectBy(ProductSpecification.byColor(Color.red));

        ProductSpecification[] specs =
            { ProductSpecification.byColor(Color.red) };
        List foundProducts =
            repository.selectBy(Arrays.asList(specs));
```

Is that better? Not really, for now we have to create a `ProductSpecification` `List` every time we want to select `Products` using a single `ProductSpecification`. That's unnecessary code, which, in quantity, will only contribute to creating a bloated system. Having two `selectBy(...)` methods is certainly better than having to create unnecessary `Lists`, however, the fundamental problem remains: from a `ProductRepository` client's perspective, single `ProductSpecifications` are handled differently from multiple `ProductSpecifications`.

There is an additional problem. At this point, `repository.selectBy(List ProductSpecs)` expects that you want to select `Products` which match *all* of the criterion passed in the list. That isn't good if you happen to have product criteria like this:

```
product.getColor() != targetColor ||
product.getPrice() < targetPrice
```

The `List`-based `selectBy(...)` method simply wasn't written to handle such `ProductSpecifications`. Changing it to handle “and” and “or” conditions is a job best handled by the *Composite* pattern. Now, if you only had “and” conditions and no “or” conditions, would it still make sense to refactor to *Composite*? My answer would be yes, because:

- It's easy to implement *Composite*
- it means there will be one `selectBy(...)` method that's simpler than the current, `List`-based one
- implementing *Composite* will make it easy to support “or” conditions, “not” conditions and whatever other kinds of operations you may need.

So let's implement this refactoring.

1. I need to create a *Composite* class in this first step. What should this *Composite* class do? It must let a client “and” together two *ProductSpecifications* so that

- two specifications may be treated like one
- client code no longer needs to create a *List* of *ProductSpecifications*.

I’ll call my composite specification *AndProductSpecification* and define it like so:

```
public class AndProductSpecification extends ProductSpecification {
    private ProductSpecification augend, addend;
    public AndProductSpecification(
        ProductSpecification augend,
        ProductSpecification addend) {
        this.augend = augend;
        this.addend = addend;
    }
    public boolean isSatisfiedBy(Product product) {
        return augend.isSatisfiedBy(product) && addend.isSatisfiedBy(product);
    }
}
```

This class will make it possible for clients to call the *selectBy(...)* method that takes a single *ProductSpecification*:

```
public class ProductRepository...
    public List selectBy(ProductSpecification productSpecification) {
        List foundProducts = new ArrayList();
        Iterator products = iterator();
        while (products.hasNext()) {
            Product product = (Product) products.next();
            if (productSpecification.isSatisfiedBy(product))
                foundProducts.add(product);
        }
        return foundProducts;
    }
}
```

2. Next, I replace client code that was producing a *List* of *ProductSpecifications* with code that uses the *AndProductSpecification* and calls the *selectBy(...)* method that takes a single *ProductSpecification* argument. Here’s an example:

```
public void testFindByColorSizeAndBelowPrice() {
    List specifications = new ArrayList();
    specifications.add(ProductSpecification.byColor(Color.red));
    specifications.add(ProductSpecification.bySize(ProductSize.SMALL));
    specifications.add(ProductSpecification.byBelowPrice(10.00f));
    List foundProducts = repository.selectBy(specifications);

    ProductSpecification specification =
        new AndProductSpecification(
            ProductSpecification.byColor(Color.red),
            new AndProductSpecification(
                ProductSpecification.bySize(ProductSize.SMALL),
                ProductSpecification.byBelowPrice(10.00f)));
    List foundProducts = repository.selectBy(specification);
}
```

I compile and test to ensure that the changes work.

3. I repeat step 2 for all client code that formerly created a list of *ProductSpecifications*, replacing it with code that instantiates an *AndProductSpecification* and calls *ProductRepository*’s *selectBy(ProductSpecification productSpecification)*

method.

4. At this point, I'd like to delete the `selectBy(...)` method that accepts a `List` of `ProductSpecifications`. To do so, I simply confirm that no client is calling that method and then I delete it.

In some cases, before you'll be able to delete methods that distinguish between one and many elements, you'll need to create additional composite classes (like an `OrProductSpecification` or a `NotProductSpecification`) by repeating steps 1-3.

The refactoring is now finished since `ProductRepository` no longer has multiple methods for distinguishing between one and many `ProductSpecifications`. To make the code even cleaner, I see that I can encapsulate the composite class, `AndProductSpecification`, using the refactoring, *Encapsulate Classes with Creation Methods* (27). To do so, I'll make the `AndProductSpecification` a private inner class of `ProductSpecification` and provide a *Creation Method* for it like so:

```
public abstract class ProductSpecification...
    private class AndProductSpecification extends ProductSpecification {
        private ProductSpecification augend, addend;
        public AndProductSpecification(
            ProductSpecification augend,
            ProductSpecification addend) {
            this.augend = augend;
            this.addend = addend;
        }
        public boolean isSatisfiedBy(Product product) {
            return augend.isSatisfiedBy(product)
                && addend.isSatisfiedBy(product);
        }
    }

    public ProductSpecification and(           // here's the new Creation method
        final ProductSpecification addend) {
        return new AndProductSpecification(this, addend);
    }
}
```

This change requires that I also update client code, for example:

```
public void testFindByColorSizeAndBelowPrice()...
    ProductSpecification specification =
        new AndProductSpecification(
            ProductSpecification.byColor(Color.red),
            new AndProductSpecification(
                ProductSpecification.bySize(ProductSize.SMALL),
                ProductSpecification.byBelowPrice(10.00f)
            )
        );
    List foundProducts = repository.selectBy(specification);
```

now becomes:

```
public void testFindByColorSizeAndBelowPrice()...
    ProductSpecification specification =
        ProductSpecification.byColor(Color.red).and(
            ProductSpecification.bySize(ProductSize.SMALL).and(
                ProductSpecification.byBelowPrice(10.00f)
            )
        );
    List foundProducts = repository.selectBy(specification);
```

Now I have an easy way of “anding” together `ProductSpecifications` into a single composite, for use by `ProductRepository`’s one and only `selectBy(...)` method.

Compose Method

It isn't easy to understand your method's logic

*Transform the logic into a small number of
intention-revealing steps at the same level of detail*

```
public boolean contains(Component c) {
    Point p = c.getLocation();
    int locX = new Double(p.getX()).intValue();
    int locY = new Double(p.getY()).intValue();
    boolean completelyWithin =
        (locX >= coords[0] &&
         locY >= coords[1] &&
         (locX+CardComponent.WIDTH) <= coords[2]) &&
         (locY+CardComponent.HEIGHT) <= coords[3];
    if (completelyWithin) return true;

    locX = locX+CardComponent.WIDTH;
    locY = locY+CardComponent.HEIGHT;
    boolean partiallyWithin =
        (locX > coords[0] &&
         locY > coords[1] &&
         (locX < coords[2]) &&
         (locY < coords[3]));

    return partiallyWithin;
}
```



```
public boolean contains(Component c) {
    return completelyWithin(c) || partiallyWithin(c);
}

private boolean completelyWithin(Component c) {
    Point p = c.getLocation();
    return (p.x >= coords[0] &&
            p.y >= coords[1] &&
            (p.x + CardComponent.WIDTH) <= coords[2] &&
            (p.y + CardComponent.HEIGHT) <= coords[3]);
}

private boolean partiallyWithin(Component c) {
    Point p = c.getLocation();
    return ((p.x + CardComponent.WIDTH) > coords[0] &&
            (p.y + CardComponent.HEIGHT) > coords[1] &&
            (p.x + CardComponent.WIDTH) < coords[2] &&
            (p.y + CardComponent.HEIGHT) < coords[3]);
}
```

Motivation

Kent Beck once said that some of his best patterns are those that he thought someone would laugh at him for writing. *Composed Method* [Beck] may be such a pattern. A Composed Method is a small, simple method that is easy to understand. Do you write a lot of Composed Methods? I like to think I do, but I often find that I don't, at first. So I have to go back and refactor to this pattern. When my code has many Composed Methods, it tends to be a easy to use, read and extend.

I find myself constantly refactoring to this pattern. For example, just the other day I was debugging a method in some code I've been writing with a friend. The method, called `contains()`, wasn't very complex, but it was complex enough that I had to think about how it was doing its job. I knew this method would be easier to debug if I refactored it first. But my ego wasn't ready for that, just then: I just wanted to get rid of the bug. So, after writing an automated test to demonstrate the bug, I wrote new code in the `contains()` method to fix the bug. That code didn't fix the bug and after two more failed attempts, I was ready to refactor. It wasn't difficult to transform `contains()` into a Composed Method. But after doing so, it was so much easier to follow the logic. And moments after the refactoring, I found and fixed my bug.

Communication	Duplication	Simplicity
It may be clear <i>what</i> a method does but not <i>how</i> the method does what it does. Make the "how" easy to understand by clearly communicating every logical step. You'll often implement part of this refactoring using <i>Extract Method</i> [F].	Duplicate code, whether blatant or subtle, clutters a method's logic. Remove the duplication to make the code smaller and simpler. Doing so often reveals further refactoring opportunities.	Composed Methods often read like English. If your method has too many lines of code, such that you can't easily explain how it does its job, simplify it by extracting logic till it is a Composed Method.

Mechanics

This is one of the most important refactorings I know of. Conceptually, it is also one of the simplest. So you'd think that this refactoring would lead to a simple set of mechanics. In fact, just the opposite is the case. While the steps themselves aren't complex, there is no simple, repeatable set of these steps. But there are guidelines for refactoring to Composed Method, some of which include:

- *Think Small* – Composed Methods are rarely more than 10 lines of code, and are usually more like 5.
- *Remove Duplication* – Reduce the amount of code in the method by getting rid of blatant and/or subtle code duplication.
- *Communicate Intention* – do so with the names of your variables and methods, and by making your code simple.
- *Simplify* – there are many ways to skin a cat. Refactor to the way that is most simple and that best communicates your intention. Simple methods may not be the most highly optimized methods. Don't worry about that. Make your code simple and optimize it later.

- *Similar Levels* – when you break up one method into chunks of behavior, make the chunks operate at similar levels. For example, if you have a piece of detailed conditional logic mixed in with some high-level method calls, you have code at different levels. Push the detail into a new or existing high-level chunk.
- *Group Related Code* – Some code is simply hard to extract into its own method. You can easily see a way to extract part of the code, but the rest remains in the original method. You now have *code at different levels*. In addition, because you have an unnatural split between related fragments of code, your code is harder to follow. In general, look for ways to *group* related code fragments, even if they aren't obvious at first.

Let's now look at three examples of refactoring to Composed Method:

Example 1

I'll start with the game example from the code sketch above. We begin with a single bulky method, called `contains()`, which figures out whether a `Component` is fully or partially contained within a rectangular area:

```
public boolean contains(Component c) {
    Point p = c.getLocation();
    int locX = new Double(p.getX()).intValue();
    int locY = new Double(p.getY()).intValue();
    boolean completelyWithin =
        (locX >= coords[0] &&
         locY >= coords[1] &&
         (locX+CardComponent.WIDTH) <= coords[2]) &&
         (locY+CardComponent.HEIGHT) <= coords[3];
    if (completelyWithin) return true;

    locX = locX+CardComponent.WIDTH;
    locY = locY+CardComponent.HEIGHT;
    boolean partiallyWithin =
        (locX > coords[0] &&
         locY > coords[1] &&
         (locX < coords[2]) &&
         (locY < coords[3]));

    return partiallyWithin;
}
```

Before we get into the refactoring, let's look at one of six test methods for the `contains()` method. The following method tests to see if a card is initially contained within the first player's play area, then moves the card out of the first player's play area and follows that with another test:

```
public void testCardOutOfPlayAreaOne() {
    Hand hand = (Hand)explanations.getCurrentPlayer().getHand();
    Card card = (Card)hand.elements().nextElement();
    CardComponent c = new CardComponent(card,explanations);
    PlayerArea area = explanations.getPlayerArea(0);
    explanations.moveCard(c, area.upperLeft());
    assertEquals("area contains card", true, area.contains(c));

    explanations.moveCard(c, CardComponent.WIDTH + 10, CardComponent.HEIGHT + 10);
    assertEquals("area does not contain card", false, area.contains(c));
}
```

The above test, and the other five tests, all pass (or "run green") before I begin refactoring. I run these tests after each of the small steps I am about to do below.

To begin, my first impulse is to make the `contains()` method *smaller*. That leads me to look at the conditional represented by the variable, `completelyWithin`:

```
boolean completelyWithin =
    (locX >= coords[0] &&
     locY >= coords[1] &&
     (locX+CardComponent.WIDTH) <= coords[2]) &&
    (locY+CardComponent.HEIGHT) <= coords[3];
```

While that variable helps make it clear what the conditional logic does, the `contains()` method would be smaller and easier to read if this fragment were in its own method. So I start with an Extract Method:

```
public boolean contains(Component c) {
    Point p = c.getLocation();
    int locX = new Double(p.getX()).intValue();
    int locY = new Double(p.getY()).intValue();
    if (completelyWithin(locX, locY)) return true;

    locX = locX+CardComponent.WIDTH;
    locY = locY+CardComponent.HEIGHT;
    boolean partiallyWithin =
        (locX > coords[0] &&
         locY > coords[1] &&
         (locX < coords[2]) &&
         (locY < coords[3]));
    return partiallyWithin;
}

private boolean completelyWithin(int locX, int locY) {
    return (locX >= coords[0] &&
            locY >= coords[1] &&
            (locX+CardComponent.WIDTH) <= coords[2]) &&
            (locY+CardComponent.HEIGHT) <= coords[3];
}
```

Next, after seeing a similar temporary variable, called `partiallyWithin`, I do another Extract Method:

```
public boolean contains(Component c) {
    Point p = c.getLocation();
    int locX = new Double(p.getX()).intValue();
    int locY = new Double(p.getY()).intValue();
    if (completelyWithin(locX, locY)) return true;
    locX = locX+CardComponent.WIDTH;
    locY = locY+CardComponent.HEIGHT;
    return partiallyWithin(locX, locY);
}

private boolean partiallyWithin(int locX, int locY) {
    return (locX > coords[0] &&
            locY > coords[1] &&
            (locX < coords[2]) &&
            (locY < coords[3]));
}
```

The `contains()` method is now smaller and simpler, but it still seems cluttered with variable assignments. I notice that the assignments to `locX` and `locY` are performed simply for use by the new methods, `completelyWithin()` and `partiallyWithin()`. I decide to let those methods deal with the `locX` and `locY` assignments. The easiest way to do this is to just pass the `Point` variable, `p`, to each of the methods:

```
public boolean contains(Component c) {
```

```

    Point p = c.getLocation();
    if (completelyWithin(p)) return true;
    return partiallyWithin(p);
}

```

Now, the `contains()` method is really looking smaller and simpler. I feel like I'm done. But then I look at that first line of code:

```
Point p = c.getLocation();
```

The level of that code seems wrong – it is a detail, while the rest of the code in the method represents core pieces of logic. The two methods I'm calling each need the `Point` variable. But each of those methods could easily obtain the `Point` variable if I just sent them `Component c`. I consider doing that, but then I worry about violating the rule of doing things *once and only once*. For if I pass variable `c`, the `Component`, to each method, each method will have to contain code to obtain a `Point` from `c`, instead of just getting one passed in directly.

Hmmmm. What is my real goal here? Is it more important to get the levels of the code right or to say things once and only once? After some reflection, I realize that my goal is to produce a method that can be read and understood in seconds. But as it stands, that first line of code takes away from the readability and simplicity of the method. So I push down the code to obtain a `Point` into the two called methods and end up with the following:

```

public boolean contains(Component c) {
    return completelyWithin(c) || partiallyWithin(c);
}

private boolean completelyWithin(Component c) {
    Point p = c.getLocation();
    int locX = new Double(p.x).intValue();
    int locY = new Double(p.y).intValue();
    return (locX >= coords[0] &&
            locY >= coords[1] &&
            (locX + CardComponent.WIDTH) <= coords[2] &&
            (locY + CardComponent.HEIGHT) <= coords[3]);
}

private boolean partiallyWithin(Component c) {
    Point p = c.getLocation();
    int locX = new Double(p.x).intValue() + CardComponent.WIDTH;
    int locY = new Double(p.y).intValue() + CardComponent.HEIGHT;
    return (locX > coords[0] &&
            locY > coords[1] &&
            (locX < coords[2]) &&
            (locY < coords[3]));
}

```

Now I think I'm really done. But whenever you think you're really done, you're not. A reviewer of this refactoring, named Andrew Swan, observed that I was converting `p.x` and `p.y` to ints, when they are already ints! So this lead to a further simplification:

```

public boolean contains(Component c) {
    return completelyWithin(c) || partiallyWithin(c);
}

private boolean completelyWithin(Component c) {
    Point p = c.getLocation();
    return (p.x >= coords[0] &&
            p.y >= coords[1] &&
            (p.x + CardComponent.WIDTH) <= coords[2] &&
            (p.y + CardComponent.HEIGHT) <= coords[3]);
}

```

```
private boolean partiallyWithin(Component c) {  
    Point p = c.getLocation();  
    return ((p.x + CardComponent.WIDTH) > coords[0] &&  
            (p.y + CardComponent.HEIGHT) > coords[1] &&  
            (p.x + CardComponent.WIDTH) < coords[2] &&  
            (p.y + CardComponent.HEIGHT) < coords[3]);  
}
```

Example 2

```
public static Vector wrap(String s) {
    Vector wrapVector = new Vector();
    String words;
    String word;
    int lastPos;
    do {
        if (s.length() > 16) {
            words="";
            word="";
            lastPos=0;
            for (int i=0;i<16;i++) {
                if (s.charAt(i)==' ' || s.charAt(i)=='-') {
                    words+=word+s.charAt(i);
                    lastPos = i+1;
                    word="";
                } else word+=s.charAt(i);
            }
            if (lastPos==0) {
                // Rare case that there was no space or dash, insert one and break
                words+=word+"-";
                lastPos=16;
            }
            wrapVector.addElement(words);
            s = s.substring(lastPos, s.length());
        }
    } while (s.length() > 16);
    if (s.length()>0) wrapVector.addElement(s);
    return wrapVector;
}
```



```
public static Vector wrap(StringBuffer cardText) {
    Vector wrapLines = new Vector();
    while (cardText.length() > 0)
        wrapLines.addElement(extractPhraseFrom(cardText));
    return wrapLines;
}

private static String extractPhraseFrom(StringBuffer cardText) {
    StringBuffer phrase = new StringBuffer("");
    StringBuffer word = new StringBuffer("");
    final int MAXCHARS = Math.min(MAX_LINE_WIDTH, cardText.length());
    for (int i=0; i<MAXCHARS; i++) {
        addCharacterTo(word, cardText.charAt(i));
        if (isCompleteWord(word, cardText))
            addCompleteWordTo(phrase, word);
    }
    addRemainingWordTo(phrase, word);
    removePhraseFrom(cardText, phrase);
    return phrase.toString();
}

private static boolean addCharacterTo(StringBuffer word, char character) ...
private static boolean isCompleteWord(StringBuffer word, StringBuffer cardText) ...
private static void addCompleteWordTo(StringBuffer phrase, StringBuffer word) ...
private static void addRemainingWordTo(StringBuffer phrase, StringBuffer word) ...
private static void removePhraseFrom(StringBuffer cardText, StringBuffer phrase)...
```

In a game I've been writing with a friend, text needs to be displayed on graphical cards. The text is typically too long to fit on one line of each card, so it must be displayed on multiple lines

of each card. To enable this behavior, we test-first programmed a `wrap()` method. Here are a few of the tests:

```
public void accumulateResult(String testString) {
    int i = 0;
    for (Enumeration e = CardComponent.wrap(testString).elements(); e.hasMoreElements();)
        result[i++] = (String)e.nextElement();
}

public void testWrap() {
    accumulateResult("Developers Misunderstand Requirements");
    assertEquals("First line", "Developers ", result[0]);
    assertEquals("Second line", "Misunderstand ", result[1]);
    assertEquals("Third line", "Requirements", result[2]);
}

public void testWrap2() {
    accumulateResult("Stories Are Too Complex");
    assertEquals("First line", "Stories Are Too ", result[0]);
    assertEquals("Second line", "Complex", result[1]);
}

public void testWrap3() {
    accumulateResult("Intention-Revealing Code");
    assertEquals("First line", "Intention-", result[0]);
    assertEquals("Second line", "Revealing Code", result[1]);
}
```

With these tests in place, I can work on refactoring the following bloated method:

```
public static Vector wrap(String s) {
    Vector wrapVector = new Vector();
    String words;
    String word;
    int lastPos;
    do {
        if (s.length() > 16) {
            words="";
            word="";
            lastPos=0;
            for (int i=0;i<16;i++) {
                if (s.charAt(i)==' ' || s.charAt(i)=='-') {
                    words+=word+s.charAt(i);
                    lastPos = i+1;
                    word="";
                } else word+=s.charAt(i);
            }
            if (lastPos==0) {
                // Rare case that there was no space or dash, insert one and break
                words+=word+"-";
                lastPos=16;
            }
            wrapVector.addElement(words);
            s = s.substring(lastPos, s.length());
        }
    } while (s.length() > 16);
    if (s.length()>0) wrapVector.addElement(s);
    return wrapVector;
}
```

The first thing I notice is that we have some blatant duplicate logic: the line, `s.length() > 16`, appears in a conditional statement at line 6 and at the end of the while statement. No good. I experiment with removing this duplication by using a while loop instead of a `do..while` loop. The tests confirm that the experiment works:

```
public static Vector wrap(String s) {
    Vector wrapVector = new Vector();
    String words;
```

```

String word;
int lastPos;
while (s.length() > 16) {
    words="";
    word="";
    lastPos=0;
    for (int i=0;i<16;i++)
        if (s.charAt(i)==' ' || s.charAt(i)=='-') {
            words+=word+s.charAt(i);
            lastPos = i+1;
            word="";
        } else word+=s.charAt(i);
    if (lastPos==0) {
        // Rare case that there was no space or dash, insert one and break
        words+=word+"-";
        lastPos=16;
    }
    wrapVector.addElement(words);
    s = s.substring(lastPos, s.length());
}
if (s.length()>0) wrapVector.addElement(s);
return wrapVector;
}

```

Next I notice more duplication. At two places in the middle of the method, the code says:

```
word+=s.charAt(i).
```

By consolidating this logic, I see a way to simplify a conditional statement:

```

for (int i=0;i<16;i++) {
    word+=s.charAt(i); // now we say this only once
    if (s.charAt(i)==' ' || s.charAt(i)=='-') {
        words+=word;
        lastPos = i+1;
        word="";
    } // else statement is no longer needed
}

```

Additional duplicate logic doesn't jump out at me just yet, so I continue to look (I know it is there!). I wonder about the variable, `lastPos`. What does it store? Can I figure out what the value of `lastPos` would be, without having to declare and set a variable for it? After a little bit of study, I try some experiments. Gradually it dawns on me that `words.length()` contains the exact value as that held by `lastPos`. This allows me to get rid of another variable, and all of the assignments to it:

```

public static Vector wrap(String s) {
    Vector wrapVector = new Vector();
    String words;
    String word;
    while (s.length() > 16) {
        words="";
        word="";
        for (int i=0;i<16;i++) {
            word+=s.charAt(i);
            if (s.charAt(i)==' ' || s.charAt(i)=='-') {
                words+=word;
                word="";
            }
        }
        if (words.length() == 0) // if no space or dash, insert one
            words+=word+"-";
        wrapVector.addElement(words);
        s = s.substring(words.length(), s.length());
    }
}

```

```

        if (s.length() > 0) wrapVector.addElement(s);
        return wrapVector;
    }

```

The code is definitely getting smaller and more manageable. But the body of the while method still seems big and bulky. I decide to *Extract Method* [F]:

```

public static Vector wrap(String s) {
    Vector wrapVector = new Vector();
    String words;
    while (s.length() > 16) {
        words = extractPhraseFrom(s);
        wrapVector.addElement(words);
        s = s.substring(words.length(), s.length());
    }
    if (s.length() > 0) wrapVector.addElement(s);
    return wrapVector;
}

private static String extractPhraseFrom(String cardText) {
    String phrase = "";
    String word = "";
    for (int i = 0; i < 16; i++) {
        word += cardText.charAt(i);
        if (cardText.charAt(i) == ' ' || cardText.charAt(i) == '-') {
            phrase += word;
            word = "";
        }
    }
    if (phrase.length() == 0) // no found space or dash, insert dash
        phrase += word + "-";
    return phrase;
}

```

We're making progress. But I'm still not happy with the wrap() method: I don't like the fact that the code is adding elements to the wrapVector both inside and outside the while loop and I also don't like the mysterious line that changes the value of the String "s" (which is a bad name for a variable that holds on to a card's text):

```

s = s.substring(words.length(), s.length());

```

So I ask myself how I can make this logic clearer? Given some card text, I would like my code to show how the text is broken up into pieces, added to a collection and returned. I decide that the best way to achieve this objective is to push all code that is responsible for creating a "phrase" into the extractPhraseFrom() method. I hope to end up with a while loop that has one line of code.

My first step is to rename and change the type of the String variable, s. I call it cardText and change it to be StringBuffer, since it will be altered by the extractPhraseFrom() method. This change requires that I make all callers of wrap() pass in a StringBuffer instead of a String. As I go about doing this work, I see that I can also get rid of the temporary variable, word, leaving the following:

```

public static Vector wrap(StringBuffer cardText) {
    Vector wrapVector = new Vector();
    while (cardText.length() > 16) {
        wrapVector.addElement(extractPhraseFrom(cardText));
        cardText.delete(0, words.length());
    }
    if (cardText.length() > 0) wrapVector.addElement(cardText.toString());
    return wrapVector;
}

```


Now I must figure out how to push the fragmented pieces of phrase-construction logic down into the `extractPhraseFrom()` method. My tests give me a lot of confidence as I go about this work. First, I go for the low-hanging fruit: the code that deletes a substring from `cardText` can easily be moved to `extractPhraseFrom()`, which yields the following:

```
public static Vector wrap(StringBuffer cardText) {
    Vector wrapVector = new Vector();
    while (cardText.length() > 16)
        wrapVector.addElement(extractPhraseFrom(cardText));
    if (cardText.length() > 0) wrapVector.addElement(cardText.toString());
    return wrapVector;
}
```

Now, I've just got the line of code after the while loop to worry about:

```
if (cardText.length() > 0) wrapVector.addElement(cardText.toString());
```

How can I get that code to live in the `extractPhraseFrom()` method? I study the while loop and see that I'm looping on a magic number, 16. First, I decide to make a constant for that number, called `MAX_LINE_WIDTH`. Then, as I continue to study the loop, I wonder why the `wrap()` method has two conditionals fragments that check `cardText.length()`, (one in the while loop and one after the while loop). I want to remove that duplication. I decide to change the while loop to do its thing while `cardText.length() > 0`.

This last change requires a few changes to the `extractPhraseFrom` method to make it capable of handling the case when a line of text isn't greater than 16 characters (now called `MAX_LINE_WIDTH`). Once the tests confirm that everything is working, `wrap()` now feels like a Composed Method, while `extractPhraseFrom()` is getting there. Here's what we have now:

```
public static Vector wrap(StringBuffer cardText) {
    Vector wrapLines = new Vector();
    while (cardText.length() > 0)
        wrapLines.addElement(extractPhraseFrom(cardText));
    return wrapLines;
}

private static String extractPhraseFrom(StringBuffer cardText) {
    String phrase = "";
    String word = "";
    final int MAX_CHARS = Math.min(MAX_LINE_WIDTH, cardText.length());
    for (int i=0; i<MAX_CHARS; i++) {
        word += cardText.charAt(i);
        if (cardText.charAt(i)==' ' || cardText.charAt(i)=='-' ||
            cardText.toString().endsWith(word)) {
            phrase += word;
            word = "";
        }
    }
    if (phrase.length() == 0)
        phrase = word + "-";
    cardText.delete(0, phrase.length());
    return phrase;
}
```

This code is simpler than the original, so we could stop here. But I'm not altogether happy with the `extractPhraseFrom()` method. It's not a Composed Method, so I'm drawn to continue refactoring it. What's wrong with it? Well, there's a lot of conditional logic in it, and that conditional logic doesn't communicate very well. For example, what does this mean:

```

if (cardText.charAt(i)==' ' || cardText.charAt(i)=='-' ||
    cardText.toString().endsWith(word)) {
    phrase += word;
    word="";
}

```

Since my pair and I wrote that code, I know that it means, “*if we’ve found a complete word, then add the word to the phrase, and blank out the word variable so we can find the next word.*” But the next reader will have to figure that out. So I’ll make the intention clear, by using Extract Method (which also requires changing some variables from Strings to StringBuffers):

```

private static String extractPhraseFrom(StringBuffer cardText) {
    StringBuffer phrase = new StringBuffer("");
    StringBuffer word = new StringBuffer("");
    final int MAXCHARS = Math.min(MAX_LINE_WIDTH, cardText.length());
    for (int i=0; i<MAXCHARS; i++) {
        word.append(cardText.charAt(i));
        if (isCompleteWord(word, cardText)) // note how more intention-revealing this is
            addCompleteWordTo(phrase, word); // same for this line
    }
    if (phrase.length() == 0)
        phrase.append(word + "-");
    cardText.delete(0, phrase.length());
    return phrase.toString();
}

private static boolean isCompleteWord(StringBuffer word, StringBuffer cardText) {
    return (word.charAt(word.length()-1) == ' ' || word.charAt(word.length()-1) == '-' ||
        cardText.toString().endsWith(word.toString()));
}

private static void addCompleteWordTo(StringBuffer phrase, StringBuffer word) {
    phrase.append(word);
    word.delete(0, word.length());
}

```

We’re getting closer. But I still don’t like the cryptic conditional statement that comes after the for loop. So I apply *Extract Method* to it:

```

private static String extractPhraseFrom(StringBuffer cardText) {
    StringBuffer phrase = new StringBuffer("");
    StringBuffer word = new StringBuffer("");
    final int MAXCHARS = Math.min(MAX_LINE_WIDTH, cardText.length());
    for (int i=0; i<MAXCHARS; i++) {
        word.append(cardText.charAt(i));
        if (isCompleteWord(word, cardText))
            addCompleteWordTo(phrase, word);
    }
    addRemainingWordTo(phrase, word); // now this code communicates intention
    cardText.delete(0, phrase.length());
    return phrase.toString();
}

private static void addRemainingWordTo(StringBuffer phrase, StringBuffer word) {
    if (phrase.length() == 0)
        phrase.append(word + "-");
}

```

The `extractPhraseFrom()` method is now 10 lines of code and reads a lot more like English. But it is still uneven! Consider these two lines of code:

```

word.append(cardText.charAt(i));

cardText.delete(0, phrase.length());

```

Both of these lines aren't complicated, but compared with the other code, which reads like English, these bits of code stick out, demanding that the reader concentrates to understand them. So I push myself to extract these 2 lines of code into 2 intention-revealing methods: `addCharacterTo()` and `removePhraseFrom()`. This yields a Composed Method:

```
private static String extractPhraseFrom(StringBuffer cardText) {
    StringBuffer phrase = new StringBuffer("");
    StringBuffer word = new StringBuffer("");
    final int MAXCHARS = Math.min(MAX_LINE_WIDTH, cardText.length());
    for (int i=0; i<MAXCHARS; i++) {
        addCharacterTo(word, cardText.charAt(i));
        if (isCompleteWord(word, cardText))
            addCompleteWordTo(phrase, word);
    }
    addRemainingWordTo(phrase, word);
    removePhraseFrom(cardText, phrase);
    return phrase.toString();
}
```

My tests run green and I'm satisfied.

Example 3

```
private void paintCard(Graphics g) {
    Image image = null;
    if (card.getType().equals("Problem")) {
        image = explanations.getGameUI().problem;
    } else if (card.getType().equals("Solution")) {
        image = explanations.getGameUI().solution;
    } else if (card.getType().equals("Value")) {
        image = explanations.getGameUI().value;
    }
    g.drawImage(image,0,0,explanations.getGameUI());

    if (highlight)
        paintCardHighlight(g);
    paintCardText(g);
}
```



```
private void paintCard(Graphics g) {
    paintCardImage(g);
    paintCardHighlight(g);
    paintCardText(g);
}
```

The above, original `paintCard()` method isn't long, nor is it complicated. It paints a card image, checks a flag to see if it must paint a card highlight, and then paints text onto the card. Painting the card highlight and card text are performed by the methods, `paintCardHighlight()` and `paintCardText()`. But the code that paints the card image lives not in a separate method but in the `paintCard()` method itself. So? Well, consider the refactored version of `paintCard()`. I can look at the refactored version and know what it does in 2 seconds, while I have to spend a few brain cycles to figure out what the previous version does. Trivial difference? *No*, not when you consider how much simpler an entire system is when it consists of *many* composed methods, like `paintCard()`.

So what was the *smell* that led to this refactoring? *Code at different levels*: raw code mixed with higher-level code. When the method contains code at the same levels, it is easier to read and understand. As the guidelines in the mechanics section say, above, Composed Methods tend to have code at the *same level*.

Implementing this refactoring was incredibly easy. I did Extract Method [F] as follows:

```
private void paintCard(Graphics g) {
    paintCardImage(g);
    if (highlight)
        paintCardHighlight(g);
    paintCardText(g);
}

private void paintCardImage(Graphics g) {
    Image image = null;
    if (card.getType().equals("Problem")) {
        image = explanations.getGameUI().problem;
    } else if (card.getType().equals("Solution")) {
        image = explanations.getGameUI().solution;
    } else if (card.getType().equals("Value")) {
        image = explanations.getGameUI().value;
    }
    g.drawImage(image,0,0,explanations.getGameUI());
}
```

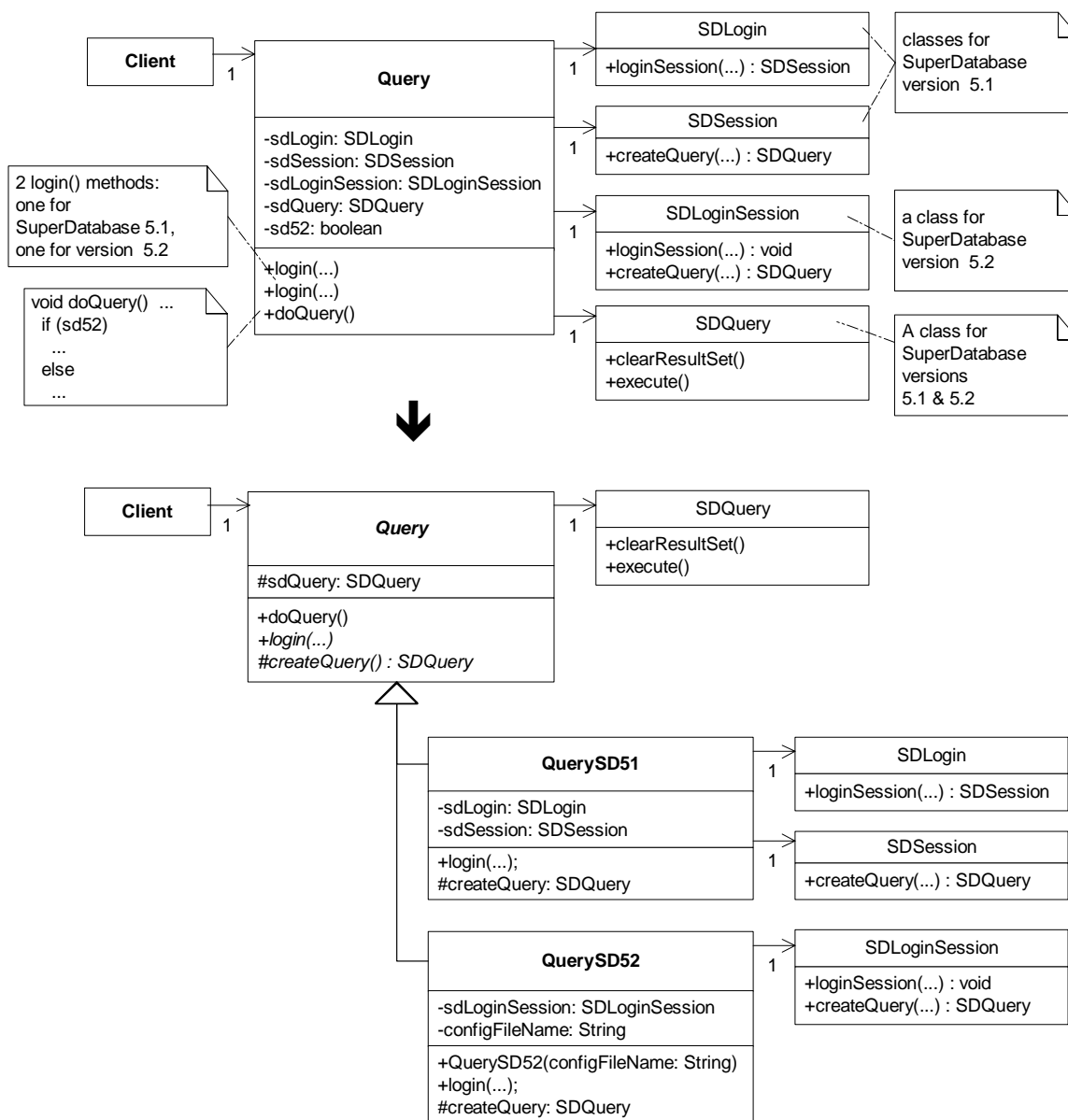
To finish this refactoring, I took the sole conditional statement in the method (`if (highlight)...`) and pushed it down into the `paintCardHighlight()` method. Why? I wanted the reader to simply see three steps: paint image, highlight image and paint card text. The detail of whether or not we do highlight the card isn't important to me – the reader can find that out if they look. But if that confuses other programmers, I'd be happy to see the method renamed to `paintCardHighlightIfNecessary(g)` or something similar.

```
private void paintCard(Graphics g) {  
    paintCardImage(g);  
    paintCardHighlight(g);  
    paintCardText(g);  
}
```

Separate Versions with Adapters

One class adapts multiple versions of a component, library, API or other entity

Write Adapters for each version



Motivation

While software must often support multiple versions of a component, library or API, code that handles these versions doesn't have to be a confusing mess. And yet, I routinely encounter code that attempts to handle multiple versions of something by overloading classes with version-specific state variables, constructors and methods. Accompanying such code are comments like "this is for version X – please delete this code when we move to version Y!" Sure, like that's ever gonna happen. Most programmers won't delete the version X code for fear that something they don't know about still relies on it. So the comments don't get deleted and many versions supported by the code remain in the code.

Now consider an alternative: for each version of something you need to support, create a separate class. The class name could even include the version number of what it supports, to be really explicit about what it does. We call such classes *Adapters* [GoF]. Adapters implement a common interface and are responsible for functioning correctly with one (and usually only one) version of some code. Adapters make it easy for client code to swap in support for one library or API version, or another. And programmers routinely rely on runtime information to configure their programs with the correct Adapter.

I refactor to Adapters fairly often. I like Adapters because they let me decide how I want to communicate with other people's code. In a fast-changing world, Adapters help me stay insulated from the highly useful but rapidly changing APIs, such as those springing eternally from the open-source world.

In several of the refactorings in this catalog, I assert the importance of not refactoring to a pattern too quickly in order to avoid overengineering. There must be a genuine need to refactor to a pattern, such as an overabundance of conditional logic, code bloat, duplication or unnecessary complexity. However, in the case of code that handles multiple versions of a component, library, API, etc., I often find compelling reasons to refactor to Adapters *early*, since not doing so can lead to a propagation of conditional or version-dependent logic throughout a system. So, while I'm not suggesting you adapt too early, be on guard for any complexity or propagating conditionality or maintenance issues accruing from code written to handle multiple versions of something. *Adapt* early and often so that it's easy to use or phase out various versions of code.

Communication	Duplication	Simplicity
A class that mixes together version-specific state variables, constructors and methods doesn't effectively communicate how each version is different or similar. Communicate version differences by isolating the differences in separate Adapter classes. Communicate how versions are similar by making each Adapter implement a common interface – either by subclassing an abstract class, implementing the same interface or a combination thereof.	When each version of a component, library, API, etc., isn't isolated in its own Adapter, but is instead accessed directly or through a single class, there tends to be the same repeating chunks of conditional logic that make version-specific calls to code. Such duplication bloats a class and makes the code harder to follow.	When a class is responsible for functioning correctly with several versions of some other code, it is rarely simple. Version-specific code tends to bloat the single class and leads to conditional logic in the client code that uses it. Adapters provide a simple way to isolate versions and give clients a simple interface to every version.

Mechanics

There are different ways to go about this refactoring, depending on what your code looks like before you begin. For example, if you have a class that uses a lot of conditional logic to handle multiple versions of something, it's likely that you can create Adapters for each version by repeatedly applying *Replace Conditional with Polymorphism* (255) [F]. If you have a case like that shown in the code sketch – in which a single class supports multiple versions of something by containing version-specific variables and methods, you'll refactor to Adapter using a slightly different approach. I'll outline the mechanics for this latter scenario.

1. Identify the overburdened class (we'll call this class, "V").
2. Apply *Extract Subclass* (330) [F] or *Extract Class* (149) [F] for a single version of the multiple versions supported by V. Copy or move all instance variables and methods used exclusively for that version into the new class.

To do this, you may need to make some private members of V public or protected. It may also be necessary to initialize some instance variables via a constructor in your new class, which will necessitate updates to callers of the new constructor.

✓ Compile and test.

3. Repeat step 2 until there is no more version-specific code in V.
4. Remove any duplication found in the new classes, by applying refactorings like *Pull Up Method* (322) [F] and *Form Template Method* (345) [F].

✓ Compile and test.

Example

The code we'll refactor in this example, which was depicted in the code sketch above, is based on real-world code that handles queries to a database using a third party library. To protect the innocent, I've renamed that library "SD," which stands for SuperDatabase.

1. We begin by identifying a class that is overburdened with support for multiple versions of SuperDatabase. This class, called `Query`, provides support for SuperDatabase versions 5.1 and 5.2, which means it is already an *Adapter* to the SuperDatabase code. It just happens to be an Adapter that is *adapting* too much.

In the code listing below, notice the version-specific instance variables, duplicate `login()` methods and conditional code in `doQuery()`:

```
public class Query . . .
    private SDLogin sdLogin;           // needed for SD version 5.1
    private SDSession sdSession;       // needed for SD version 5.1
    private SDLoginSession sdLoginSession; // needed for SD version 5.2
    private boolean sd52;               // tells if we're running under SD 5.2
    private SDQuery sdQuery;           // this is needed for SD versions 5.1 & 5.2

    // this is a login for SD 5.1
    // NOTE: remove this when we convert all applications to 5.2
    public void login(String server, String user, String password) throws QueryException {
        sd52 = false;
        try {
```



```

        sdSession = sdLogin.loginSession(server, user, password);
    } catch (SDLoginFailedException lfe) {
        throw new QueryException(QueryException.LOGIN_FAILED,
            "Login failure\n" + lfe, lfe);
    } catch (SDSocketInitFailedException ife) {
        throw new QueryException(QueryException.LOGIN_FAILED,
            "Socket fail\n" + ife, ife);
    }
}

// 5.2 login
public void login(String server, String user, String password, String
sdConfigFileName) throws QueryException {
    sd52 = true;
    sdLoginSession = new SDLoginSession(sdConfigFileName, false);
    try {
        sdLoginSession.loginSession(server, user, password);
    } catch (SDLoginFailedException lfe) {
        throw new QueryException(QueryException.LOGIN_FAILED,
            "Login failure\n" + lfe, lfe);
    } catch (SDSocketInitFailedException ife) {
        throw new QueryException(QueryException.LOGIN_FAILED,
            "Socket fail\n" + ife, ife);
    } catch (SDNotFoundException nfe) {
        throw new QueryException(QueryException.LOGIN_FAILED,
            "Not found exception\n" + nfe, nfe);
    }
}

public void doQuery() throws QueryException {
    if (sdQuery != null)
        sdQuery.clearResultSet();
    if (sd52)
        sdQuery = sdLoginSession.createQuery(SDQuery.OPEN_FOR_QUERY);
    else
        sdQuery = sdSession.createQuery(SDQuery.OPEN_FOR_QUERY);
    executeQuery();
}

```

2. Because Query doesn't already have subclasses, I decide to apply *Extract Subclass (330)* [F] to isolate code that handles SuperDatabase 5.1 queries. My first step is to define the subclass and create a constructor for it:

```

class QuerySD51 extends Query {
    public QuerySD51() {
        super();
    }
}

```

Next, I find all calls to the constructor of Query and, where appropriate, change the code to call the QuerySD51 constructor. For example, I find the following:

```

public void loginToDatabase(String db, String user, String password)...
    query = new Query();
    try {
        if (usingSDVersion52()) {
            query.login(db, user, password, getSD52ConfigFileName()); // Login to SD 5.2
        } else {
            query.login(db, user, password); // Login to SD 5.1
        }
        ...
    } catch (QueryException qe)...

```

And change this to:

```

public void loginToDatabase(String db, String user, String password)...
    try {

```

```

        if (usingSDVersion52()) {
            query = new Query();
            query.login(db, user, password, getSD52ConfigFileName()); // Login to SD 5.2
        } else {
            query = new QuerySD51();
            query.login(db, user, password); // Login to SD 5.1
        }
        ...
    } catch (QueryException qe) {

```

Next, I apply *Push Down Method* (328) [F] and *Push Down Field* (329) [F] to outfit QuerySD51 with the methods and instance variables it needs. During this step, I have to be careful to consider the clients that make calls to public Query methods, for if I move a public method like login() from Query to a QuerySD51, the caller will not be able to call the public method unless its type is changed to QuerySD51. Since I don't want to make such changes to client code, I proceed cautiously, sometimes copying and modifying public methods instead of completely removing them from Query. While I do this, I generate duplicate code, but that doesn't bother me now - I'll get rid of the duplication in step 5.

```

class Query...
    private SDLogin sdLogin;
    private SDSession sdSession;
    protected SDQuery sdQuery;

    // this is a login for SD 5.1
    public void login(String server, String user, String password) throws QueryException {
        // I make this a do-nothing method
    }

    public void doQuery() throws QueryException {
        if (sdQuery != null)
            sdQuery.clearResultSet();
        if (sd52)
            sdQuery = sdLoginSession.createQuery(SDQuery.OPEN_FOR_QUERY);
        else
            sdQuery = sdSession.createQuery(SDQuery.OPEN_FOR_QUERY);
            executeQuery();
    }
}

class QuerySD51 {
    private SDLogin sdLogin;
    private SDSession sdSession;

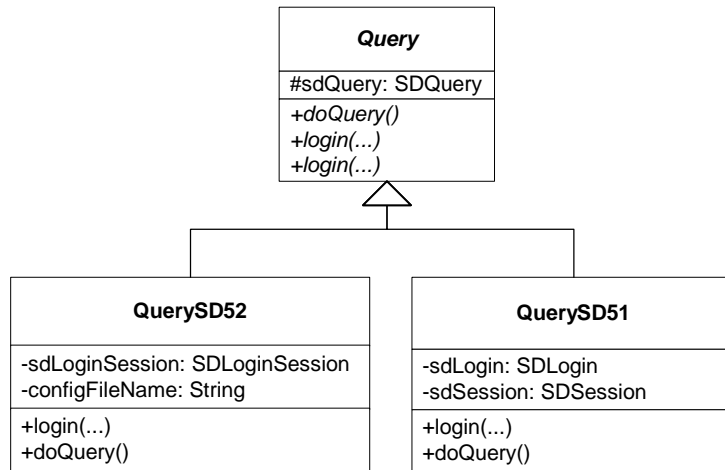
    public void login(String server, String user, String password) throws QueryException {
        sd52 = false;
        try {
            sdSession = sdLogin.loginSession(server, user, password);
        } catch (SDLoginFailedException lfe) {
            throw new QueryException(QueryException.LOGIN_FAILED,
                "Login failure\n" + lfe, lfe);
        } catch (SDSocketInitFailedException ife) {
            throw new QueryException(QueryException.LOGIN_FAILED,
                "Socket fail\n" + ife, ife);
        }
    }

    public void doQuery() throws QueryException {
        if (sdQuery != null)
            sdQuery.clearResultSet();
        if (sd52)
            sdQuery = sdLoginSession.createQuery(SDQuery.OPEN_FOR_QUERY);
        else
            sdQuery = sdSession.createQuery(SDQuery.OPEN_FOR_QUERY);
        executeQuery();
    }
}

```

I compile and test that QuerySD51 works. No problems.

3. Next, I repeat step 2 to create QuerySD52. Along the way, I can make the Query class abstract, along with the doQuery() method. Here's what I have now:



Query is now free of version-specific code, but it is not free of duplicate code.

4. I now go on a mission to remove duplication. I quickly find some in the two implementations of doQuery():

```

abstract class Query...
    public abstract void doQuery() throws QueryException;

class QuerySD51...
    public void doQuery() throws QueryException {
        if (sdQuery != null)
            sdQuery.clearResultSet();

        sdQuery = sdSession.createQuery(SDQuery.OPEN_FOR_QUERY);
        executeQuery();
    }

class QuerySD52...
    public void doQuery() throws QueryException {
        if (sdQuery != null)
            sdQuery.clearResultSet();

        sdQuery = sdLoginSession.createQuery(SDQuery.OPEN_FOR_QUERY);
        executeQuery();
    }
    
```

Each of the above methods simply initializes the sdQuery instance in a different way. This means that I can apply *Introduce Polymorphic Creation with Factory Method* (43) and *Form Template Method* (345) [F] to create a single superclass version of doQuery():

```

public abstract class Query ...
    protected abstract SDQuery createQuery();                // a Factory Method [GoF]

    public void doQuery() throws QueryException {            // a Template Method [GoF]
        if (sdQuery != null)
            sdQuery.clearResultSet();
        sdQuery = createQuery();                            // call to the Factory Method
    }
    
```

```
        executeQuery();
    }

class QuerySD51...
    protected SDQuery createQuery() {
        return sdSession.createQuery(SDQuery.OPEN_FOR_QUERY);
    }

class QuerySD52...
    protected SDQuery createQuery() {
        return sdLoginSession.createQuery(SDQuery.OPEN_FOR_QUERY);
    }
```

After compiling and testing the changes, I now face a more obvious duplication problem: Query still declares public method for the SD 5.1 & 5.2 login() methods, even though they don't do anything anymore (i.e. the real login work is handled by the subclasses). The signatures for these two login() method are identical, except for 1 parameter:

```
// SD 5.1 login
public void login(String server, String user, String password) throws QueryException ...

// SD 5.2 login
public void login(String server, String user,
                  String password, String sdConfigFileName) throws QueryException ...
```

I decide to make the login() signatures the same, by simply supplying QuerySD52 with the sdConfigFileName information via its constructor:

```
class QuerySD52 ...
    private String sdConfigFileName;
    public QuerySD52(String sdConfigFileName) {
        super();
        this.sdConfigFileName = sdConfigFileName;
    }
```

Now Query has only one abstract login() method:

```
abstract class Query ...
    public abstract void login(String server, String user,
                              String password) throws QueryException ...
```

And client code is updated as follows:

```
public void loginToDatabase(String db, String user, String password)...
    if (usingSDVersion52())
        query = new QuerySD52(getSD52ConfigFileName());
    else
        query = new QuerySD51();

    try {
        query.login(db, user, password);
        ...
    } catch(QueryException qe)...
```

I'm nearly done. Since Query is an abstract class, I decide to rename it AbstractQuery, which communicates more about its nature. But making that name change necessitates changing client code to declare variables of type AbstractQuery instead of Query. Since I don't want to do that, I apply *Extract Interface (341)* [F] on AbstractQuery to obtain a Query interface that AbstractQuery can implement:

```
interface Query {
    public void login(String server, String user, String password) throws QueryException;
    public void doQuery() throws QueryException;
}

abstract class AbstractQuery implements Query ...
public abstract void login(String server, String user,
    String password) throws QueryException ...
```

Now, subclasses of AbstractQuery implement login(), while AbstractQuery doesn't even need to declare the login() method, since it is an abstract class.

I compile and test and everything works as planned. Each version of SuperDatabase is now fully *adapted*. The code is smaller and treats each version in a more uniform way, all of which makes it easier to

- see similarities and differences between the versions
- remove support for older, unused versions
- add support for newer versions

Adapting with Anonymous Inner Classes

JDK 1.0 included an interface called Enumeration, which was used to iterate over collections like Vectors or Hashtables. Over time, better collections classes were added to the JDK, along with a new interface, called Iterator. To make it possible to interoperate with code written using the Enumeration interface, the JDK provided the following Creation Method, which uses Java's anonymous inner class capability to adapt an Iterator with an Enumeration:

```
public class Collections...
    public static Enumeration enumeration(final Collection c) {
        return new Enumeration() {
            Iterator i = c.iterator();

            public boolean hasMoreElements() {
                return i.hasNext();
            }
            public Object nextElement() {
                return i.next();
            }
        };
    }
}
```

Adapting Legacy Systems

An organization has an extremely sophisticated system which brings in most of their income, but which happens to be written in about 2 million lines of COBOL, little of which was ever refactored over a decade of development. Sound familiar? Systems like this are usually hard to extend because they were never refactored. And as a result, organizations that maintain such systems can't easily add new features to them, which makes them less competitiveness, which can ultimately put them out of business.

What to do? One popular approach is to use Adapters to model new views of the legacy system. Client code talks to the Adapters, which in turn talk to the legacy code. Over time, teams rewrite entire systems by simply writing new implementations for each Adapter. The process goes like this:

- Identify a subsystem of your legacy system
- Write Adapters for that subsystem
- Write new client programs that rely on calls to the Adapters
- Create versions of each Adapter using newer technologies
- Test that the newer and older Adapters function identically
- Update client code to use the new Adapters
- Repeat for the next subsystem

This is an example of applying *Separate Versions with Adapte (134)*, only it is performed across an entire system or subsystem, so the mechanics are a bit different.

Adapt Interface

Your class implements an interface but only provides code for some of the interface's methods.

Move the implemented methods to an Adapter of the interface and make the Adapter accessible from a Creation Method.

```
public class CardComponent extends Container implements MouseMotionListener ...
public CardComponent(Card card, Explanations explanations) {
    ...
    addMouseMotionListener(this);
}
public void mouseDragged(MouseEvent e) {
    e.consume();
    dragPos.x = e.getX();
    dragPos.y = e.getY();
    setLocation(getLocation().x+e.getX()-currPos.x,
                getLocation().y+e.getY()-currPos.y);
    repaint();
}
public void mouseMoved(MouseEvent e) {
}
```



```
public class CardComponent extends Container ...
public CardComponent(Card card, Explanations explanations) {
    ...
    addMouseMotionListener(createMouseMotionAdapter());
}
private MouseMotionAdapter createMouseMotionAdapter() {
    return new MouseMotionAdapter() {
        public void mouseDragged(MouseEvent e) {
            e.consume();
            dragPos.x = e.getX();
            dragPos.y = e.getY();
            setLocation(getLocation().x+e.getX()-currPos.x,
                        getLocation().y+e.getY()-currPos.y);
            repaint();
        }
    };
}
```

Motivation

Empty methods in concrete classes bother me. I often find that they're there because a class needs to satisfy a contract by implementing an interface, but only really needs code for *some* of the interface's methods. The rest of the methods get declared, but remain empty: they were added to satisfy a compiler rule. I find that these empty methods add to the heftiness of a class's interface (i.e. it's public methods), falsely advertise behavior (I'm a class that can, among other things, do X(), Y() and Z() – only I really only provide code for X()), and forces me to do work (like declaring empty methods) that I'd rather not do.

The Adapter pattern provides a nice way to refactor this kind of code. By implementing empty methods for every method defined by an interface, the Adapter lets me subclass it to supply just the code I need. In Java, I don't even have to formally declare an Adapter subclass: I

can just create an anonymous inner Adapter class and supply a reference to it from a Creation Method.

Communication	Duplication	Simplicity
Empty methods on a class don't communicate very much at all. Either someone forgot to delete the empty method, or it is just there because an interface forces you to have it there. It is far better to communicate only what you actually implement, and an Adapter can make this feasible.	If more than one of your classes partially implements an interface, you'll have numerous empty methods in your classes. You can remove this duplication by letting each of the classes work with an Adapter which handles the empty method declarations.	It is always simpler to supply less code than more. This refactoring gives you a way to cut down on the number of methods your classes declare. In addition, when used to adapt multiple interfaces, it can provide a nice way to partition methods in each of their respective adapters.

Mechanics

1. If you don't already have an adapter for the interface (which we'll call A), create a class that implements the interface and provides do-nothing behavior. Then write a Creation Method that will return a reference to an instance of your Adapter (which we'll call AdapterInstance).
2. Delete every empty method in your class that's solely there because your class implements A.
3. For those methods specified by A for which you have code, move each to your AdapterInstance.
4. Remove code declaring that your class implements A.
5. Supply the AdapterInstance to clients who need it.

Example

We'll use the example from the code sketch above. In this case we have a class called `CardComponent` that extends the `JDK Component` class and implements the `JDK's MouseMotionListener` interface. However, it only implements one of the two methods declared by the `MouseMotionListener` interface. So our task here is to replace a partially implemented interface with an Adapter.

1. The first step involves writing a Creation Method for our AdapterInstance. If we don't have an AdapterInstance, we need to create one using the refactoring, *Adapt Interface*. In this case, the `JDK` already supplies us with an adapter for the `MouseMotionListener` interface. It's called `MouseMotionAdapter`. So we create the following new method on the `CardComponent` class, using Java's handy anonymous inner class capability:

```
private MouseMotionAdapter createMouseMotionAdapter() {
    return new MouseMotionAdapter() {
    };
}
```

2. Next, we delete the empty method(s) that `CardComponent` declared because it implemented `MouseMotionListener`. In this case, it implemented `mouseDragged()`, but did not implement `mouseMoved()`.


```
public void mouseMoved(MouseEvent e) {}
```

3. We're now ready to move the `mouseDragged()` method from `CardComponent` to our instance of the `MouseMotionAdapter`:

```
private MouseMotionAdapter createMouseMotionAdapter() {  
    return new MouseMotionAdapter() {  
        public void mouseDragged(MouseEvent e) {  
            e.consume();  
            dragPos.x = e.getX();  
            dragPos.y = e.getY();  
            setLocation(getLocation().x+e.getX()-currPos.x,  
                        getLocation().y+e.getY()-currPos.y);  
            repaint();  
        }  
    };  
}
```

4. Now we can remove the `implements MouseMotionListener` from `CardComponent`.

```
public class CardComponent extends Container implements MouseMotionListener {
```

5. Finally, we must supply the new adapter instance to clients that need it. In this case, we must look at the constructor. It has code that looks like this:

```
public CardComponent() {  
    ...  
    addMouseMotionListener(this);  
}
```

This needs to be changed to call our new, private, Creation Method:

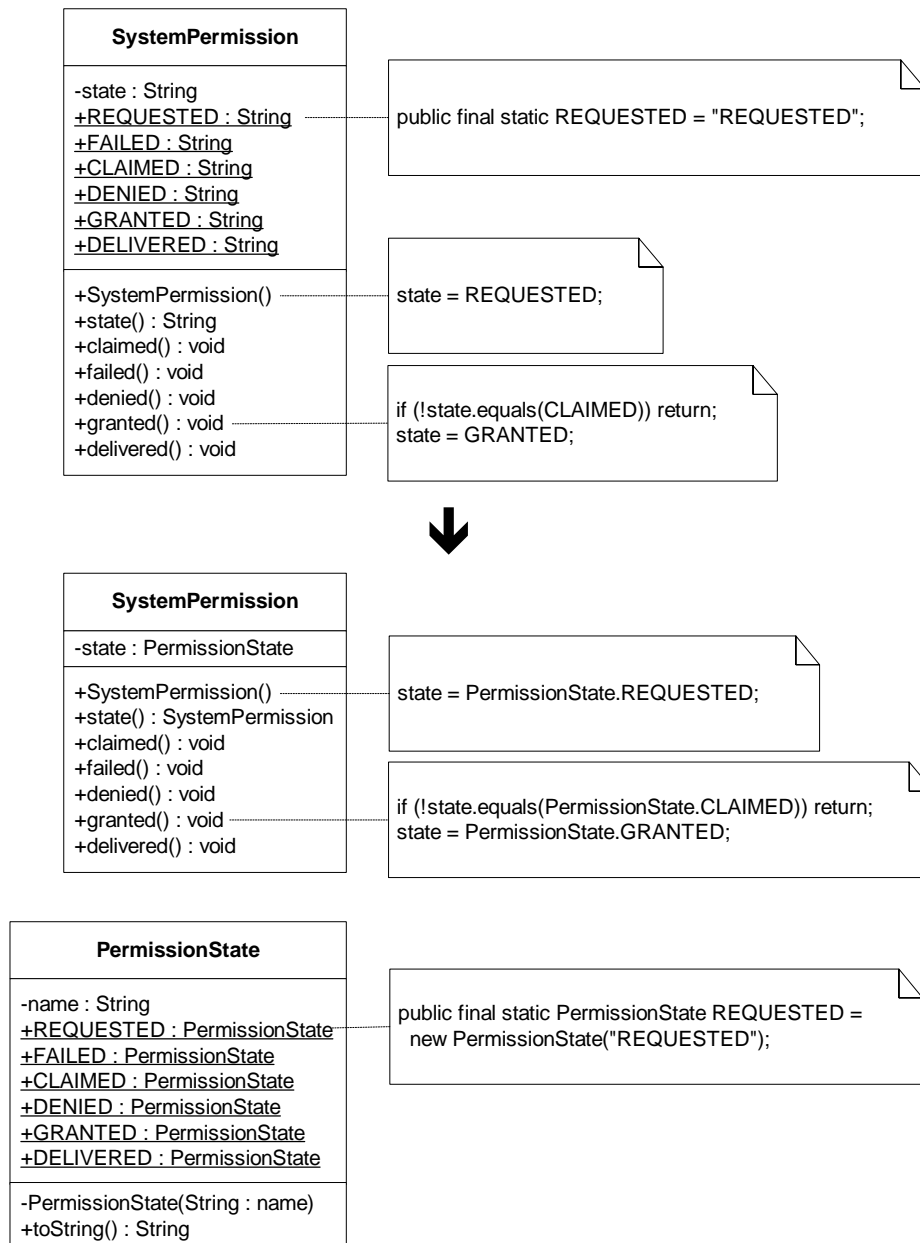
```
public CardComponent() {  
    ...  
    addMouseMotionListener(createMouseMotionAdapter());  
}
```

Now we test. Unfortunately, since this is mouse related code, I don't have automated unit tests. So I resort to some simple manual testing and confirm that everything is ok.

Replace Enum with Type-Safe Enum

A field's type (e.g. String, int, etc.) fails to protect it from unsafe assignments and invalid equality comparisons

Constrain the assignments and equality comparisons by making the field type-safe



Motivation

A Type-Safe Enum bundles together a user-defined type with a set of constant instances of that type. A primary motivation for refactoring to Type-Safe Enum is to constrain the possible values that may be assigned to or equated with a variable.

To understand the value of this pattern, it helps to study code that isn't type-safe. Consider the following test case:

```
public void testPermissionRequest() {
    SystemPermission permission = new SystemPermission();
    assertEquals("permission state", permission.REQUESTED, permission.state());
    assertEquals("permission state", "REQUESTED", permission.state());
}
```

The first line of code creates a `SystemPermission` object. The constructor for this object sets its state instance variable equal to the `SystemPermission.REQUESTED` state:

```
public SystemPermission() {
    state = REQUESTED;
}
```

Other methods within `SystemPermission` assign state to system permission states such as `GRANTED` and `DENIED`. Now, given that each of these state types was defined using `String` constants (like `public final static String REQUESTED = "REQUESTED"`), and state was defined as type `String`, then the two tests above would both evaluate to true since state - accessible via `permission.state()` - would be considered equal to `SystemPermission.REQUESTED` and the `String`, "REQUESTED."

What's the problem with that? Glad you asked. The `String` "REQUESTED" represents one object reference while the constant `String`, `SystemPermission.REQUESTED`, represents a different object reference, and yet the instance variable, `state`, is considered equal to both of them? That's not good, for just after a `SystemPermission` is instantiated, we want its `state` to be equal to the object reference, `SystemPermission.REQUESTED`, and no other object reference. A Type-Safe Enum can easily accomplish this.

Another motivation for refactoring to a Type-Safe Enum occurs when callers can change the value of an instance variable to an invalid value. For example, consider this code:

```
public class SystemPermission...
    public void setState(String newState){
        state = newState;
    }

    permission.setState("thinking"); // "thinking" is not a valid SystemPermission state
```

If one didn't use a Type-Safe Enum to prevent such spurious assignments, you'd have to fill your classes with lots of unnecessary validation logic.

Communication	Duplication	Simplicity
It is useful to communicate the availability of a type and constant values of that type. A Type-Safe Enum does this well because it is a class that exists solely to define the type and constants.	Duplication isn't an issue with respect to this refactoring.	A family of constants defined using a language-based type is slightly simpler to declare than a family of Type-Safe Enums, but because Type-Safe Enums prevent spurious assignments from occurring, they often help us simplify code.

Mechanics

(Please note: I'll be thoroughly changing these mechanics soon as I've found a better way to implement this refactoring Stay tuned -jk)

Identify a type-unsafe field – i.e. an instance variable declared as a language-defined type, which is assigned to or compared against a family of constant values.

Compile and test.

Declare a new class to store the family of constant values, naming the class after the kinds of types it will store. This will be your Type-Safe Enum.

In the class that declared the type-unsafe field, create a type-safe version of it by declaring a field whose type is the Type-Safe Enum class. Create any necessary getting/setting methods for this field, mirroring the getting/setting methods declared for the type-unsafe field.

Choose one constant value that the type-unsafe field is assigned to and/or compared against and define a new version of this constant in your Type-Safe Enum class by creating a public final static constant that is an instance of the Type-Safe Enum class.

Wherever the type-unsafe field is assigned to the constant value chosen for step 4, *add* code to assign the type-safe field equal to the type-safe enum constant created during the previous step.

Wherever the type-unsafe field is compared against the constant value chosen in step 4, *add* code to compare it against the type-safe enum constant created during step 4.

Whenever possible, add this comparison code after the logic that compares the type-unsafe field to the constant.

Compile and test.

Repeat steps 4 through 6 for every constant in the family of constant values.

Delete the type-unsafe field and anything related to it: getting/setting methods, direct assignments to it, comparisons against it and all of the type-unsafe constants.

Compile and test.

Example

This example, which was shown in the code sketch and mentioned in the Motivation section, deals with handling permission requests to access software systems. We'll begin by looking at relevant parts of the class, `SystemPermission`:

```
public class SystemPermission {  
    private String state;  
    private boolean granted;
```

```
private boolean failed;

public final static String REQUESTED = "REQUESTED";
public final static String FAILED = "FAILED";
public final static String CLAIMED = "CLAIMED";
public final static String DENIED = "DENIED";
public final static String GRANTED = "GRANTED";
public final static String DELIVERED = "DELIVERED";

public SystemPermission() {
    state = REQUESTED;
    failed = false;
    granted = false;
}

public boolean isGranted() {
    return granted;
}

public boolean hasFailed() {
    return failed;
}

public String state() {
    return state;
}

public void claimed() {
    if (state.equals(REQUESTED))
        state = CLAIMED;
}

public void failed() {
    if (!state.equals(REQUESTED)) return;
    state = FAILED;
    failed = true;
}

public void denied() {
    if (state.equals(CLAIMED))
        state = DENIED;
}

public void granted() {
    if (!state.equals(CLAIMED)) return;
    state = GRANTED;
    granted = true;
}

public void delivered() {
    if (state.equals(GRANTED) || state.equals(DENIED) )
        state = DELIVERED;
}
}
```

1. The field we're interested in here is called `state`, since it can be assigned to or compared against a family of `String` constants also defined inside `SystemPermission`. Our goal is to make `state` type-safe.

The first step is to rename `state` and its associated getting/setting methods. I'll rename it to `old_state`, and, since `state` only has a getting method and no setting method, I'll create a method called `old_state()` and update client code to use it:

```
public class SystemPermission...
    private String old_state;

    public SystemPermission() {
        old_state = REQUESTED;
    }
}
```

```
    ...  
}  
  
public String old_state() {  
    return old_state;  
}  
  
// etc.
```

Here is some client code I also update:

```
public class SystemPermissionTest extends TestCase...  
    public void testPermissionRequest() {  
        assertEquals("request", SystemPermission.REQUESTED, permission.old_state());  
    }  
}
```

Note: It's best to use an automated refactoring tool to handle the renaming of the variable and method(s). Now I compile and test to make sure the name changes didn't break anything.

2. Now I create a class called `PermissionState`, which will be my Type-Safe Enum class:

```
public final class PermissionState {  
}
```

I make it final because it will not need to be subclassed.

3. I create a new type-safe field inside `SystemPermission`, using the type, `PermissionState`. Since `old_state` only had a getting method and not a setting method, I only need to create a getting method for `state`:

```
public class SystemPermission...  
    private PermissionState state;  
  
    public PermissionState state() {  
        return state;  
    }  
}
```

4. Now I choose one constant value that the type-unsafe instance variable is assigned to or compared against, and I create a version of this constant in `PermissionState`, making it a public constant `PermissionState` member variable and instance of `PermissionState`:

```
public final class PermissionState {  
    public final static PermissionState REQUESTED = new PermissionState();  
}
```

This new type-safe constant will be easier to work with if I can query its `toString()` method to see which `PermissionState` type it is. So I make the following change:

```
public final class PermissionState {  
    private String name;  
  
    private PermissionState(String name) {  
        this.name = name;  
    }  
  
    public final static PermissionState REQUESTED = new PermissionState("REQUESTED");  
  
    public String toString() {  
        return name;  
    }  
}
```

5. Wherever I find code that assigns `old_state` to `SystemPermission.REQUESTED`, I must *add* code to assign state to `PermissionState.REQUESTED`:

```
public class SystemPermission...
    public SystemPermission() {
        old_state = REQUESTED;
        state = PermissionState.REQUESTED;
        failed = false;
        granted = false;
    }
}
```

Note: I'll delete the `old_state` assignment code later, when doing so won't cause logic problems with code that expects it to have a certain value.

6. Wherever `old_state` is compared against `SystemPermission.REQUESTED`, I must *add* code to also compare state against `PermissionState.REQUESTED`:

Here is some test code that needs updating:

```
public class SystemPermissionTest extends TestCase...
    private SystemPermission permission;

    public void setUp() {
        permission = new SystemPermission();
    }
    public void testPermissionRequest() {
        assertEquals("request", SystemPermission.REQUESTED, permission.old_state());
    }
}
```

The `testPermissionRequest` method becomes:

```
public void testPermissionRequest() {
    assertEquals("request", SystemPermission.REQUESTED, permission.old_state());
    assertEquals("request", PermissionState.REQUESTED, permission.state());
}
```

The following code in `SystemPermission` also needs updating:

```
public class SystemPermission...
    public void claimed() {
        if (old_state.equals(REQUESTED))
            old_state = CLAIMED;
    }

    public void failed() {
        if (!old_state.equals(REQUESTED)) return;
        old_state = FAILED;
        failed = true;
    }
}
```

I change this to:

```
public class SystemPermission...
    public void claimed() {
        if (old_state.equals(REQUESTED) && state.equals(PermissionState.REQUESTED))
            old_state = CLAIMED;
    }

    public void failed() {
        if (!old_state.equals(REQUESTED)) return;
        if (!state.equals(PermissionState.REQUESTED)) return;
        old_state = FAILED;
        failed = true;
    }
}
```

```
        old_state = FAILED;
        failed = true;
    }
```

I compile and test to see that everything is still working smoothly.

7. Next, I repeat steps 4 through 6 for every constant in the family of constant values. I'll spare you the details.

8. Now I have the pleasure of deleting `old_state`, the getting method, `old_state()`, all assignments made to `old_state`, any comparisons made to `old_state` and the entire family of `SystemPermission` type-unsafe constants. Here are a few of the deletions:

```
public class SystemPermissionTest extends TestCase...
    public void testPermissionRequest() {
        assertEquals("request", SystemPermission.REQUESTED, permission.old_state());
        assertEquals("request", PermissionState.REQUESTED, permission.state());
    }
    // etc...

public class SystemPermission...
    private String old_state;

    public final static String REQUESTED = "REQUESTED";
    public final static String FAILED = "FAILED";
    public final static String CLAIMED = "CLAIMED";
    public final static String DENIED = "DENIED";
    public final static String GRANTED = "GRANTED";
    public final static String DELIVERED = "DELIVERED";

    public SystemPermission() {
        old_state = REQUESTED;
        state = PermissionState.REQUESTED;
        ...
    }

    public String old_state() {
        return old_state;
    }

    public void claimed() {
        if (old_state.equals(REQUESTED) && state.equals(PermissionState.REQUESTED)) {
            old_state = CLAIMED;
            state = PermissionState.CLAIMED;
        }
    }

    // and so on...
```

I compile and test after all of the deletions. Now `state` is type-safe:

```
public class SystemPermission {
    private PermissionState state;
    private boolean granted;
    private boolean failed;

    public SystemPermission() {
        state = PermissionState.REQUESTED;
        failed = false;
        granted = false;
    }

    public boolean isGranted() {
        return granted;
    }
}
```



```
public boolean hasFailed() {
    return failed;
}

public PermissionState state() {
    return state;
}

public void claimed() {
    if (state.equals(PermissionState.REQUESTED))
        state = PermissionState.CLAIMED;
}

public void failed() {
    if (!state.equals(PermissionState.REQUESTED))
        return;
    state = PermissionState.FAILED;
    failed = true;
}

public void denied() {
    if (state.equals(PermissionState.CLAIMED))
        state = PermissionState.DENIED;
}

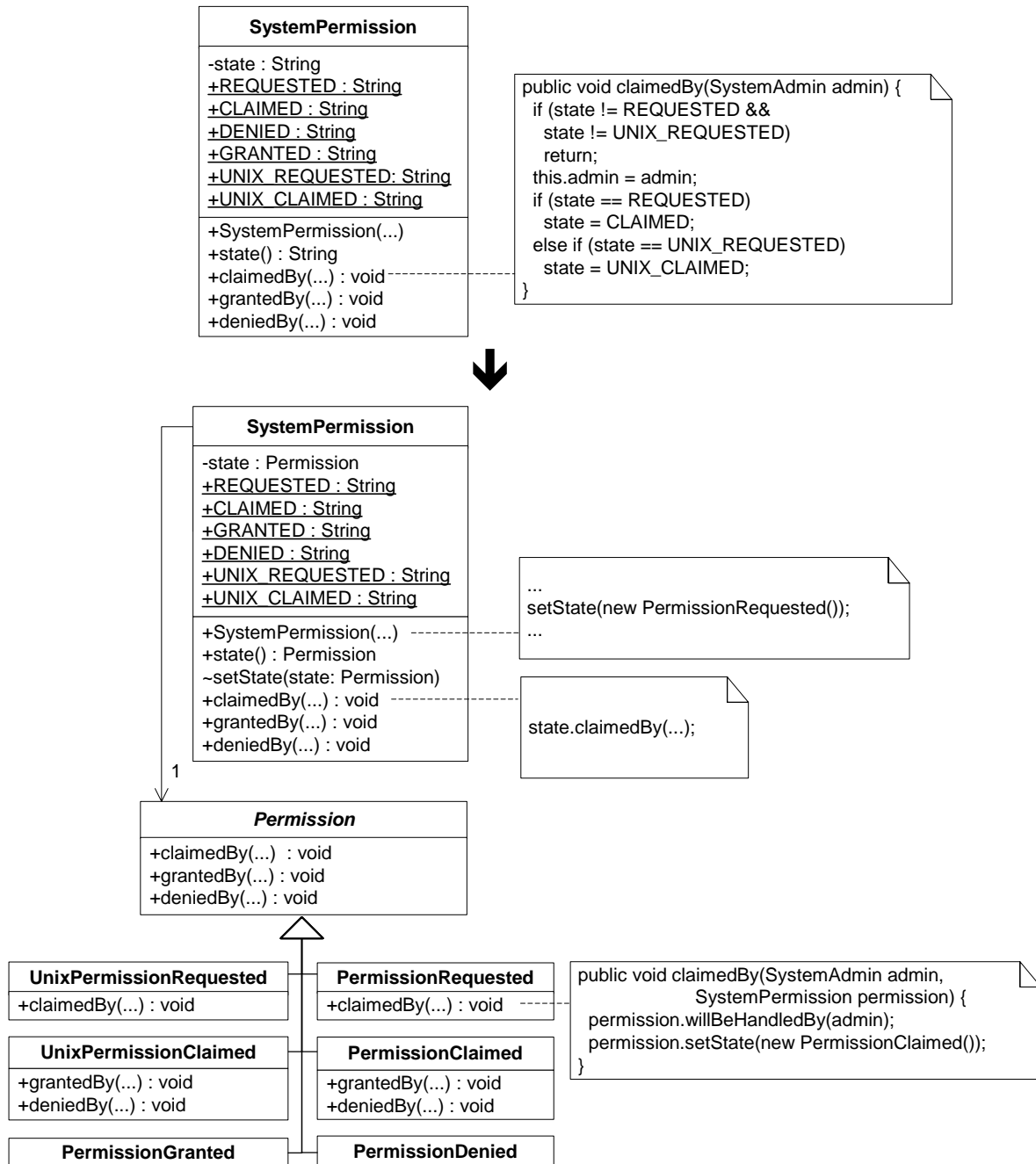
public void granted() {
    if (!state.equals(PermissionState.CLAIMED))
        return;
    state = PermissionState.GRANTED;
    granted = true;
}

public void delivered() {
    if (state.equals(PermissionState.GRANTED)
        || state.equals(PermissionState.DENIED))
        state = PermissionState.DELIVERED;
}
}
```

Replace State-Altering Conditionals with State

Complex conditional expressions control an object's state transitions

Replace the conditionals with State classes that handle specific states and transitions between them



Motivation

The primary reason for refactoring to the State pattern is to tame overly-complex state-altering conditional logic. Such logic, which tends to spread itself throughout a class, controls an object's state, including how states transition to other states. When you implement this pattern you create state classes that represent specific states of an object and the transitions between those states. The object that has its state changed is known as the *context*. A context delegates state-changing behavior to a state object. State objects make state transitions at runtime by making the context point to a different state object.

If you don't know the State pattern very well, you'll understand this refactoring better if you study it in *Design Patterns* [GoF]. If you know this pattern, perhaps your using it when you don't need to be: i.e. when simple state-altering conditional logic would suffice. This refactoring is concerned with helping you identify the place where state-altering conditional logic is no longer easy to follow or extend and when the State pattern can really make a difference.

Before I refactor to State, I always see if I can implement a simpler solution by applying low-level refactorings, such as *Extract Method* (110) [F]. If such refactorings still don't tame the conditional logic, I know I'm ready for State. The State pattern can remove or reduce many lines of conditional logic, yielding clean, simple and extensible code.

If your state objects have no instance variables (i.e. they are *stateless*), you can optimize memory usage by having context objects share instances of the stateless state instances. The patterns Flyweight [GoF] or Singleton [GoF] are often used to implement sharing (for example, see (169)). However, it's always best to add state-sharing code *after* your users experience system delays and a profiler points you to the state-instantiation code as a prime bottleneck.

This refactoring is different from *Replace Type Code with State/Strategy* (227) [F] in a few areas. First, I don't have a single refactoring for the State and Strategy patterns because I view them as different patterns, I refactor to them for different reasons (see

Replace Conditional Calculations with Strategy (50)) and the mechanics of the refactorings to each pattern differ. Second, Martin deliberately doesn't document a full refactoring to the State pattern, since the complete implementation depends on a further refactoring he wrote, called *Replace Conditional with Polymorphism* (225) [F]. While I respect that decision, I thought it would be more helpful to readers to understand how the refactoring works from end to end, so my mechanics and example sections delineate all of the steps to get you from conditional state-changing logic to a State pattern implementation.

Communication	Duplication	Simplicity
Many lines of state-altering conditional logic don't communicate intent very well. Communicate the state changing logic clearly by splitting out the states into well-named classes, each of which knows how to transition itself to other states.	When you have lots of state-altering conditional logic, you tend to see the same conditional phrases repeated throughout the methods of a class. Implementing the State pattern allows you to remove much of this conditional logic.	One of the main reasons to perform this refactoring is to simplify complex state-changing logic. If you can't easily follow the state-changing logic in your class, it may be a good time to refactor to State.

Mechanics

1. A class (which we'll call the *context class*) contains a field (which we'll call the *original state field*) that gets assigned to or compared against a family of constants during state transitions. Rename this field and self-encapsulate it (*Self-Encapsulate Field* [F]).

✓ Compile and test.

2. Declare a new abstract class and name it based on the name or general purpose of the *original state field*. This will be the *state superclass*.

Declare this class to be non-public (e.g. in Java, consider using package-protection).

3. Declare subclasses of the *state superclass*, one for each of the states the *context class* may enter.

Declare these classes to be non-public (e.g. in Java, consider using package-protection).

4. Create a non-public field (which we'll call the *state field*) in the *context class*, making its type that of the *state superclass*. Create a setting method for it.

5. Wherever an assignment is made to the *original state field*, add a similar assignment statement to the *state field*, using a new instance of the appropriate subclass for the assignment value.

6. Change the getting method for the *original state field* so that its return value is obtained from the *state field*. This will require making each of the state subclasses capable of returning the correct value.

✓ Compile and test.

7. In the *context class*, for each constant in the family of constants that identify states, replace its value with a constant value obtained from the corresponding state subclass.

✓ Compile and test.

8. Find a *context class* method that changes the value of the *original state field* based on state-transition logic. Copy this method to the *state superclass*, making the simplest changes possible to make the new method work. (A common, *simple* change is to pass the *context class* to the method in order to have code call methods on the *context class*). Finally, replace the body of the *context class* method with a delegation call to the new method.

✓ Compile and test.

Repeat this step for every *context class* method that changes the value of the *original state field* based on state-transition logic.

9. Choose a state that the *context class* can enter and identify which *state superclass* methods make this state transition to other states. Copy the identified method(s), if any, to the subclass associated with the chosen state and remove all unrelated logic.

Unrelated logic usually includes verifications of a current state or logic that transitions to unrelated states.

✓ Compile and test.

Repeat for all states the *context class* can enter.

10. Delete the bodies of each of the methods copied to the *state superclass* during step 8 to produce an empty implementation for each method.

✓ Compile and test.

11. In the *context class*, delete the *original state field*, the setting method for it and all calls to the setting method.

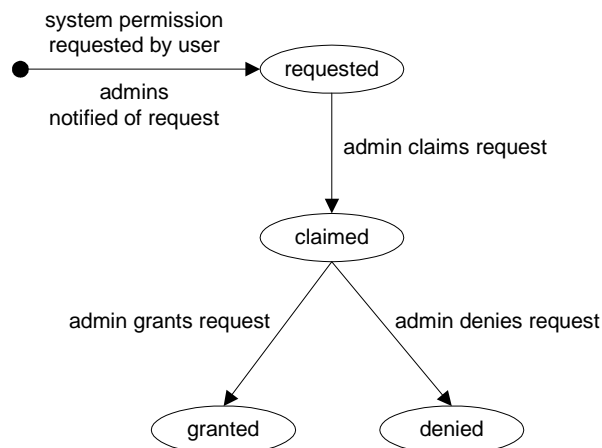
✓ Compile and test.

12. Rename the *original state field's* getting method (perhaps to its original name before step 1) and update all callers.

✓ Compile and test.

Example

To understand when it makes sense to refactor to the State pattern, it helps to study a class that manages its state without requiring the sophistication of the State pattern. `SystemPermission` is such a class. It uses simple conditional logic to keep track of the state of a permission request to access a software system. Over the lifetime of a `SystemPermission` object an instance variable named `state` transitions between the states *requested*, *claimed*, *denied* and *granted*. Here is a UML representation of the possible transitions:



Below is the code for `SystemPermission` and a fragment of test code to show how the class gets used:

```
public class SystemPermission {
    private SystemProfile profile;
    private SystemUser requestor;
    private SystemAdmin admin;
    private boolean isGranted;
    private String state;

    public final static String REQUESTED = "REQUESTED";
    public final static String CLAIMED = "CLAIMED";
    public final static String GRANTED = "GRANTED";
    public final static String DENIED = "DENIED";

    public SystemPermission(SystemUser requestor, SystemProfile profile) {
        this.requestor = requestor;
        this.profile = profile;
        state = REQUESTED;
        isGranted = false;
        notifyAdminOfPermissionRequest();
    }

    public String state() {
        return state;
    }

    public void claimedBy(SystemAdmin admin) {
        if (state != REQUESTED)
            return;
        this.admin = admin;
        state = CLAIMED;
    }

    public void deniedBy(SystemAdmin admin) {
        if (state != CLAIMED)
            return;
        if (this.admin != admin) return;
        isGranted = false;
        state = DENIED;
        notifyUserOfPermissionRequestResult();
    }

    public void grantedBy(SystemAdmin admin) {
        if (state != CLAIMED)
            return;
        if (this.admin != admin) return;
        state = GRANTED;
        isGranted = true;
        notifyUserOfPermissionRequestResult();
    }

    public boolean isGranted() {
        return isGranted;
    }

    public void notifyAdminOfPermissionRequest() {
        // ...
    }
    public void notifyUserOfPermissionRequestResult() {
        // ...
    }
}

public class TestStates extends TestCase ...
    public void testGrantedBy() {
        permission.grantedBy(admin);
    }
}
```

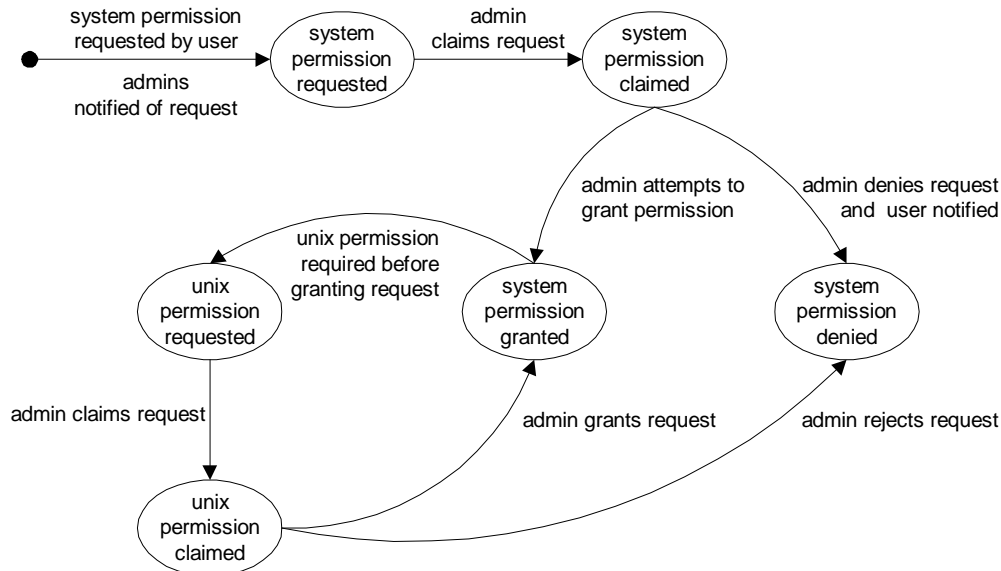
```

assertEquals("requested", permission.REQUESTED, permission.state());
assertEquals("not granted", false, permission.isGranted());
permission.claimedBy(admin);
permission.grantedBy(admin);
assertEquals("granted", permission.GRANTED, permission.state());
assertEquals("granted", true, permission.isGranted());
}

```

Notice how the instance variable, `state`, gets assigned to different values as clients call specific `SystemPermission` methods. Now look at the overall conditional logic in `SystemPermission`. This logic is responsible for transitioning between states, but the logic isn't very complicated so the code doesn't require the sophistication of the State pattern.

This conditional state changing logic can quickly become hard to follow as more real-world behavior gets added to the `SystemPermission` class. For example, I helped design a security system in which users needed to obtain unix and/or database permissions before the user could be granted general permission to access a given software system. The state transition logic that requires unix permission before general permission may be granted looks like this:



Adding support for unix permission makes `SystemPermission`'s state-altering conditional logic a lot more complicated than it used to be. Consider the following:

```

public class SystemPermission...
    public void claimedBy(SystemAdmin admin) {
        if (state != REQUESTED &&
            state != UNIX_REQUESTED)
            return;
        this.admin = admin;
        if (state == REQUESTED)
            state = CLAIMED;
        else if (state == UNIX_REQUESTED)
            state = UNIX_CLAIMED;
    }

    public void deniedBy(SystemAdmin admin) {
        if (state != CLAIMED &&
            state != UNIX_CLAIMED) return;
    }

```



```

    if (this.admin != admin) return;
    isGranted = false;
    isUnixPermissionGranted = false;
    state = DENIED;
    notifyUserOfPermissionRequestResult();
}

public void grantedBy(SystemAdmin admin) {
    if (state != CLAIMED &&
        state != UNIX_CLAIMED) return;
    if (this.admin != admin) return;

    if (profile.isUnixPermissionRequired() &&
        state == UNIX_CLAIMED)
        isUnixPermissionGranted = true;
    else if (profile.isUnixPermissionRequired() &&
        !isUnixPermissionGranted()) {
        state = UNIX_REQUESTED;
        notifyUnixAdminsOfPermissionRequest();
        return;
    }
    state = GRANTED;
    isGranted = true;
    notifyUserOfPermissionRequestResult();
}

```

An attempt can be made to simplify the above code by applying *Extract Method (110)* [F]. For example, one could refactor the `grantedBy()` method like so:

```

public void grantedBy(SystemAdmin admin) {
    if (!isInClaimedState()) return;
    if (this.admin != admin) return;
    if (isUnixPermissionRequestedAndClaimed())
        isUnixPermissionGranted = true;
    else if (isUnixPermissionDesiredButNotRequested()) {
        state = PermissionState.UNIX_REQUESTED;
        notifyUnixAdminsOfPermissionRequest();
        return;
    }
    ...
}

```

That's a little better but now the `SystemPermission` class has lots of state-specific boolean logic (i.e. methods like `isUnixPermissionRequestedAndClaimed()`) and yet `grantedBy()` still isn't simple. It's time to simplify things by refactoring to the State pattern.

1. `SystemPermission` has a field called `state`, numerous statements that assign a constant value to `state` and an accessor method called `state()`. I rename `state` to `oldState` and self-encapsulate it:

```

public class SystemPermission...
    private String oldState;

    public SystemPermission(SystemUser requestor, SystemProfile profile) {
        ...
        setOldState(REQUESTED);
        ...
    }

    private void setOldState(String oldState) {
        this.oldState = oldState;
    }

    public String oldState() {
        return oldState;
    }
}

```

```
public void claimedBy(SystemAdmin admin) {
    if (oldState() != REQUESTED && oldState() != UNIX_REQUESTED)
        return;
    this.admin = admin;
    if (oldState() == REQUESTED)
        setOldState(CLAIMED);
    else if (oldState() == UNIX_REQUESTED)
        setOldState(UNIX_CLAIMED);
}

// etc...
```

I make sure all client and test code is updated, compile and test that everything still works.

2. Next, I create a *state superclass* for all of the permission states that a `SystemPermission` can enter. “Permission” sounds like a good name for this class:

```
abstract class Permission {}
```

I declare this class in the same package as `SystemPermission` and I give it Java’s package-protection (instead of making it public) since only `SystemPermission` will need to know about it

3. Now it’s time to create `Permission` subclasses for each of the states that a `SystemPermission` can enter. Again, I declare these classes in the same package as `SystemPermission` and I give them Java’s package-protection access:

```
class PermissionRequested extends Permission {}
class PermissionClaimed extends Permission {}
class PermissionGranted extends Permission {}
class PermissionDenied extends Permission {}
class UnixPermissionRequested extends Permission {}
class UnixPermissionClaimed extends Permission {}
```

4. Now I create a private `Permission` field in `SystemPermission` and provide a setting method for it:

```
public class SystemPermission...
    private Permission state;

    private void setState(Permission state) {
        this.state = state;
    }
}
```

5. Wherever an assignment is made to `oldState`, I make a similar assignment to `state`, using new instances of the recently-created subclasses as assignment values:

```
public class SystemPermission {
    public SystemPermission(SystemUser requestor, SystemProfile profile) {
        ...
        setOldState(REQUESTED);
        setState(new PermissionRequested());
        ...
    }

    public void claimedBy(SystemAdmin admin) {
        if (oldState() != REQUESTED && oldState() != UNIX_REQUESTED)
            return;
        this.admin = admin;
        if (oldState() == REQUESTED) {
            setOldState(CLAIMED);
        }
    }
}
```

```
        setState(new PermissionClaimed());
    }
    else if (oldState() == UNIX_REQUESTED) {
        setOldState(UNIX_CLAIMED);
        setState(new UnixPermissionClaimed());
    }
}

// etc...
```

You might wonder why I wrote the above code to directly instantiate state subclasses, such as `PermissionClaimed`, each time a method like `claimedBy(...)` is called. The answer is that I'm not concerned with memory optimization during this refactoring. If system users complain that their system is slow and a profiler can show me that significant delays result from repeated instantiations of state subclasses, I'll refactor the above code to obtain pre-instantiated state subclasses, perhaps using a refactoring like *Replace Multiple Instantiations with Flyweight* (193). Until then I won't optimize prematurely.

6. Now I refactor the `PermissionState` accessor method, `oldState()`, to return the correct value using the `state` field instead of the `oldState` field. This requires adding code to the *state superclass*, `Permission`, and its subclasses:

```
public class SystemPermission...
    public String oldState() {
        return state.name();
    }

abstract class Permission {
    public abstract String name();
}

class PermissionRequested extends Permission {
    public String name() {
        return "REQUESTED";
    }
}

class PermissionClaimed extends Permission {
    public String name() {
        return "CLAIMED";
    }
}

// etc...
```

I compile and test to confirm that the changes worked.

7. Now I update each of the state-identifying constants in `SystemPermission` to obtain their values from corresponding state subclasses. For example, here's the original code for the `REQUESTED` constant:

```
public class SystemPermission...
    public final static String REQUESTED = "REQUESTED";
```

I change this to:

```
public class SystemPermission...
    public final static String REQUESTED = PermissionRequested.NAME;

class PermissionRequested extends Permission...
    public final static String NAME = "REQUESTED";
```

```

public String name() {
    return NAME;
}

```

You might wonder why I just don't move all of `SystemPermission`'s state-identifying constants to each of the state subclasses. I considered doing just that, but then re-thought the decision because it would mean coupling client code (for example, test code) to the State pattern implementation. By leaving the constants in `SystemPermission`, my client code works regardless of whether I've implemented the State pattern.

8. Now I find a method on `SystemPermission` that changes the value of `oldState` based on state-transition logic. There are three such methods in `SystemPermission`: `claimedBy(...)`, `deniedBy(...)` and `grantedBy(...)`. I start by working with `claimedBy(...)`. I must copy this method to `Permission`, making enough changes to get it to compile and then replacing the body of the original `claimedBy(...)` method with a call to the new `Permission` version of it:

```

public class SystemPermission...
    void setOldState(String old_state) { // changed from private to package-protected
        this.oldState = old_state;
    }

    void setState(Permission state) {      // changed from private to package-protected
        this.state = state;
    }

    public void claimedBy(SystemAdmin admin) {
        state.claimedBy(admin, this);
    }

    public void willBeHandledBy(SystemAdmin admin) {
        this.admin = admin;
    }

abstract class Permission {
    public abstract String name();

    public void claimedBy(SystemAdmin admin, SystemPermission permission) {
        if (permission.oldState() != permission.REQUESTED &&
            permission.oldState() != permission.UNIX_REQUESTED)
            return;
        permission.willBeHandledBy(admin);
        if (permission.oldState() == permission.REQUESTED) {
            permission.setOldState(permission.CLAIMED);
            permission.setState(new PermissionClaimed());
        }
        else if (permission.oldState() == permission.UNIX_REQUESTED) {
            permission.setOldState(permission.UNIX_CLAIMED);
            permission.setState(new UnixPermissionClaimed());
        }
    }
}

```

After I compile and test to see that the changes worked, I repeat this step for `deniedBy(...)` and `grantedBy(...)`.

9. Now I choose a state that `PermissionRequest` can enter and identify which `Permission` methods make this state transition to other states. I'll start with the `REQUESTED` state. This state can only transition to the `CLAIMED` state and the transition happens in the

Permission.claimedBy(...) method. So I copy that method to the PermissionRequested class:

```
class PermissionRequested extends Permission...
public void claimedBy(SystemAdmin admin, SystemPermission permission) {
    if (permission.oldState() != permission.REQUESTED &&
        permission.oldState() != permission.UNIX_REQUESTED)
        return;
    permission.willBeHandledBy(admin);
    if (permission.oldState() == permission.REQUESTED) {
        permission.setOldState(permission.CLAIMED);
        permission.setState(new PermissionClaimed());
    }
    else if (permission.oldState() == permission.UNIX_REQUESTED) {
        permission.setOldState(permission.UNIX_CLAIMED);
        permission.setState(new UnixPermissionClaimed());
    }
}
}
```

There's a lot of logic in this method that's no longer needed. For example, anything related to the UNIX_REQUESTED state isn't needed because we're only concerned with the REQUESTED state in the PermissionRequested class. We also don't need to check if our current state is REQUESTED, because the fact that we're in the PermissionRequested class tells us that. So I can reduce this code to the following:

```
class PermissionRequested extends Permission...
public void claimedBy(SystemAdmin admin, SystemPermission permission) {
    permission.willBeHandledBy(admin);
    permission.setOldState(permission.CLAIMED);
    permission.setState(new PermissionClaimed());
}
}
```

As always, I compile and test to make sure I didn't break anything. Now I repeat this step for the 5 additional states. Let's look at what is required to produce the PermissionClaimed and PermissionGranted states.

The CLAIMED state can transition to DENIED, GRANTED or UNIX REQUESTED. The grantedBy(...) or deniedBy(...) methods take care of these transitions so I copy those methods to the PermissionClaimed class and delete unnecessary logic:

```
class PermissionClaimed extends Permission...
public void deniedBy(SystemAdmin admin, SystemPermission permission) {
    if (permission.oldState() != permission.CLAIMED &&
        permission.oldState() != permission.UNIX_CLAIMED)
        return;
    if (permission.admin() != admin)
        return;
    permission.setIsGranted(false);
    permission.setIsUnixPermissionGranted(false);
    permission.setOldState(permission.DENIED);
    permission.setState(new PermissionDenied());
    permission.notifyUserOfPermissionRequestResult();
}

public void grantedBy(SystemAdmin admin, SystemPermission permission) {
    if (permission.oldState() != permission.CLAIMED &&
        permission.oldState() != permission.UNIX_CLAIMED)
        return;
    if (permission.admin() != admin)
        return;

    if (permission.profile().isUnixPermissionRequired() &&
```

```

        permission.setOldState(permission.UNIX_CLAIMED);
        permission.setIsUnixPermissionGranted(true);
    } else if (
        permission.profile().isUnixPermissionRequired() &&
        !permission.isUnixPermissionGranted() {
        permission.setOldState(permission.UNIX_REQUESTED);
        permission.setState(new UnixPermissionRequested());
        permission.notifyUnixAdminsOfPermissionRequest();
        return;
    }
    permission.setOldState(permission.GRANTED);
    permission.setState(new PermissionGranted());
    permission.setIsGranted(true);
    permission.notifyUserOfPermissionRequestResult();
}
}
}

```

For `PermissionGranted`, my job is easy. Once a `SystemPermission` reaches the `GRANTED` state, it has no further states it can transition to (i.e. it's at an end state). So this class doesn't need to implement any transition methods (`claimedBy(...)`, etc). In fact, it really needs to inherit empty implementations of the transition methods, which is exactly what will happen after the next step in the refactoring.

10. On the `Permission` class, I can now delete the bodies of `claimedBy(...)`, `deniedBy(...)` and `grantedBy(...)`, leaving the following:

```

abstract class Permission {
    public abstract String name();
    public void claimedBy(SystemAdmin admin, SystemPermission permission) {}
    public void deniedBy(SystemAdmin admin, SystemPermission permission) {}
    public void grantedBy(SystemAdmin admin, SystemPermission permission) {}
}

```

I compile and test to confirm that the states continue to behave correctly.

11. Next, I delete `SystemPermission`'s `oldState` field, the setting method for it and any code that was calling the setting method

```

public class SystemPermission...
    private String oldState;

    public SystemPermission(SystemUser requestor, SystemProfile profile)...
        ...
        setOldState(REQUESTED);
        ...
    }

    void setOldState(String old_state){
        this.oldState = old_state;
    }

class PermissionRequested extends Permission...
    public void claimedBy(SystemAdmin admin, SystemPermission permission) {
        permission.willBeHandledBy(admin);
        permission.setOldState(permission.CLAIMED);
        permission.setState(new PermissionClaimed());
    }
}

// etc...

```

I compile and test.

12. For the final step, I rename the `SystemPermission` getting method, `oldState()`, to its original name, `state()`:

```
public class SystemPermission...
    public String state() {
        return state.name();
    }
}
```

I update all callers, such as this one:

```
public class TestStates extends TestCase...
    public void testRequestedBy() {
        assertEquals("requested", permission.REQUESTED, permission.state());
    }
}
```

And that completes the refactoring.

Inner States

In languages that let you define classes within classes, you can refactor the above example so that the *state superclass* and its subclasses are completely encapsulated in the *context class*. Here's what `Permission` and its subclasses look like when they are defined as inner classes of `SystemPermission`:

```
public class SystemPermission...
    private abstract class Permission {
        public abstract String name();
        public void claimedBy(SystemAdmin admin) {}
        public void deniedBy(SystemAdmin admin) {}
        public void grantedBy(SystemAdmin admin) {}
    }

    private Permission state;

    private class PermissionRequested extends Permission {
        public final static String NAME = "REQUESTED";

        public String name() {
            return NAME;
        }

        public void claimedBy(SystemAdmin admin) {
            willBeHandledBy(admin);
            setState(new PermissionClaimed());
        }
    }

    private class PermissionGranted extends Permission {
        public final static String NAME = "GRANTED";
        public String name() {
            return NAME;
        }
    }

    // etc...
```

The *state superclass*, `Permission`, and its subclasses can now be declared as private, since client code (inside or outside `SystemPermission`'s package) doesn't need to know that those classes exist. In addition, since the state subclasses reside inside `SystemPermission`, it's no longer necessary to pass a `SystemPermission` reference to each of their state-transitioning

methods, `claimedBy(...)`, etc. So the `SystemPermission` methods that call these state-transitioning methods now look like this:

```
public class SystemPermission...
    public void claimedBy(SystemAdmin admin) {
        state.claimedBy(admin); // used to be state.claimedBy(admin, this);
    }

    public void deniedBy(SystemAdmin admin) {
        state.deniedBy(admin); // used to be state.deniedBy(admin, this);
    }

    public void grantedBy(SystemAdmin admin) {
        state.grantedBy(admin); // used to be state.grantedBy(admin, this);
    }
}
```

This solution also makes it possible to disable any client from calling `SystemPermission`'s setting method, `setState(...)`, since it can now be declared as private:

```
public class SystemPermission...
    private void setState(Permission state) {
        this.state = state;
    }
}
```

Using inner classes in this way adds a lot of code to the *context class* but completely encapsulates the state classes. If you want to give your state subclasses complete access to context class methods without having to change the protection levels of those methods, inner classes are a good idea. If you need to add new state classes and state transitions without having to touch (or recompile) your *context class*, avoid using inner classes for your states.

Replace Invariable Behavior with Template Method

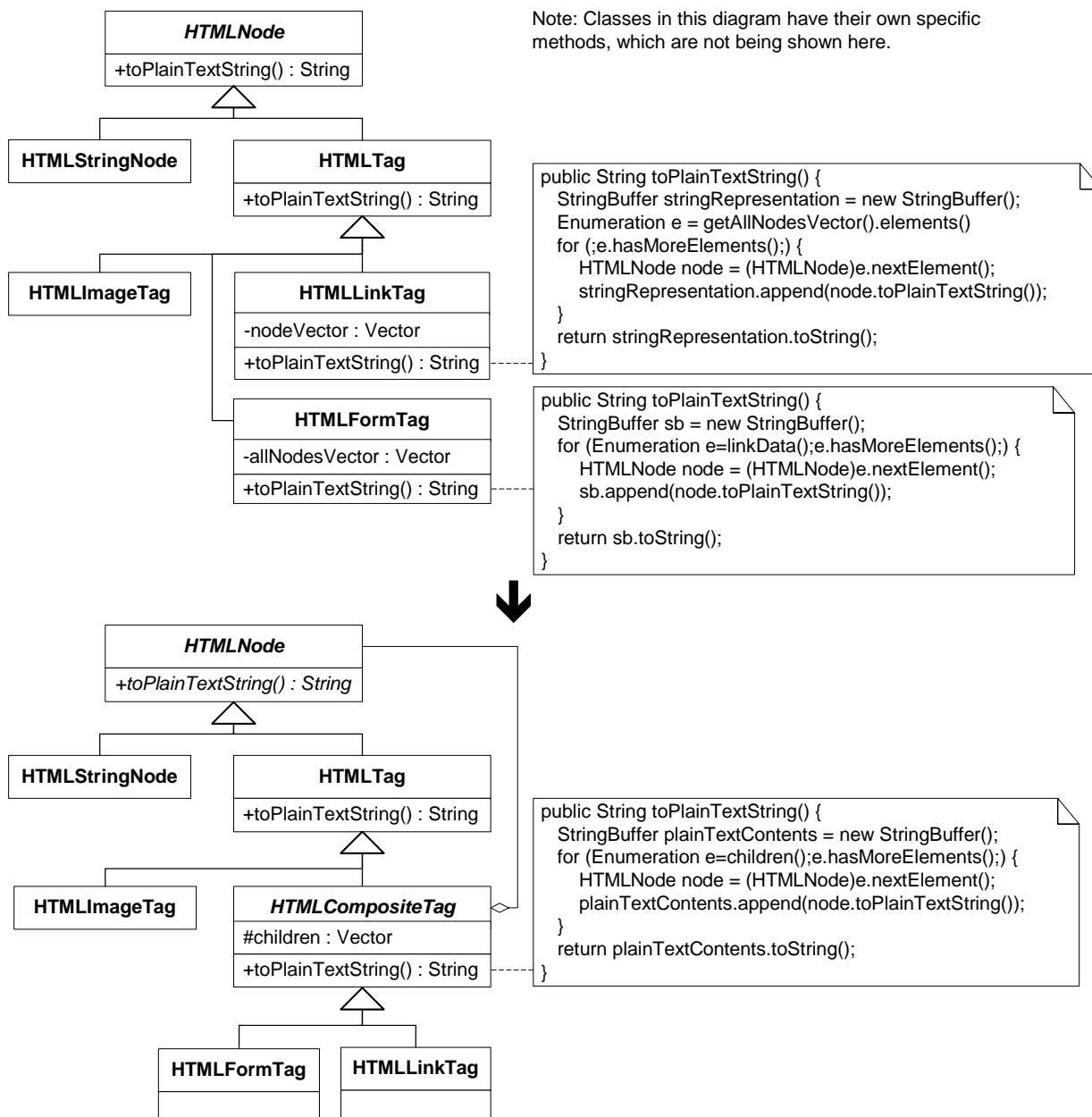
Subclasses implement a behavior in the same order but in variable ways

*Move the invariable parts of the behavior to a Template Method
and let the subclasses implement what varies*

Extract Composite

Subclasses in a hierarchy store children from the hierarchy and have duplicated methods that operate on the children

Create a Composite superclass and move the duplicated fields and logic from the subclasses to the Composite



Motivation

In *Extract Superclass* [F], Martin Fowler explains that if you have two or more classes with similar features, it makes sense to move the common features to a superclass. This refactoring is similar only it's concerned with when the common features duplicated across subclasses ought to be extracted to a *Composite* [GoF].

Subclasses in hierarchies that store collections of children and have methods for reporting information about those children are common. When the children being collected happen to be classes in the same hierarchy, there's a good chance that much duplicate code could be removed by refactoring to *Composite*.

Duplication removal is at the heart of this refactoring and *Extract Superclass* [F]. When you have the type of child-related code described above, following the mechanics for either refactoring will lead you to the creation of a *Composite*. So why did I choose to write this special case refactoring? Mostly because I think it's useful to give folks examples of duplication that are more subtle than others and to demonstrate different ways to remove the duplication. In addition, it would be excellent if the narrower mission of this refactoring (i.e., the removal of duplicate code related only to collections of children and related methods) made it easier for tools vendors to automate this refactoring.

Communication	Duplication	Simplicity
To-do	To-do	To-do

Mechanics

Since this refactoring is a generalized case of Martin Fowler's refactoring, *Extract Superclass* [F], it's mechanics are similar. The main difference is the extent to which you'll extract code to a superclass. For this refactoring, the extracted code is the child-handling logic that is similar across classes in a hierarchy. Once you finish this refactoring, you can continue to pull common functionality up into your newly created Composite by fully applying *Extract Superclass* [F].

1. Create a *composite*, an abstract superclass, named to reflect that it will contain children (e.g. `CompositeTag`).
2. Make each *child-container* (i.e. a class in the hierarchy that contains duplicate child-handling code) a subclass of *composite*.
3. In a *child-container*, find a child-processing method that is *purely-duplicated* or *partially-duplicated* across the *child-containers*. A *purely-duplicated* method will have the same method body with the same or different method names across *child-containers*. A *partially-duplicated* method will have a method body with common and uncommon code and the same or different method names across *child-containers*.

Whether you've found a *purely-duplicated* or *partially-duplicated* method, if its name isn't consistent across *child-containers*, make it consistent by applying *Rename Method* [F].

For a *purely-duplicated* method, move the child collection field referenced by the method to the *composite* by applying *Pull Up Field* [F]. Rename this field if its name doesn't make sense for all *child-containers*. Now move the method to the *composite* by applying *Pull Up Method* [F]. If the pulled-up method relies on constructor code still residing in *child-containers*, pull up that code to the *composite*'s constructor.

For a *partially-duplicated* method, see if the method body can be made consistent across all *child-containers* using *Substitute Algorithm* [F]. If so, refactor it as a *purely-duplicated* method. Otherwise, extract the code that is common across all *child-container* implementations using *Extract Method* [F] and pull it up to the composite using *Pull Up Method* [F]. If the method body follows the same sequence of steps, some of which are implemented differently, see if you can apply *Replace Invariable Behavior with Template Method* (169).

✓ Compile and test after each refactoring.

4. Repeat step 3 for child-processing methods in the child-containers that contain *purely-duplicated* or *partially-duplicated* code.
5. Check each client of each *child-container* to see if it can now communicate with the *child-container* using the composite interface. If it can, make it do so.

Example

I've been refactoring some code on an open-source project that was started by my colleague, Somik Raha. The project is an HTML parser (see <http://sourceforge.net/projects/htmlparser>). When the parser parses a piece of HTML, it identifies and creates objects representing HTML tags and pieces of text. For example, here's some HTML:

```
<HTML>
  <BODY>
    Hello, and welcome to my web page! I work for
    <A HREF="http://industriallogic.com">
      <IMG SRC="http://industriallogic.com/images/logo141x145.gif">
    </A>
  </BODY>
</HTML>
```

Given such HTML, the parser would create objects of the following types:

```
HTMLTag (for the <BODY> tag)
HTMLStringNode (for the String, "Hello, and welcome...")
HTMLLinkTag (for the <A HREF="..."> tag)
```

Since the link tag () contains an image tag () tag, you might wonder what the parser does with it. The image tag, which the parser treats as an HTMLImageTag, is treated as a child of the HTMLLinkTag. When the parser notices that the link tag contains an image tag, it constructs and gives one HTMLImageTag object as a child to the HTMLLinkTag object.

Additional tags in the parser, such as HTMLFormTag, HTMLTitleTag and others, are also child-containers. As I studied some of these classes, it didn't take long to spot duplicate code for storing and handling child nodes. For example, consider the following:

```
public class HTMLLinkTag extends HTMLTag...
    private Vector nodeVector;

    public String toPlainTextString() {
        StringBuffer sb = new StringBuffer();
        HTMLNode node;
        for (Enumeration e=linkData();e.hasMoreElements();)
        {
            node = (HTMLNode)e.nextElement();
            sb.append(node.toPlainTextString());
        }
        return sb.toString();
    }

public class HTMLFormTag extends HTMLTag...
    protected Vector allNodesVector;

    public String toPlainTextString() {
        StringBuffer stringRepresentation = new StringBuffer();
        HTMLNode node;
        for (Enumeration e=getAllNodesVector().elements();e.hasMoreElements();) {
            node = (HTMLNode)e.nextElement();
            stringRepresentation.append(node.toPlainTextString());
        }
        return stringRepresentation.toString();
    }
}
```

Since HTMLFormTag and HTMLLinkTag both contain children, they both have a Vector for storing children, though it goes by a different name in each class. Since both classes need to

support the `toPlainTextString()` operation, which outputs the non-HTML-formatted text of the tag's children, both contain logic to iterate over their children and produce plain text. Yet the code to do this operation is nearly identical in these classes! In fact, there are several nearly-identical methods in the child-container classes, all of which reek from duplication. So let's apply *Extract Composite* to this code:

1. I must first create an abstract class that will become the superclass of the *child-container* classes. Since the child-container classes, like `HTMLLinkTag` and `HTMLFormTag`, are already subclasses of `HTMLTag`, I create the following:

```
public abstract class CompositeTag extends HTMLTag {
    public CompositeTag(
        int tagBegin,
        int tagEnd,
        String tagContents,
        String tagLine) {
        super(tagBegin, tagEnd, tagContents, tagLine);
    }
}
```

2. Now I make the child-containers subclasses of `CompositeTag`:

```
public class HTMLFormTag extends CompositeTag
public class HTMLLinkTag extends CompositeTag
```

and so on...

(Note, for the remainder of this refactoring, I'll only show code from two *child-containers*, `HTMLLinkTag` and `HTMLFormTag`, even though there are others in the code base).

3. I look for a *purely-duplicated* method across all *child-containers* and find `toPlainTextString()`. Since this method has the same name in each *child-container*, I don't have to change its name anywhere. My first step is to pull up the child `Vector` that stores children. I do this using the `HTMLLinkTag` class:

```
public class HTMLLinkTag extends CompositeTag...
    private Vector nodeVector;

public abstract class CompositeTag extends HTMLTag...
    protected Vector nodeVector; // pulled-up field
```

Since I want `HTMLFormTag` to use the same newly-pulled-up `Vector`, `nodeVector` (yes, it's an awful name, I'll change it soon), I rename its local child `Vector` to be `nodeVector`:

```
public class HTMLFormTag extends CompositeTag...
    protected Vector allNodesVector;
    protected Vector nodeVector;
    ...
```

And then I delete this local field (since `HTMLFormTag` inherits it):

```
public class HTMLFormTag extends CompositeTag...
    protected Vector nodeVector;
```

Now I can rename `nodeVector` in the composite:

```
public abstract class CompositeTag extends HTMLTag...
protected Vector nodeVector;
protected Vector children;
```

I'm now ready to pull up the `toPlainTextString()` method to `CompositeTag`. My first attempt at doing this using an automated refactoring tool fails because the two methods aren't identical in `HTMLLinkTag` and `HTMLFormTag`. The trouble is, `HTMLLinkTag` gets an iterator on its children by means of the method, `linkData()`, while `HTMLFormTag` gets an iterator on its children by means of the `getAllNodesVector().elements()`:

```
public class HTMLLinkTag extends CompositeTag
{
    public Enumeration linkData()
    {
        return children.elements();
    }

    public String toPlainTextString()...
        for (Enumeration e=linkData();e.hasMoreElements();)
            ...
}

public class HTMLFormTag extends CompositeTag...
{
    public Vector getAllNodesVector() {
        return children;
    }
    public String toPlainTextString()...
        for (Enumeration e=getAllNodesVector().elements();e.hasMoreElements();)
            ...
}
```

To fix this problem, I must create a consistent method for getting access to a `CompositeTag`'s children. I won't bore you with the steps. I end up with:

```
public abstract class CompositeTag extends HTMLTag...
{
    public Enumeration children() {
        return children.elements();
    }
}
```

...and a version of `toPlainTextString()` that's now identical in `HTMLLinkTag` and `HTMLFormTag`:

```
public String toPlainTextString() {
    StringBuffer plainTextContents = new StringBuffer();
    HTMLNode node;
    for (Enumeration e=children();e.hasMoreElements();) {
        node = (HTMLNode)e.nextElement();
        plainTextContents.append(node.toPlainTextString());
    }
    return plainTextContents.toString();
}
```

The automated refactoring in my IDE now lets me easily pull up `toPlainTextString()` to `CompositeTag`. I run my tests and everything is ok.

4. In this step I repeat step 3 for additional methods that may be pulled-up from the *child-containers* to the *composite*. There happen to be several of these methods. I'll show you one that involves a method called `toHTML()`. This method outputs the HTML of a given node. Both `HTMLLinkTag` and `HTMLFormTag` have their own implementations for this methods. To implement step 3, I must first decide if `toHTML()` is *purely-duplicated* or *partially-duplicated*.

Here's a look at how `HTMLLinkTag` implements the method:

```

public class HTMLLinkTag extends CompositeTag
{
    public String toHTML() {
        StringBuffer sb = new StringBuffer();
        putLinkStartTagInto(sb);
        //sb.append(tagContents.toString());
        HTMLNode node;
        for (Enumeration e = children(); e.hasMoreElements();) {
            node = (HTMLNode)e.nextElement();
            sb.append(node.toHTML());
        }
        sb.append("</A>");
        return sb.toString();
    }

    public void putLinkStartTagInto(StringBuffer sb) {
        sb.append("<A ");
        String key,value;
        int i = 0;
        for (Enumeration e = parsed.keys(); e.hasMoreElements();) {
            key = (String)e.nextElement();
            i++;
            if (key!=TAGNAME) {
                value = getParameter(key);
                sb.append(key+"=\""+value+"\"");
                if (i<parsed.size()-1) sb.append(" ");
            }
        }
        sb.append(">");
    }
}

```

After creating a buffer, `putLinkStartTagInto(...)` deals with getting the contents of the start tag into the buffer, along with any attributes it may have. The start tag would be something like `` or ``, where `HREF` or `NAME` represent attributes of the tag. The tag could have children, such as an `HTMLStringNode`, as in `I'm a string node` or child `HTMLImageTags`. Finally there is the end tag, ``, which must be added to the result buffer before the HTML representation of the tag is returned.

Let's now see how `HTMLFormTag` implements this method:

```

public class HTMLFormTag extends CompositeTag...
{
    public String toHTML() {
        StringBuffer rawBuffer = new StringBuffer();
        HTMLNode node,prevNode=null;
        rawBuffer.append("<FORM METHOD=\""+formMethod+"\" ACTION=\""+formURL+"\"");
        if (formName!=null && formName.length()>0)
            rawBuffer.append(" NAME=\""+formName+"\"");
        Enumeration e = children.elements();
        node = (HTMLNode)e.nextElement();
        HTMLTag tag = (HTMLTag)node;
        Hashtable table = tag.getParsed();
        String key,value;
        for (Enumeration en = table.keys(); en.hasMoreElements();) {
            key=(String)en.nextElement();
            if (!(key.equals("METHOD")
                || key.equals("ACTION")
                || key.equals("NAME")
                || key.equals(HTMLTag.TAGNAME))) {
                value = (String)table.get(key);
                rawBuffer.append(" "+key+"="+value+"");
            }
        }
        rawBuffer.append(">");
        rawBuffer.append(lineSeparator);
        for (; e.hasMoreElements();) {
            node = (HTMLNode)e.nextElement();
            if (prevNode!=null) {
                if (prevNode.elementEnd()>node.elementBegin()) {

```



```

        // Its a new line
        rawBuffer.append(lineSeparator);
    }
}
rawBuffer.append(node.toHTML());
prevNode=node;
}
return rawBuffer.toString();
}

```

This implementation has some similarities and differences to the `HTMLLinkTag` implementation. Therefore, according to the definition in the mechanics, `toHTML()` should be treated as a *partially-duplicated* child-container method. That means that my next step is to see if I can make one implementation of this method by applying the refactoring, *Substitute Algorithm* [F].

It turns out I can. It is easier than it looks because both versions of `toHTML()` essentially do the same three things: output the start tag along with any attributes, output any child tags, output the closed tag. Knowing that, I arrive at a common method for dealing with the start tag, which I pull up to *composite*:

```

public abstract class CompositeTag extends HTMLTag...
{
    public void putStartTagInto(StringBuffer sb) {
        sb.append("<" + getTagName() + " ");
        String key,value;
        int i = 0;
        for (Enumeration e = parsed.keys();e.hasMoreElements();) {
            key = (String)e.nextElement();
            i++;
            if (key!=TAGNAME) {
                value = getParameter(key);
                sb.append(key+"=\""+value+"\"");
                if (i<parsed.size()) sb.append(" ");
            }
        }
        sb.append(">");
    }
}

public class HTMLLinkTag extends CompositeTag...
{
    public String toHTML() {
        StringBuffer sb = new StringBuffer();
        putStartTagInto(sb);
        ...
    }
}

public class HTMLFormTag extends CompositeTag
{
    public String toHTML() {
        StringBuffer rawBuffer = new StringBuffer();
        putStartTagInto(rawBuffer);
        ...
    }
}

```

I perform similar operations to make a consistent way of obtaining HTML from child nodes and from an end tag and all of that work enables me to pull-up one generic `toHTML()` method to the composite:

```

public abstract class CompositeTag extends HTMLTag...
{
    public String toHTML() {
        StringBuffer htmlContents = new StringBuffer();
        putStartTagInto(htmlContents);
        putChildrenTagsInto(htmlContents);
        putEndTagInto(htmlContents);
        return htmlContents.toString();
    }
}

```

To complete this part of the refactoring, I'll continue to move child-related methods to the `CompositeTag`, though I'll spare you the details.

5. The final step involves checking clients of *child-containers* to see if they can now communicate with the child-containers using the `CompositeTag` interface. In this case, there are no such cases in the parser itself, so I'm finished with the refactoring.

Replace Subclasses with Visitor

You've extended one or more hierarchies
with subclasses to obtain new behavior

*Replace the new subclasses with a Visitor
that can provide the new behavior by
visiting hierarchy objects at runtime*

Motivation

Ralph Johnson, one of the authors of *Design Patterns* [GoF], once observed, "Most of the time you don't need Visitor, but when you do need Visitor, you *really* need Visitor!" So when do you *really* need Visitor? I've needed it when I've needed to accumulate new information from an object structure, but the classes of objects in the structure didn't support the behavior I needed.

To accumulate that information, I initially considered two implementations: jamming the new behavior into existing classes using new methods or defining new classes to make the information I needed accessible. This refactoring was inspired from choosing the latter option.

Adding new classes (or, in the case of this example)

the information I needed and to serve as the new class types for the objects in the could take the place of the existing classes when representing particular objects in the object structure.

first tried could add methods to existing classes or create new classes. These two options are what many programmers use to add new behavior to hierarchies. I op

Since all of the classes insimply by defining new classes. Since the objects in the object structure I'm trying to accumulate information from all share the same base class, I

composed of objectwrite new subclasses extend the hierarchy of classes used to produce the objects in the object structure.

to write new subclasses of a class hierarchy that can provide the information you need to accumulate. That solution may work, but it is often much less simple than

Adding new hierarchy subclasses to provide the information I needed and the classes in a hierarchy extend a hierarchy with new behavior and after trying different options, Visitor provided the simplest solution.

Adding new behavior to a hierarchy often occurs by adding new methods to existing hierarchy classes or by adding new hierarchy subclasses.

In both cases, the extended behavior usually has little to do with the main responsibility of the classes in the hierarchy. So a big issue here is the clarity that comes from a good separation of

directwhen the user doesn't find the behavior they need in the existing hierarchy so they add it, irregardless of whether or not it makes sense to add it

Extended behavior in a hierarchy simply means

”, let's first review what Visitors are used for. While they are certainly capable of such sophisticated behavior as altering objects structures while visiting them at runtime, in general, Visitors are used to accumulate information. Visitors obtain their information by visiting

- classes that implement the Composite pattern
- classes that are from a single hierarchy
- classes that are from different hierarchies

When I need to accumulate information from objects at runtime, Visitor is usually the last implementation I consider. That's because there's usually some simpler solution. For example, if an *iterator* is available and can be used to easily accumulate the information I need, I certainly don't need a Visitor. If an iterator is available but it is awkward to use to accumulate the data I'm , I'm done. In general, Here's a loose ordering of implementations I consider when I need to accumulate specific data from objects:

Iterator – if I have easy

The main reason for needing a Visitor is when you want classes in a hierarchy (or two) to provide behavior that doesn't quite belong in the classes themselves. Such behavior is either not the core responsibility of the classes or is so abundant in variations that adding it to the classes would heavily pollute their interfaces.

the classes themselves. Such behavior is usually not core to the purpose of the classes.

In practice, I find I need one when

The I need the classes in a hierarchy to support extended

I've got some behavior I need to add to a I'd I need doesn't quite fit with the core behavior of an existing hierarchy or two.

How right he is! Since 1995 (when I first learned this pattern) I've needed it all of three times and it is likely I could've used something simpler on at least one of those occasions.

Programmers either love or hate this pattern. The ones who love it (or who are patterns-happy

needed it and one of them So I'm always 've wanted to use it but have rarely done so because simpler solutions often do the job.

found the need to use it. right fit. certainly wanted to use it since it is so cool, but I've rarely encountered situations in which implementing this pattern was justified. Of course, as Ralph says, there are times when you need it.

encounter two main situations in which

I've seen plenty of cases where programmers thought they needed Visitor but really didn't (something far simpler would be far du

[discuss reasons to choose Move Accumulation to Collecting Parameter over this refactoring]

[finish this section]

Visitor isn't a pattern In over ten years of programming I've used the Visitor pattern on real systems all of three times and I was probably over-engineering one of them. My friends have similar numbers to report.

Communication	Duplication	Simplicity
ToDo	ToDo	ToDo

Mechanics

1. Define a concrete Visitor class and pass it the data it needs to get its job done (it's common to pass this data via a constructor).

If you're creating your 2nd Visitor class, apply *Extract Superclass* [F] on your first Visitor to produce a base class Visitor, make your new Visitor extend the base class, and change message signatures on *visitees* (defined in step 3) so they accept a base class Visitor parameter instead of a concrete Visitor.

2. Define a method or methods on the Visitor to obtain the information it will accumulate. The method(s) need not be on the interface of the base class Visitor, assuming you created a base class Visitor during step 1.
3. Identify the *visitee*, a class that the Visitor must visit in order to accumulate data it needs. Add to the *visitee* an accept method that takes as a parameter the Visitor or the Visitor's base class (e.g. `public void accept(Visitor visitor)`). Then add a method to the Visitor (and its base class, if there is one) called *visit*, which takes as a parameter the *visitee* (e.g. `public void visitSomeClass(SomeClass someClass)`).
4. Implement the visit method on the Visitor (created during the previous step), making any modifications necessary to give the Visitor access to the necessary *visitee* data and/or logic.

✓ Compile and test that the Visitor now accumulates correct data from the *visitee*.

5. Repeat steps 3 and 4 for every *visitee* that the Visitor must visit in order to accumulate the information it needs.
6. Add a method to the outermost class that you'll be passing the visitor to when you need your visitor to accumulate information. This method will likely iterate over a collection of objects, passing the visitor to the *visitees*. If some of the objects being iterated over don't have an accept method (i.e. aren't *visitees*), you may wish to create a do-nothing accept method on their base class (if there is one) in order to make processing uniform.

✓ Compile and test.

7. Replace all client calls to the custom classes with calls to the Visitor.

✓ Compile and test.

Example 1

It takes a good deal of patience to find a real-world case in which refactoring to Visitor actually makes sense. I found such a case while pair-programming some code that uses an open-source, streaming HTML parser (see <http://sourceforge.net/projects/htmlparser>). To understand this refactoring, I'll need to give you a brief overview of one of the parser's main features: scanners.

As the parser parses HTML (or XML), it recognizes tags and text. For example, consider the following HTML:

```
<HTML>
  <BODY>
    Hello, and welcome to my web page! I work for
    <A HREF="http://industriallogic.com">
      <IMG SRC="http://industriallogic.com/images/logo141x145.gif">
    </A>
  </BODY>
</HTML>
```

In its plain vanilla configuration, the parser identifies the following objects in the above HTML:

```
HTMLTag (for the <BODY> tag)
  HTMLStringNode (for the String, "Hello, and welcome...")
  HTMLTag (for the <A HREF="..."> tag)
    HTMLTag (for the <IMG SRC="..."> tag)
  HTMLEndTag (for the </A> tag)
HTMLEndTag (for the </BODY> tag)
```

By default, the parser finds HTMLTags, HTMLStringNodes and HTMLEndTags. If you want the parser to create more specific objects for common HTML tags, like links, images and forms, you have to register the appropriate scanner for each tag you're interested in. Fortunately, the parser provides a convenience method for registering all of the most common scanners in one method call (it's called `registerScanners()`). If that method had been called, prior to obtaining the found tag objects in the above HTML, the parser would've recognize the following objects:

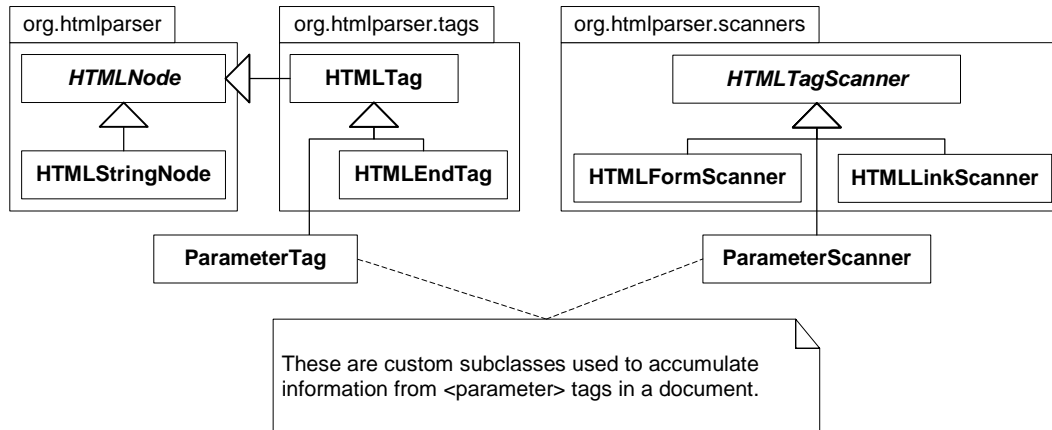
```
HTMLTag (for the <BODY> tag)
  HTMLStringNode (for the String, "Hello, and welcome...")
  HTMLLinkTag (for the <A HREF="...">...</A> tags)
    HTMLImageTag (for the <IMG SRC="..."> tag)
HTMLEndTag (for the </BODY> tag)
```

In other words, once the most common scanners have been registered, the parser can now recognize that `` is an `HTMLLinkTag` object and `` is an `HTMLImageTag` object.

The parser allows programmers to create and register their own custom scanners. Such scanners are usually written to identify non-standard tags in HTML or XML. For example, while working on some code to transform HTTP parameters into XML, my colleague, Somik Raha, and I found that we needed the ability to find the number of `<parameter>` tags in XML like this:

```
<input>
  <parameter name='key1'>value1</param>
  <parameter name='key2'>value2</param>
</input>
```

Since the parser has no idea what a `<parameter>` tag is, we wrote a custom scanner (`ParameterScanner`) and custom tag (`ParameterTag`) to make the parser capable of recognizing `ParameterTag` objects.



We created the `ParameterScanner` and `ParameterTag` test-first. Here's one of our tests:

```
public class UtilsTest extends TestCase...
    public void testCreateParameterMapFromXML() throws Exception {
        String inputXML =
            "<input>" +
            "  <param name='key1'>value1</param>" +
            "  <param name='key2'>value2</param>" +
            "</input>";
        Map paramsMap = Utils.createParameterMapFrom(inputXML);
        String data1 = (String)paramsMap.get("key1");
        String data2 = (String)paramsMap.get("key2");
        assertNotNull("Data 1 should have been found",data1);
        assertNotNull("Data 2 should have been found",data2);
        assertEquals("Data 1","value1",data1);
        assertEquals("Data 2","value2",data2);
    }
}
```

`Utils.createParameterMapFrom(...)` does most of the work in the above test. Here's what it does:

```
public class Utils {
    public static Map createParameterMapFrom(String inputXML)
        throws Exception {
        HTMLReader reader = new HTMLReader(new StringReader(inputXML), "");
        HTMLParser parser = new HTMLParser(reader);
        parser.addScanner(new ParameterScanner("-p"));
        String lastKeyFound=null;
        Map paramsMap = new HashMap();
        for (HTMLEnumeration e = parser.elements();e.hasMoreNodes();) {
            HTMLNode node = e.nextNode();
            if (node instanceof ParameterTag) {
                ParameterTag paramTag = (ParameterTag)node;
                paramsMap.put(paramTag.getKey(),paramTag.getValue());
            }
        }
        return paramsMap;
    }
}
```



```
}
```

Let me explain the contents of the above method. After creating a parser instance, we register our custom scanner by calling `addScanner(...)`. Next, we iterate over all the `HTMLNodes` (the base class for all tag and string classes) which the parser encounters. Since our registered scanner is capable of recognizing `<parameter>` tags, when the parser gets to such a tag, it gives our custom scanner the chance to handle it. Our scanner, which recognizes that the parser has found a tag it cares about, briefly takes control from the parser to produce and return back a `ParameterTag` object. Since all of that happens during iteration of the `HTMLNodes`, it is transparent to our client code. So, to accumulate the map of `ParameterTags` we need, all we have to do is recognize when an `HTMLNode` is an instance of a `ParameterTag`, which is easy using Java's instanceof syntax (`if (node instanceof ParameterTag)`).

Now, while I don't want to make you an expert in scanners, I do want to explain the basics of how they work, so you can understand the refactoring we performed. Essentially, when the parser encounters a tag during parsing, it gives its registered scanners a chance to process the tag in order to produce an instance of an `HTMLTag` subclass. The parser does this by essentially giving up parsing control to the scanner that is interested in the current tag. When our `ParameterScanner` encounters a `<parameter>` tag, which the parser normally recognizes as a plain-vanilla `HTMLTag` instance, it reads in the `HTMLTag`, and then proceeds to read in the next found `HTMLNode`, which is an `HTMLStringNode` that represents the value in `<parameter>value`. At that point, our scanner has all it needs to create and return back a `ParameterTag` instance. Once it returns that object, the parser regains parsing control, which means it can continue to search for remaining `HTMLNodes`.

If I've thoroughly confused you at this point, don't worry: to understand this refactoring you don't have to understand scanners. All you really need to understand is three things:

- specific scanners identify, create and return specific `HTMLNode` subclasses when parsing occurs
- the `ParameterScanner` processes three `HTMLNodes` (`HTMLTag`, `HTMLStringNode`, `HTMLEndTag`) for the text, `<parameter>value</parameter>`, and uses the information in the `HTMLNodes` to produce and return back one `ParameterTag` instance
- the returned `ParameterTags` are used by code in our system to create a map of data we happen to need

Now, without going into the detail of what it takes to write a custom scanner or custom tag, I simply want to show you what the code looks like so you can appreciate how much work needed to be done simply to accumulate information about `<parameter>` tags in our XML:

```
public class ParameterTag extends CompositeTag {
    public ParameterTag(
        HTMLTagData tagData,
        HTMLCompositeTagData compositeTagData) {
        super(tagData, compositeTagData);
    }

    public String getKey() {
        return startTag.getParameter("NAME");
    }

    public String getValue() {
        return ((HTMLStringNode)childTags.elementAt(0)).getText();
    }
}
```

```

public class ParameterScanner extends HTMLTagScanner {
    public ParameterScanner(String filter) {
        super(filter);
    }

    public HTMLTag scan(HTMLTag tag, String url, HTMLReader reader, String currLine)
        throws HTMLParserException {
        HTMLTag startTag = tag;
        HTMLTag endTag = null;
        boolean endTagFound = false;
        HTMLNode node;
        Vector childVector = new Vector();
        do {
            node = reader.readElement();
            if (node instanceof HTMLEndTag) {
                endTag = (HTMLTag)node;
                if (endTag.getText().equalsIgnoreCase("parameter"))
                    endTagFound = true;
            } else
                childVector.addElement(node);
        } while (endTagFound==false && node!=null);
        return new ParameterTag(new HTMLTagData(
            startTag.elementBegin(),
            endTag.elementEnd(),
            startTag.getText(),
            currLine
        ),new HTMLCompositeTagData(
            startTag,endTag,childVector
        )
        );
    }

    public String[] getID() {
        return new String [] {"PARAM"};
    }
}

```

In addition to writing the above custom classes, code to do the accumulation (shown in the method, `Utils.createParameterMapFrom(...)`, above) also had to be written. All that work got me thinking that there had to be an easier way to rip `<parameter>` information out of an XML document.

So I began considering other ways to accumulate the information. I experimented with a SAX (Simple API for XML) implementation, but realized I could come up with a simpler solution using the html parser. Since the parser, by default, visits three `HTMLNode` types, `HTMLTag`, `HTMLStringNode` and `HTMLEndTag`, I wondered if I could write a `Visitor` to visit each encountered `HTMLNode` and figure out what to accumulate. Would such a solution be simpler? Implementing the `Visitor` pattern usually isn't a lightweight solution, but compared with the custom scanner and custom tag code, it had the potential to be simpler and easier. So I attempted the refactoring and was pleased to discover that it did result in simpler, smaller code. You can decide for yourself what you think. Here are the mechanics I followed:

1. I begin by defining a `Visitor`. I do this test-first. Here's a fragment of my first test:

```

public class ParameterVisitorTest extends TestCase...
    public void testVisitTag() {
        ParameterVisitor visitor = new ParameterVisitor();
    }

```

The mechanics instruct me to pass any data my `Visitor` may need using its constructor. In this case, the `Visitor` doesn't need outside data. To make the above test-fragment compile, I define the `ParameterVisitor`:

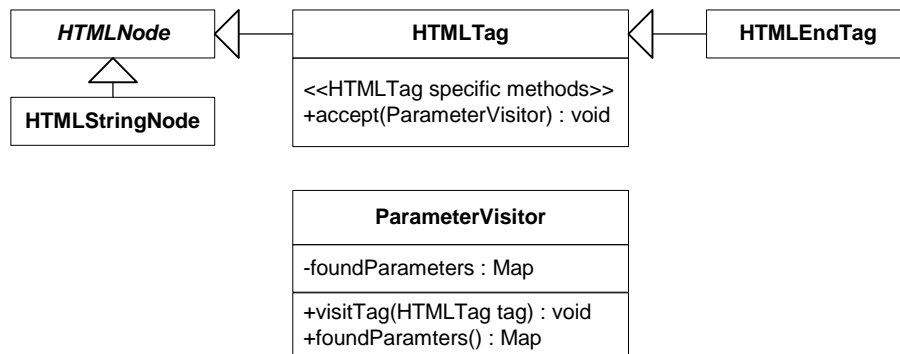
```
public class ParameterVisitor {
}
```

2. Now I define a method on `ParameterVisitor` to return the data I want it to accumulate. In this case, I want it to accumulate a `Map` containing the key and value pairs from the `<parameter>` tags that it will find. So I write the following code, which will be tested shortly:

```
public class ParameterVisitor {
    private Map foundParameters = new HashMap();

    public Map foundParameters() {
        return foundParameters;
    }
}
```

3. I'm now ready to select a *visitee* and pass it a `ParameterVisitor`. When the parser does its parsing, it identifies instances of classes that extend from `HTMLNode`, including `HTMLTag`, `HTMLEndTag` and `HTMLStringNode`. Since the `<parameter>` tag will be treated as an `HTMLTag`, I need to pass a `ParameterVisitor` instance to all `HTMLTag` instances that the parser finds. To do that, I'll need to write an `accept(ParameterVisitor visitor)` method on `HTMLTag`:



I'll implement the `accept(...)` method test-first, which will involve making a more robust test:

```
public class ParameterVisitorTest extends TestCase...
    private static final String INPUT_XML =
        "<input>" +
        "  <param name='key1'>value1</param>"+
        "  <param name='key2'>value2</param>"+
        "</input>";

    public void testVisitTag() throws Exception {
        HTMLParser parser = HTMLParser.createParser(INPUT_XML);
        ParameterVisitor visitor = new ParameterVisitor();
        HTMLEnumeration nodes = parser.elements();
        while (nodes.hasMoreNodes()) {
            HTMLNode node = (HTMLNode)nodes.nextHTMLNode();
            if (node instanceof HTMLTag &&
                !(node instanceof HTMLEndTag)) {
                HTMLTag tag = (HTMLTag)node;
                tag.accept(visitor);
            }
        }
        Map parameters = visitor.foundParameters();
        assertNotNull("parameters not null", parameters);
        Set keys = parameters.keySet();
        assertTrue("key1 found", keys.contains("key1"));
        assertTrue("key2 found", keys.contains("key2"));
    }
}
```

```
}
```

The above test code won't compile until I implement an `accept()` method on `HTMLTag`. That's easy:

```
public class HTMLTag extends HTMLNode...
    public void accept(ParameterVisitor visitor) {
    }
```

My test now runs, but it fails to find "key1" since the `ParameterVisitor` isn't yet populating its `foundParameters` Map. No problem. I must now have the `HTMLTag` visit with the `ParameterVisitor`. This means having the visatee (`HTMLTag`) pass itself to the visitor (`ParameterVisitor`), a technique that is known by the fancy term, *double dispatch*. This isn't hard to implement at all:

```
public class HTMLTag extends HTMLNode...
    public void accept(ParameterVisitor visitor) {
        visitor.visitTag(this);
    }
```

Ofcourse that code won't compile until I declare the `visitTag` method in `ParameterVisitor`:

```
public class ParameterVisitor...
    public void visitTag(HTMLTag tag) {
    }
```

All I have to do now is implement the body of `visitTag` and hopefully the above test will pass. That work occurs in step 4.

4. I'm now going to implement the method body for the new `ParameterVisitor` method, `visitTag(...)`. The mechanics inform me that this step could involve making changes that would give `ParameterVisitor` access to what might be private or protected information in `HTMLTag`. For this code, I don't need to make any access changes, since the `HTMLTag` information needed by `ParameterVisitor` is already publicly available. So I proceed to implement the method:

```
public class ParameterVisitor...
    public void visitTag(HTMLTag tag) {
        if (tag.getTagName().equals("PARAMETER")) {
            String nameAttribute = tag.getParameter("NAME");
            foundParameters.put(nameAttribute, null);
        }
    }
```

This method first checks if the tag argument has the name "PARAMETER." If it does, the code grabs the name attribute from the tag and uses it as a key value when adding an entry to the `foundParameters` map. You might wonder why a null value is being placed into the map. That's because `ParameterVisitor` doesn't yet know how to read the value string contained within an open and closed `<parameter>` tag. It will soon.

The test now passes, which means it's time to go to step 5.

5. I must now repeat steps 3 and 4 for any remaining *visitees* that the `ParameterVisitor` must visit. The next visatee to visit is an `HTMLStringNode`. Why? Because the value string

enclosed within `<parameter>` tags is recognized by the parser as an `HTMLStringNode` and the `ParameterVisitor` needs to get access to value string.

Ok, so we'll take another quick spin through the process. First, I'll update my test to assert that certain `<parameter>` values are correctly accumulated by the visitor:

```
public class ParameterVisitorTest extends TestCase...
    public void testVisitTag() throws Exception {
        HTMLParser parser = HTMLParser.createParser(INPUT_XML);
        ParameterVisitor visitor = new ParameterVisitor();
        HTMLEnumeration nodes = parser.elements();
        while (nodes.hasMoreNodes()) {
            HTMLNode node = (HTMLNode)nodes.nextHTMLNode();
            if (node instanceof HTMLTag &&
                !(node instanceof HTMLEndTag)) {
                HTMLTag tag = (HTMLTag)node;
                tag.accept(visitor);
            }
            if (node instanceof HTMLStringNode) {
                HTMLStringNode string = (HTMLStringNode)node;
                string.accept(visitor);
            }
        }
        Map parameters = visitor.parameters();
        assertNotNull("parameters not null", parameters);
        assertEquals("key1 & value1 found", "value1", parameters.get("key1"));
        assertEquals("key2 & value2 found", "value2", parameters.get("key2"));
        if (parameters.keySet().contains("key3"))
            fail("should not find key3");
    }
}
```

You'll notice a few changes in the above test. First, I updated the code to pass the visitor to `HTMLStringNodes` (that code doesn't yet compile), and second, I've updated the assertions to look for specific key values and to make sure that keys which shouldn't be present, aren't present. This code drives me to implement the `accept(...)` method on `HTMLStringNode`:

```
public class HTMLStringNode extends HTMLNode...
    public void accept(ParameterVisitor visitor) {
    }
}
```

This method needs to do the double-dispatch trick:

```
public class HTMLStringNode extends HTMLNode...
    public void accept(ParameterVisitor visitor) {
        visitor.visitStringNode(this);
    }
}
```

Now `ParameterVisitor` needs to include the `visitStringNode(...)` method:

```
public class ParameterVisitor...
    public void visitStringNode(HTMLStringNode stringNode) {
    }
}
```

Once all of the above code is in place, I can confirm that my test still doesn't pass, since there's no code to deal with the parameter values. So I must implement the body of `visitStringNode(...)`. This method must obtain a string contained within `HTMLTags` bearing the name "PARAMETER." Since I know that `HTMLTag` and `HTMLStringNode` are found in succession by the parser when it encounters `<parameter name="...">value`, I can write the following:

```
public class ParameterVisitor {
    private Map foundParameters = new HashMap();
}
```

```

    private String lastKeyVisited = null;

    public Map foundParameters() {
        return foundParameters;
    }

    public void visitTag(HTMLTag tag) {
        if (tag.getTagName().equals("PARAMETER"))
            lastKeyVisited = tag.getParameter("NAME");
    }

    public void visitStringNode(HTMLStringNode stringNode) {
        if (lastKeyVisited != null) {
            String parameterValue = stringNode.getText();
            foundParameters.put(lastKeyVisited, parameterValue);
            lastKeyVisited = null;
        }
    }
}

```

The test now passes. I'll write more tests to deal with odd cases, but I won't bore you with the details. I'm now ready for the final steps of the refactoring.

6. My job now is to write a method on a class that makes the most sense to receive my visitor and have it visit all `HTMLNodes`. That class is the `HTMLParser`. I want to instantiate it with data to parse and then give it a `Visitor` to obtain information about that data. This means that when I pass an instance of the `Visitor` to the parser, it will turn around and send the `Visitor` on a mission to visit all of the `HTMLNodes` it finds. Once again, I'll do this all test first:

```

public class ParameterVisitorTest extends TestCase...
    public void testVisitTag() throws Exception {
        HTMLParser parser = HTMLParser.createParser(INPUT_XML);
        ParameterVisitor visitor = new ParameterVisitor();
        parser.visitAllNodesWith(visitor);
        Map parameters = visitor.foundParameters();
        assertNotNull("parameters not null", parameters);
        assertEquals("key1 & value1 found", "value1", parameters.get("key1"));
        assertEquals("key2 & value2 found", "value2", parameters.get("key2"));
        if (parameters.keySet().contains("key3"))
            fail("should not find key3");
    }
}

```

To compile the above test, I write

```

public class HTMLParser...
    public void visitAllNodesWith(ParameterVisitor visitor) {
    }
}

```

Looking at that method signature, you might wonder why I'm passing in a very specific `ParameterVisitor` to the `HTMLParser`. Why not program to an interface, not a specific implementation? I will only do that when I have 2 or more concrete `Visitors`. When that time comes, step 2 in the mechanics will guide me through the process of updating method signatures to accept a `Visitor` base class instead of a concrete `Visitor`.

It's now time to implement the above method. I write the following code:

```

public class HTMLParser...
    public void visitAllNodesWith(ParameterVisitor visitor)
        throws HTMLParserException {
        HTMLEnumeration nodes = elements();
        while (nodes.hasMoreNodes()) {
            HTMLNode node = nodes.nextHTMLNode();
            node.accept(visitor);
        }
    }
}

```

```
    }  
}
```

This code doesn't compile because the abstract class, `HTMLNode`, doesn't have an `accept(...)` method. The older test code obtained instances of specific `HTMLNode` subclasses and called their `accept` methods directly. That was fine at the time, but now it's awkward. I want one simple way to have a `Visitor` visit all visited `HTMLNodes`. So I implement a do-nothing implementation of `HTMLNode.accept(...)`:

```
public abstract class HTMLNode...  
    public void accept(ParameterVisitor visitor) {  
    }
```

Now the above code compiles and my tests run. The `ParameterVisitor` is done! Now it's time to get rid of a lot of unnecessary code.

7. I now go back to the client code that used the `ParameterScanner` and `ParameterTag` and alter it to use the `ParameterVisitor`:

```
public class Utils {  
    public static Map createParameterMapFrom(String inputXML)  
        throws Exception {  
        HTMLReader reader = new HTMLReader(new StringReader(inputXML), "");  
        HTMLParser parser = new HTMLParser(reader);  
        ParameterVisitor paramVisitor = new ParameterVisitor();  
        parser.visitAllNodesWith(paramVisitor);  
        return paramVisitor.getParameterMap();  
  
        parser.addScanner(new ParameterScanner("-p-"));  
        String lastKeyFound=null;  
        Map paramsMap = new HashMap();  
        for (HTMLEnumeration e = parser.elements(); e.hasMoreNodes();) {  
            HTMLNode node = e.nextNode();  
            if (node instanceof ParameterTag) {  
                ParameterTag paramTag = (ParameterTag)node;  
                paramsMap.put(paramTag.getKey(), paramTag.getValue());  
            }  
  
        }  
        return paramsMap;  
    }  
}
```

Finally, I delete the now unnecessary `ParameterScanner` and `ParameterTag`.

Move Rendering To Visitor

Replace Multiple Instantiations with Flyweight

Motivation

Mechanics

Example

Replace Singleton with Constant

Motivation

Mechanics

Example

```
public void denied(ApplicationPermission permission) {  
    permission.setState(ApplicationPermissionDenied.getInstance());  
}
```

becomes

```
public void denied(ApplicationPermission permission) {  
    permission.setState(permission.DENIED);  
}
```

References

[Beck]

Beck, Kent. *Smalltalk Best Practice Patterns*. Upper Saddle River, N.J.: Prentice Hall, 1997.

[Bloch]

Bloch, Joshua. *Effective Java*. Addison-Wesley, 2001.

[Evans&Fowler] (fill in this reference)

[F]

Fowler, Martin. *Refactoring: Improving the Design of Existing Code*. Addison-Wesley, 1999.

[GOF]

Erich Gamma, Richard Helm, Ralph Johnson, John Vlissides. *Design Patterns: Elements of Reusable Object Oriented Software*. Reading, Mass.: Addison-Wesley, 1995.

[JUnit]

Kent Beck and Erich Gamma. JUnit Testing Framework. Available on the Web (<http://www.junit.org>).

Appendix A – Naming Conventions

[describe forName, writeOn, claimedBy, etc].

Appendix B – Loan Terminology

A few of the code fragments used in the examples in this book are based on financial systems that calculate numbers for Loans. If you don't have experience writing systems like that, you may find that the example code is hard to understand. No problem. You don't need to know much to make sense of this code. The following describes the three major loan types used in the example code:

- Term Loan: often abbreviated as a TL, is the simplest of loans: I give you \$100 and ask you to pay it back by some date, which is known as the maturity date of the loan.
- Revolver: a Revolver is an instrument that provides “revolving credit”, like a credit card with a spending limit and expiry date. Financial companies often abbreviate Revolvers as “RC.”
- RCTL – this is a combination of a Revolver and Term Loan. The loan starts its life as a Revolver, and on its expiry date, becomes a Term Loan. RCTLs have both expiry and maturity dates.

It is common to calculate numbers for loans, such as capital, risk-adjusted capital, return on capital, etc. When we do risk-adjusted calculations, we often need to use the numbers from some risk-table. In general, the kinds of calculations done on the various loan types shouldn't effect your understanding of the refactoring steps.

Conclusion

Stay tuned for more refactorings. This work is only the beginning of a larger work on this subject. I welcome your thoughts and feedback. If you are interested in seeing the latest copies of this work, please visit <http://industriallogic.com/xp/refactoring/>

Acknowledgements

I'd like to thank my wife, Tracy, for her loving support and continuous encouragement.

Eric Evans has contributed more than any one else to making this work what it is today. I want to thank him for his continued support, thoughtful conversations, great ideas and feedback.

I'd also like to thank the following people:

- Russ Rufer and all of the many great members of the Silicon Valley Patterns Group (Tracy Bialik, Alan Harriman, Chris Lopez, Charlie Toland, Bob Evans, John Brewer, Jeff Miller, Patrick Manion, Debbie Utley, Carol Thistlethwaite, Summer Misherghi, Ted Young, Siqing Zhang). Your feedback has been invaluable.
- Robert Hirshfeld, for helping clarify the Decorator mechanics section.
- Martin Fowler for inspiration and encouragement, for giving me the advice that I once gave him (i.e. use code sketches at the beginning of each refactoring) and for numerous helpful suggestions and ideas.
- Kent Beck for his reviews, suggestions and some illuminating pair-programming sessions in Alghero, Italy.
- John Vlissides for his reviews and suggestions.
- Ralph Johnson, Brian Foote, Brian Marick, Don Roberts, John Brant and others from the University of Illinois.
- Somik Raha – for many great pairing sessions, refactoring ideas and a small amount of poor code he once wrote when he was tired which provided great refactoring material.
- Ken Thomases – for pointing out an error in the Type-Safe Enum mechanics section.
- Sven Gorts – for being the quickest reviewer on the net. I had only to announce
- Many thanks to the following folks who provided excellent suggestions: Rob Mee, Jeff Grigg, Kaoru Hosokawa, Don Hinton, Andrew Swan, Erik Meade, Craig Demyanovich, Dave Hoover.

