Introduction to Refactoring

1

Sutee Sudprasert

Credits

- Refactoring : Improving the design of existing code Martin Fowler
- Design Patterns GOF

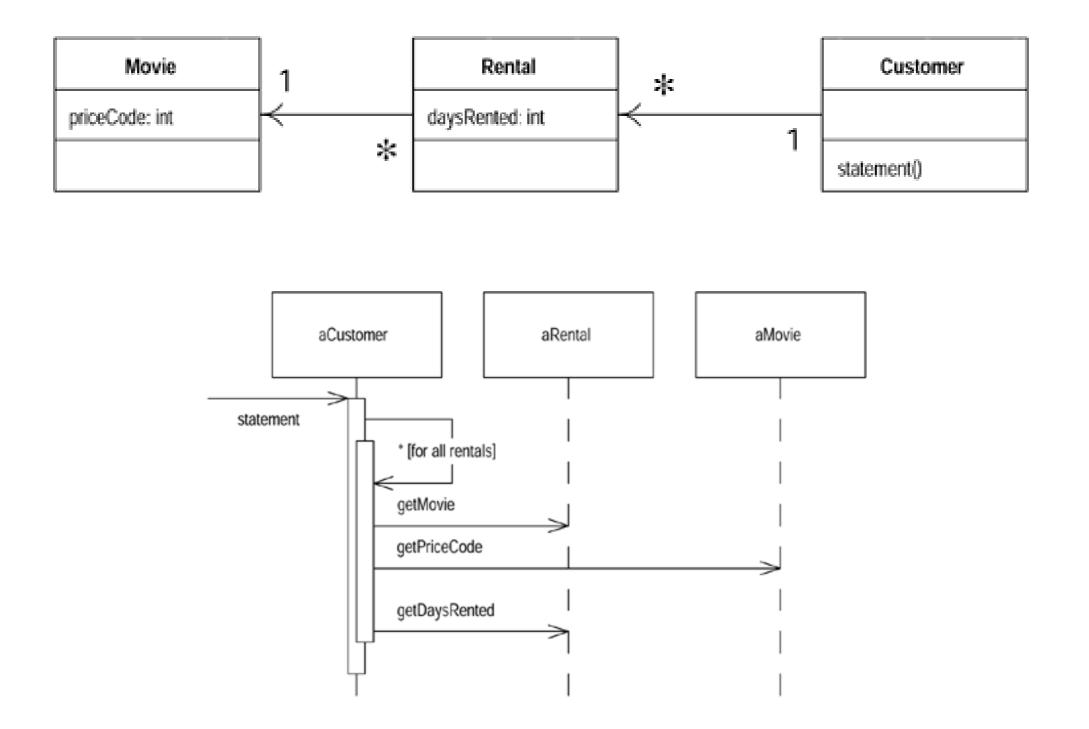
What is refactoring?

- "Refactoring is the process of changing a software system in such a way that it does not alter the external behavior of the code yet improves its internal structure." - Refactoring : Preface
 - You are improving the design of the code after it has been writing.
 - A good design comes first, and the coding comes second.
 - A good design may turn to bad over time the code will be modified.
- With refactoring you can take a bad design, and rework it into well-designed code.

Refactoring, a First Example (In Python)

- The sample program is a program to calculate and print a statement of a customer's charges at a video store.
 - Input: movies that a customer rented and for how long
 - Output: the charges which depend on
 - how long the movie is rented
 - identifies the type movie (regular, children's, new releases)

The starting point



The starting point

• python code

What's wrong with this code?

- It is not well designed and certainly no object oriented.
 - There's nothing wrong with a *quick and dirty simple program*
 - But there are some real problems with this program if this is a representative fragment of a *more complex system*
- The statement routine in Customer class is too long and does many of things that it does should really be done by the other classes
- A poorly designed system is hard to change because it is hard to figure out where the changes are needed (it is easy to make a mistake and introduce bugs)

What's wrong of this code?

- Suppose the users would like a statement printed in HTML
 - it is impossible to reuse any of behavior of the current statement method
 - you can just copy the statement method and make whatever changes you need
- What happens when the charging rules change?
 - you have to fix both statement and htmlStatement
 - if you are writing a program that you don't expect to change, then cut and paste is fine.

What's wrong of this code?

- What if the user want to make changes to the way the classify movies, but they haven't yet decided on the change they are going to make? These changes will affect both
 - the way renters are charged for movie
 - the way that frequent renter points are calculated
- The statement method is where the changes have to be made to deal with changes in classification and charging rules
- Furthermore, as the rules grow in complexity it's going to be harder to figure out where to make the changes and harder to make them without making a mistake.

First step : Extract Method

```
# determine amount for each line
if rental.movie.price_code == Movie.REGULAR:
    this_amount += 2.0
    if rental.days_rented > 2:
        this_amount += (rental.days_rented - 2) * 1.5
elif rental.movie.price_code == Movie.NEW_RELEASE:
    this_amount += rental.days_rented * 3
elif rental.movie.price_code == Movie.CHILDRENS:
    this_amount += 1.5
    if rental.days_rented > 3:
        this_amount += (rental.days_rented - 3) * 1.5
```

amount_for(rental) : return this_amount

```
it doesn't make sense for this context
def amount_fc*(self, rental):
    this_amount = 0.0
    # determine amount for each line
    if rental.movie.price_code == Movie.REGULAR:
        this_amount += 2.0
        if rental.days_rented > 2:
            this_amount += (rental.days_rented - 2) * 1.5
    elif rental.movie.price_code == Movie.NEW_RELEASE:
        this_amount += rental.days_rented * 3
    elif rental.movie.price_code == Movie.CHILDRENS:
        this_amount += 1.5
        if rental.days_rented > 3:
            this_amount += (rental.days_rented - 3) * 1.5
    return this amount
```

```
def amount_for(self, rental):
    result = 0.0
    # determine amount for each line
    if rental.movie.price_code == Movie.REGULAR:
        result += 2.0
        if rental.days_rented > 2:
            result += (rental.days_rented - 2) * 1.5
    elif rental.movie.price_code == Movie.NEW_RELEASE:
        result += rental.days_rented * 3
    elif rental.movie.price_code == Movie.CHILDRENS:
        result += 1.5
        if rental.days_rented > 3:
        result += (rental.days_rented - 3) * 1.5
    result += (rental.days_rented - 3) * 1.5
```

Is renaming worth the effort?

```
def amount_for(self, rental):
    result = 0.0
    # determine amount for each line
    if rental.movie.price_code == Movie.REGULAR:
        result += 2.0
        if rental.days_rented > 2:
            result += (rental.days_rented - 2) * 1.5
    elif rental.movie.price_code == Movie.NEW_RELEASE:
        result += rental.days_rented * 3
    elif rental.movie.price_code == Movie.CHILDRENS:
        result += 1.5
        if rental.days_rented > 3:
        result += (rental.days_rented - 3) * 1.5
    result += (rental.days_rented - 3) * 1.5
```

Is renaming worth the effort?

```
def amount_for(self, rental):
    result = 0.0
    # determine amount for each line
    if rental.movie.price_code == Movie.REGULAR:
        result += 2.0
        if rental.days_rented > 2:
            result += (rental.days_rented - 2) * 1.5
    elif rental.movie.price_code == Movie.NEW_RELEASE:
        result += rental.days_rented * 3
    elif rental.movie.price_code == Movie.CHILDRENS:
        result += 1.5
        if rental.days_rented > 3:
            result += (rental.days_rented - 3) * 1.5
    return result
```

Any fool can write code that a computer can understand. Good programmers write code that humans can understand.

```
class Customer(object):
. . .
def amount for(self, rental):
    result = 0.0
    # determine amount for each line
    if rental.movie.price code == Movie.REGULAR:
        result += 2.0
        if rental.days rented > 2:
            result += (rental.days_rented - 2) * 1.5
    elif rental.movie.price code == Movie.NEW RELEASE:
        result += rental.days rented * 3
    elif rental.movie.price code == Movie.CHILDRENS:
        result += 1.5
        if rental.days rented > 3:
            result += (rental.days rented - 3) * 1.5
    return result
```

This method doesn't use any data from Customer class

```
class Rental(object):
. . .
def get charge(self):
    result = 0.0
    # determine amount for each line
    if self.movie.price code == Movie.REGULAR:
        result += 2.0
        if self.days rented > 2:
            result += (self.days rented - 2) * 1.5
    elif self.movie.price_code == Movie.NEW_RELEASE:
        result += self.days_rented * 3
    elif self.movie.price code == Movie.CHILDRENS:
        result += 1.5
        if self.days_rented > 3:
            result += (self.days rented - 3) * 1.5
    return result
```

```
class Customer(object):
...
def amount_for(self, rental):
    return rental.get_charge()
def statement(self):
    total_amount = 0.0
    frequent_renter_points = 0
    result = 'Rental Record for %s\n' % (self.name)
    for rental in self._rentals:
        this_amount = self.amount_for(rental)
    ...
```

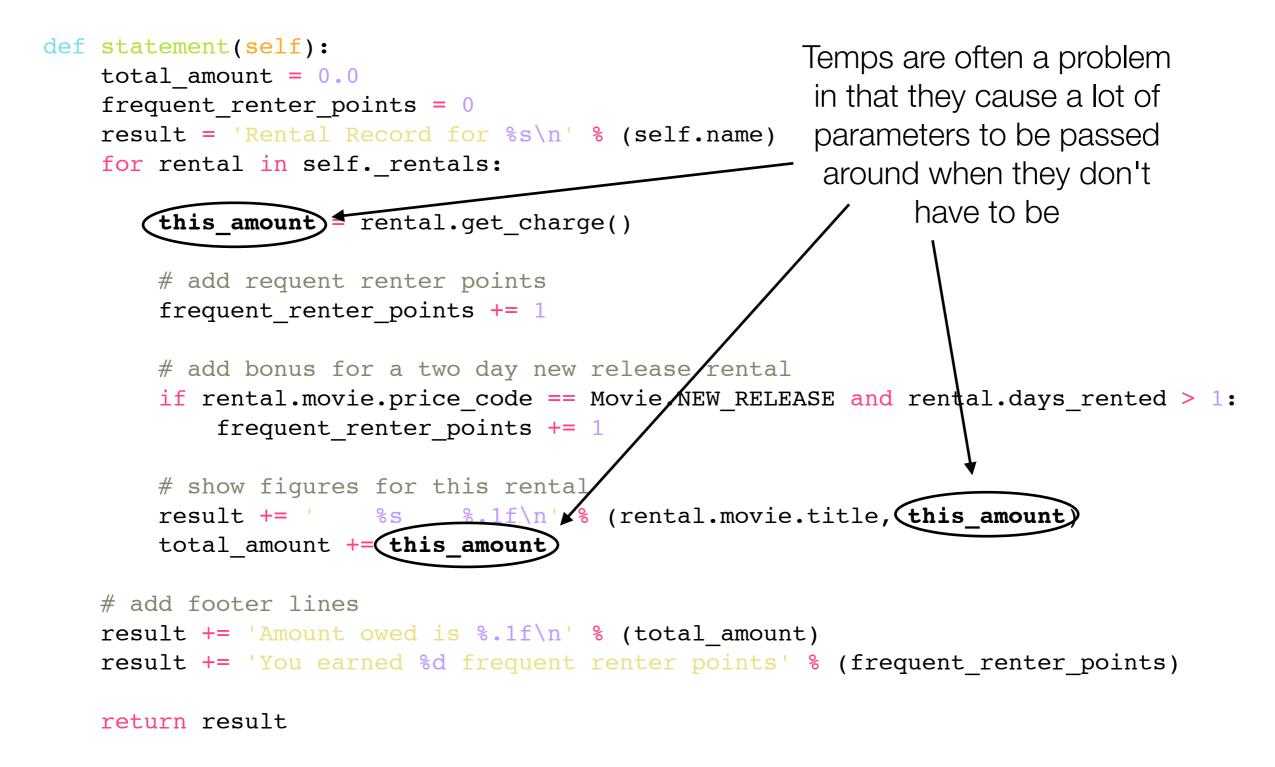
```
class Customer(object):
. . .
def amount for(self, rental):
    return rental.get charge()
def statement(self):
    total amount = 0.0
    frequent_renter_points = 0
    result = 'Rental Record for %s\n' % (self.name)
    for rental in self. rentals:
        this amount = self.amount for(rental)
    . . .
class Customer(object):
. . .
def statement(self):
    total amount = 0.0
    frequent_renter_points = 0
    result = 'Rental Record for %s\n' % (self.name)
    for rental in self._rentals:
        this_amount = rental.get_charge()
    . . .
```

Forth step : Replace Temp with Query

```
def statement(self):
   total amount = 0.0
    frequent renter_points = 0
    result = 'Rental Record for %s\n' % (self.name)
    for rental in self. rentals:
        this amount = rental.get charge()
       # add requent renter points
        frequent renter points += 1
       # add bonus for a two day new release rental
        if rental.movie.price code == Movie.NEW RELEASE and rental.days rented > 1:
            frequent renter points += 1
       # show figures for this rental
        result += ' %s %.lf\n' % (rental.movie.title, this_amount)
       total amount += this amount
    # add footer lines
    result += 'Amount owed is .1f n'  (total amount)
    result += 'You earned %d frequent renter points' % (frequent renter points)
```

return result

Forth step : Replace Temp with Query



Forth step : Replace Temp with Query

```
def statement(self):
    total amount = 0.0
    frequent renter_points = 0
    result = 'Rental Record for %s\n' % (self.name)
    for rental in self. rentals:
        # add requent renter points
        frequent renter points += 1
        # add bonus for a two day new release rental
        if rental.movie.price code == Movie.NEW RELEASE and rental.days rented > 1:
            frequent renter points += 1
        # show figures for this rental
        result += ' %s %.lf\n' % (rental.movie.title, rental.get_charge())
        total amount += rental.get charge()
    # add footer lines
    result += 'Amount owed is %.lf\n' % (total_amount)
    result += 'You earned %d frequent renter points' % (frequent renter points)
    return result
```

Fifth step : Extract Method

```
def statement(self):
    total_amount = 0.0
    frequent_renter_points = 0
    result = 'Rental Record for %s\n' % (self.name)
    for rental in self._rentals:
```

```
# add requent renter points
frequent_renter_points += 1
# add bonus for a two day new release rental
if rental.movie.price_code == Movie.NEW_RELEASE and rental.days_rented > 1:
    frequent_renter_points += 1
```

```
# show figures for this rental
result += ' %s %.lf\n' % (rental.movie.title, rental.get_charge())
total_amount += rental.get_charge()
```

```
# add footer lines
result += 'Amount owed is %.lf\n' % (total_amount)
result += 'You earned %d frequent renter points' % (frequent_renter_points)
```

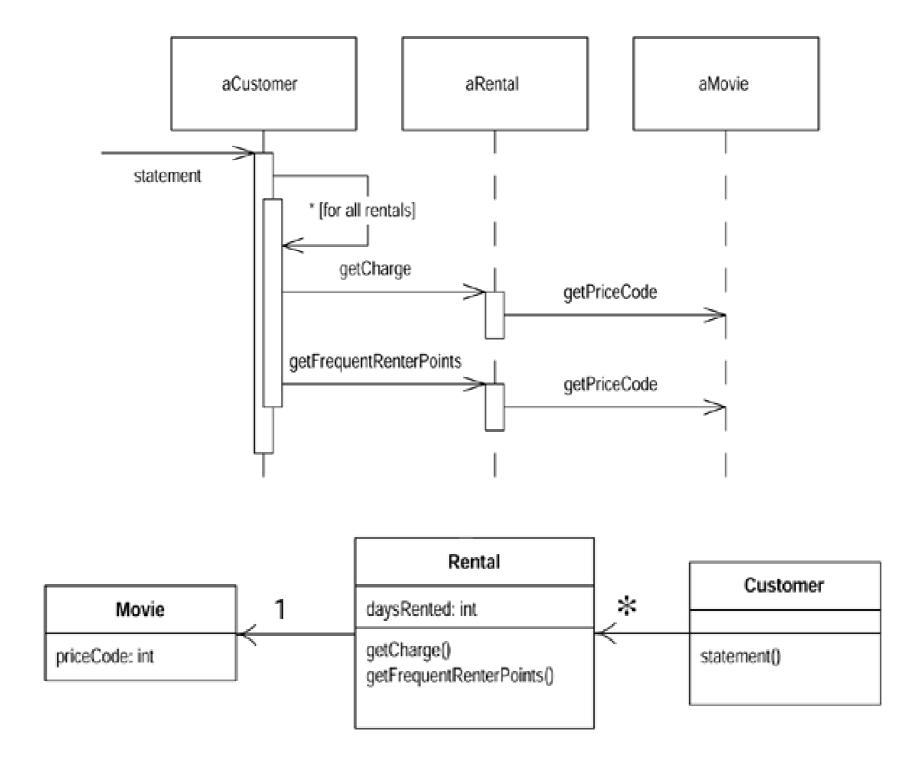
```
return result
```

Fifth step : Extract Method

```
def statement(self):
    total amount = 0.0
    frequent renter points = 0
    result = 'Rental Record for %s\n' % (self.name)
    for rental in self. rentals:
        frequent renter points += rental.get frequent renter points()
        # show figures for this rental
        result += ' %s %.lf\n' % (rental.movie.title, rental.get charge())
        total amount += rental.get charge()
    # add footer lines
    result += 'Amount owed is %.lf\n' % (total amount)
    result += 'You earned %d frequent renter points' % (frequent renter points)
    return result
class Rental(object):
 . . .
```

```
def get_frequent_renter_points(self):
    if self.movie.price_code == Movie.NEW_RELEASE and self.days_rented > 1:
        return 2
    else:
        return 1
```

Sequence diagrams and Class diagram



One loop, One function

```
def statement(self):
    total_amount = 0.0
    frequent_renter_points = 0
    result = 'Rental Record for %s\n' % (self.name)
    for rental in self._rentals:
        frequent_renter_points += rental.get_frequent_renter_points()
        # show figures for this rental
        result += ' %s %.lf\n' % (rental.movie.title, rental.get_charge())
        total_amount += rental.get_charge()
    # add footer lines
    result += 'Amount owed is %.lf\n' % (total_amount)
    result += 'You earned %d frequent renter points' % (frequent_renter_points)
    return result
```

You should make each loop perform only one function.

One loop, One function

add footer lines
result += 'Amount owed is %.lf\n' % (total_amount)
result += 'You earned %d frequent renter points' % (frequent_renter_points)
return result

You should make each loop perform only one function.

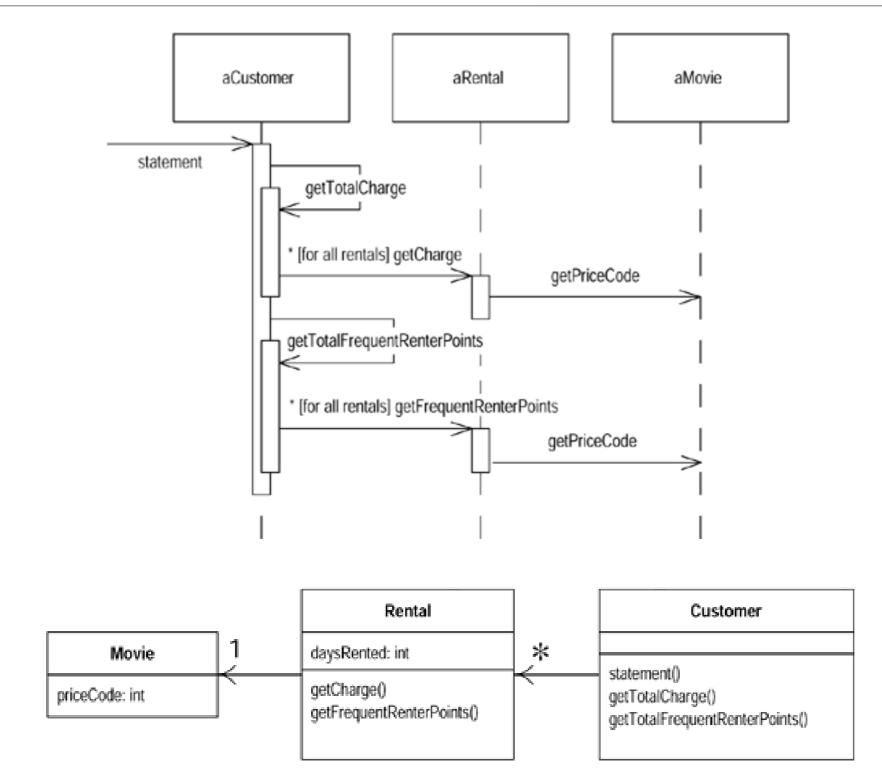
Sixth step : Replace Temp with Query

```
def statement(self):
    total_amount = 0.0
    frequent_renter_points = 0
    result = 'Rental Record for %s\n' % (self.name)
    for rental in self._rentals:
        frequent_renter_points += rental.get_frequent_renter_points()
        # show figures for this rental
        result += ' %s %.lf\n' % (rental.movie.title, rental.get_charge())
        total_amount += rental.get_charge()
        # add footer lines
        result += 'Amount owed is %.lf\n' % (total_amount)
        result += 'You earned %d frequent renter points' % (frequent_renter_points)
        return result
```

Sixth step : Replace Temp with Query

```
def get total charge(self):
    result = 0.0
    for rental in self. rentals:
        result += rental.get charge()
    return result
def get total frequent renter points(self):
    result = 0.0
    for rental in self. rentals:
        result += rental.get frequent renter points()
    return result
def statement(self):
    result = 'Rental Record for %s\n' % (self.name)
    for rental in self. rentals:
        # show figures for this rental
        result += ' %s %.lf\n' % (rental.movie.title, rental.get charge())
    # add footer lines
    result += 'Amount owed is .1f n'  (self.get total charge())
    result += 'You earned %d frequent renter points' % \
        (self.get total frequent renter points())
    return result
```

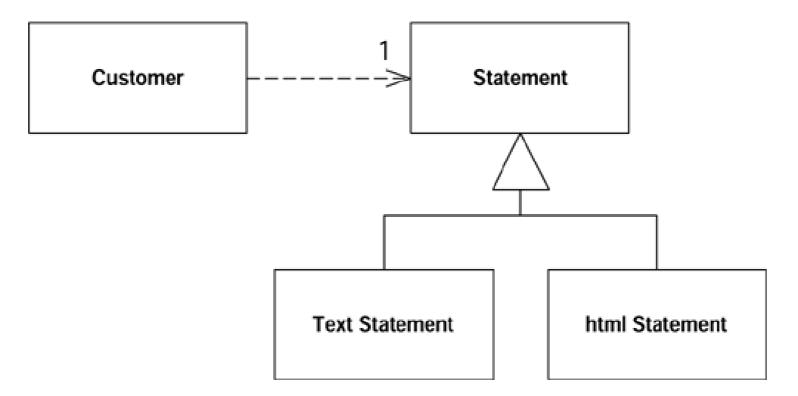
Sequence diagrams and Class diagram



HTML Statement

```
def html_statement(self):
    result = '<hl>Rentals for <em>%s</em></hl>\n' % (self.name)
    for rental in self._rentals:
        # show figures for this rental
        result += '%s : %.lf<br>\n' % (rental.movie.title, rental.get_charge())
    # add footer lines
    result += 'You owe <em>%.lf</em>\n' % (self.get_total_charge())
    result += 'On this rental you earned <em>%d</em> frequent renter points' % \
        (self.get_total_frequent_renter_points())
    return result
```

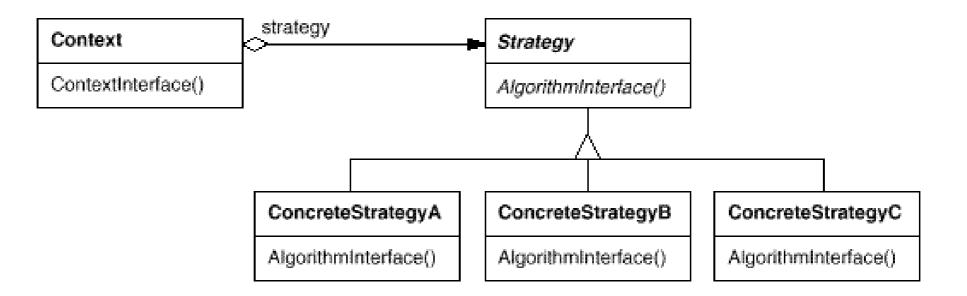
statement and html_statement methods perform similar steps in the same order, yet the steps are different.



First, we have to create a separate **strategy** hierarchy for printing the statements and move the two statement methods over to the subclasses.

Design Patterns : Strategy

- Define a family of algorithms, encapsulate each one, and make them interchangeable. Strategy <u>lets the algorithm vary independently from clients</u> that use it.
- Structure



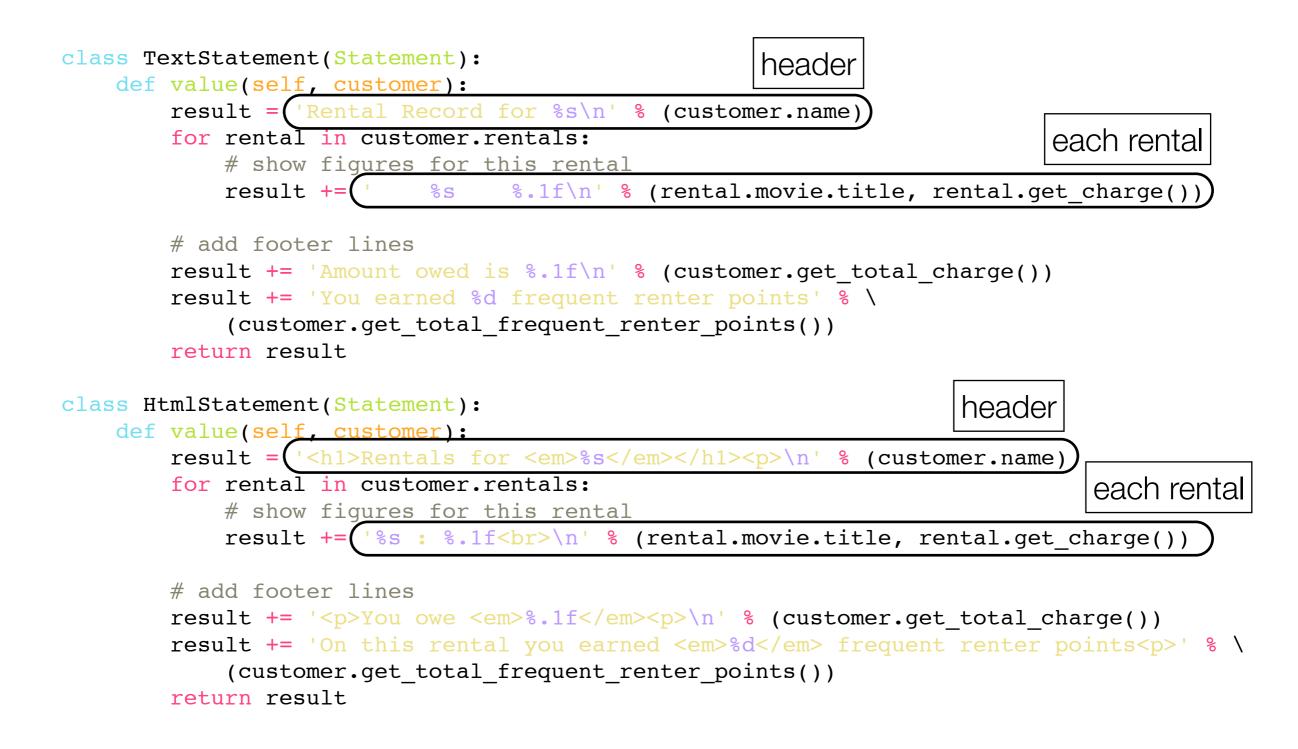
```
class Statement(object):
   def value(self, customer):
        raise NotImplementedError
class TextStatement(Statement):
   def value(self, customer):
        result = 'Rental Record for %s\n' % (customer.name)
        for rental in customer.rentals:
            # show figures for this rental
           result += ' %s %.lf\n' % (rental.movie.title, rental.get charge())
       # add footer lines
        result += 'Amount owed is %.lf\n' % (customer.get total charge())
        result += 'You earned %d frequent renter points' % \
            (customer.get total frequent renter points())
        return result
class HtmlStatement(Statement):
   def value(self, customer):
        result = '<h1>Rentals for <em>%s</em></h1>\n' % (customer.name)
        for rental in customer.rentals:
            # show figures for this rental
            result += '%s : %.1f<br>\n' % (rental.movie.title, rental.get charge())
        # add footer lines
        result += 'You owe <em>%.lf</em>\n' % (customer.get total charge())
        result += 'On this rental you earned <em>%d</em> frequent renter points' % \
            (customer.get total frequent renter points())
        return result
```

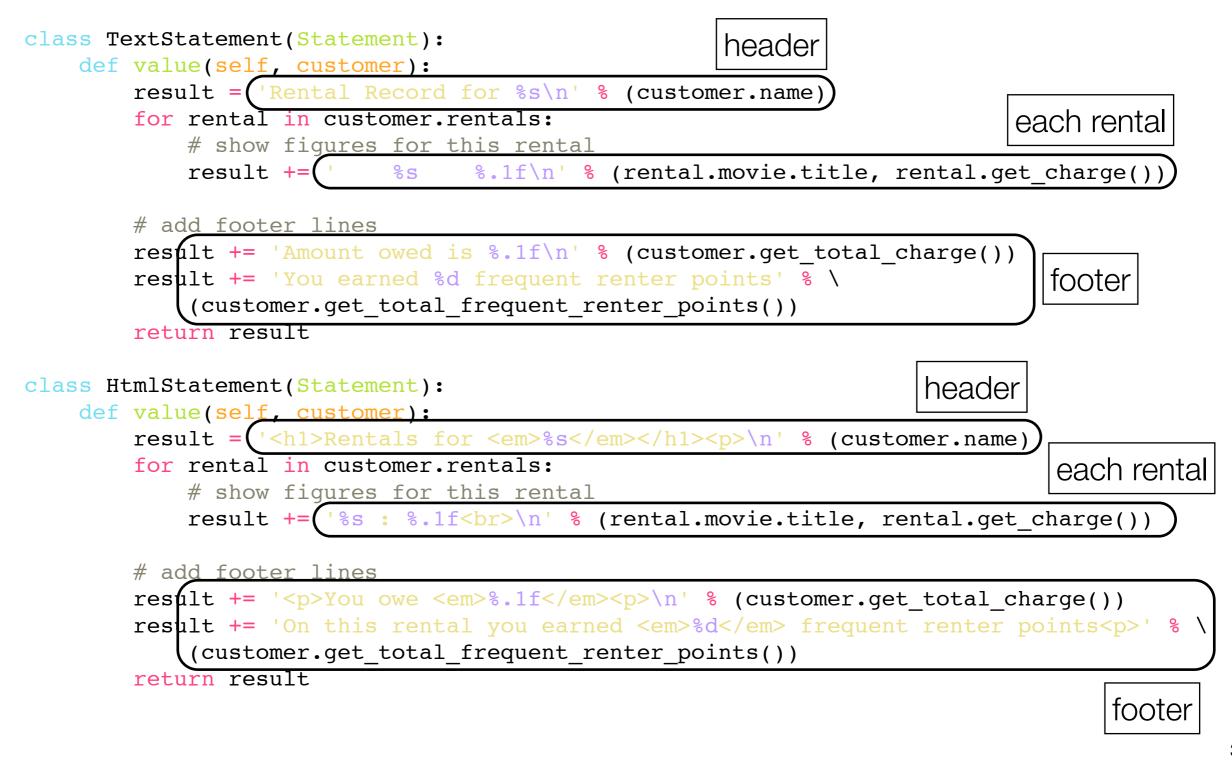
```
class Customer(object):
...
def html_statement(self):
   return HtmlStatement().value(self)
def statement(self):
   return TextStatement().value(self)
```

Now, we can separate the varying code from the similar code by using Extract Method to extract the pieces that are different between the two methods.

```
class TextStatement(Statement):
   def value(self, customer):
       result = 'Rental Record for %s\n' % (customer.name)
       for rental in customer.rentals:
           # show figures for this rental
           result += ' %s %.lf\n' % (rental.movie.title, rental.get_charge())
       # add footer lines
       result += 'Amount owed is .1f n'  (customer.get total charge())
       result += 'You earned %d frequent renter points' % \
            (customer.get total frequent renter points())
       return result
class HtmlStatement(Statement):
   def value(self, customer):
       result = '<h1>Rentals for <em>%s</em></h1>\n' % (customer.name)
       for rental in customer.rentals:
           # show figures for this rental
           result += '%s : %.1f<br>\n' % (rental.movie.title, rental.get charge())
       # add footer lines
       result += 'You owe <em>%.lf</em>\n' % (customer.get total charge())
       result += 'On this rental you earned <em>%d</em> frequent renter points' % \
            (customer.get total frequent renter points())
       return result
```

```
class TextStatement(Statement):
                                                   header
   def value(self, customer):
       result = ('Rental Record for %s\n' % (customer.name)
       for rental in customer.rentals:
           # show figures for this rental
           result += ' %s %.lf\n' % (rental.movie.title, rental.get_charge())
       # add footer lines
       result += 'Amount owed is %.lf\n' % (customer.get total charge())
       result += 'You earned %d frequent renter points' % \
            (customer.get total frequent renter points())
       return result
                                                                 header
class HtmlStatement(Statement):
   def value(self, customer):
       result = ( '<h1>Rentals for <em>%s</em></h1>\n' % (customer.name)
       for rental in customer.rentals:
           # show figures for this rental
           result += '%s : %.1f<br>\n' % (rental.movie.title, rental.get charge())
       # add footer lines
       result += 'You owe <em>%.lf</em>\n' % (customer.get total charge())
       result += 'On this rental you earned <em>%d</em> frequent renter points' % \
            (customer.get total frequent renter points())
       return result
```





```
class TextStatement(Statement):
   def value(self, customer):
       result = self.header string(customer)
       for rental in customer.rentals:
           # show figures for this rental
            result += self.each rental string(rental)
       # add footer lines
       result += self.footer string(customer)
       return result
    def header string(self, customer):
       return 'Rental Record for s \in n'  (customer.name)
    def each_rental_string(self, rental):
       return ' %s %.lf\n' % (rental.movie.title, rental.get charge())
   def footer string(self, customer):
       return 'Amount owed is %.lf\n' % (customer.get total charge()) + \
               'You earned %d frequent renter points' % \
                    (customer.get total frequent renter points())
```

```
class HtmlStatement(Statement):
   def value(self, customer):
       result = self.header string(customer)
       for rental in customer.rentals:
           # show figures for this rental
           result += self.each rental string(rental)
       # add footer lines
       result += self.footer string(customer)
       return result
   def header string(self, customer):
       return '<h1>Rentals for <em>%s</em></h1>\n' % (customer.name)
   def each_rental_string(self, rental):
       return '%s : %.lf<br>\n' % (rental.movie.title, rental.get charge())
   def footer string(self, customer):
       return 'You owe <em>%.lf</em>\n' % (customer.get total charge()) + \
               'On this rental you earned <em>%d</em> frequent renter points' %
                    (customer.get total frequent renter points())
```

```
class Statement(object):
    def value(self, customer):
        result = self.header_string(customer)
        for rental in customer.rentals:
            # show figures for this rental
            result += self.each_rental_string(rental)
        # add footer lines
        result += self.footer_string(customer)
        return result
    def header_string(self, customer):
        raise NotImplementedError
    def each_rental_string(self, rental):
        raise NotImplementedError
```

Finally, pull the value method from two subclasses to their super class

Another Ugly Code

```
def get_charge(self):
    result = 0.0
    # determine amount for each line
    if self.movie.price_code == Movie.REGULAR:
        result += 2.0
        if self.days_rented > 2:
            result += (self.days_rented - 2) * 1.5
    elif self.movie.price_code == Movie.NEW_RELEASE:
        result += self.days_rented * 3
    elif self.movie.price_code == Movie.CHILDRENS:
        result += 1.5
        if self.days_rented > 3:
            result += (self.days_rented - 3) * 1.5
    result += (self.days_rented - 3) * 1.5
```

What are the problems of this code?

Another Ugly Code

- It is a bad idea to do a switch based on an attribute of another object. If you
 must use a switch statement, it should be on your own data, not on someone
 else's.
- Keeping getCharge in the Movie class has the least ripple effect from adding new movie types or editing the existing ones.

Eighth step : Move Method

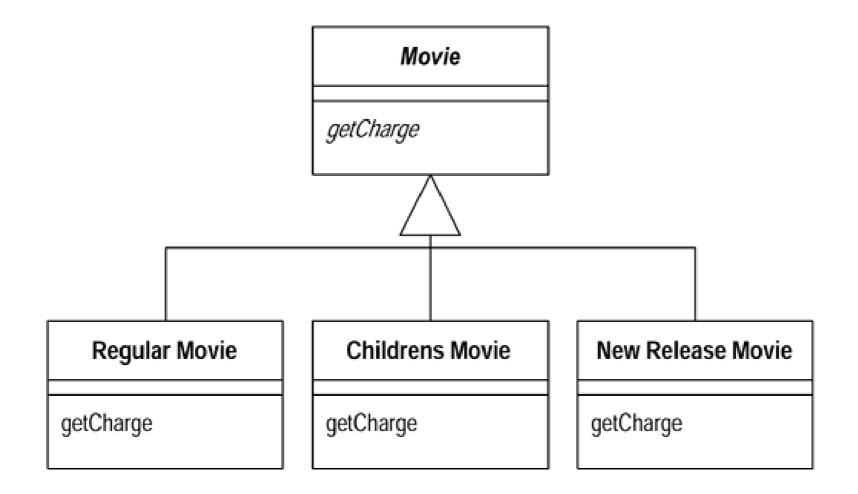
```
class Movie(object):
 def get_charge(self, days_rented):
     result = 0.0
     if self.price code == Movie.REGULAR:
         result += 2.0
         if days rented > 2:
             result += (days rented -2) * 1.5
     elif self.price code == Movie.NEW RELEASE:
         result += days rented * 3
     elif self.price code == Movie.CHILDRENS:
         result += 1.5
         if days rented > 3:
             result += (days rented -3) * 1.5
     return result
 def get frequent renter points(self, days rented):
     if self.price code == Movie.NEW RELEASE and days rented > 1:
         return 2
     else:
         return 1
class Rental(object):
. . .
def get charge(self):
    return self.movie.get charge(self.days rented)
def get frequent renter points(self):
    return self.movie.get frequent renter points(self.days rented)
```

This code still be ugly (why?)

```
def get_charge(self, days_rented):
    result = 0.0
    # determine amount for each line
    if self.price_code == Movie.REGULAR:
        result += 2.0
        if days_rented > 2:
            result += (days_rented - 2) * 1.5
    elif self.price_code == Movie.NEW_RELEASE:
        result += days_rented * 3
    elif self.price_code == Movie.CHILDRENS:
        result += 1.5
        if days_rented > 3:
            result += (days_rented - 3) * 1.5
    return result
```

This code performs 3 tasks depending on the type of the movie.

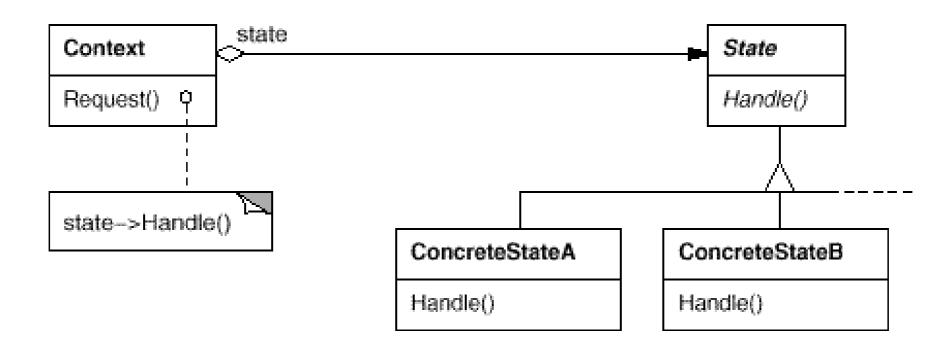
Can we have subclasses of movie?



What happens if the movie type is changed? Because an object cannot change its class during its lifetime.

Design Patterns : State

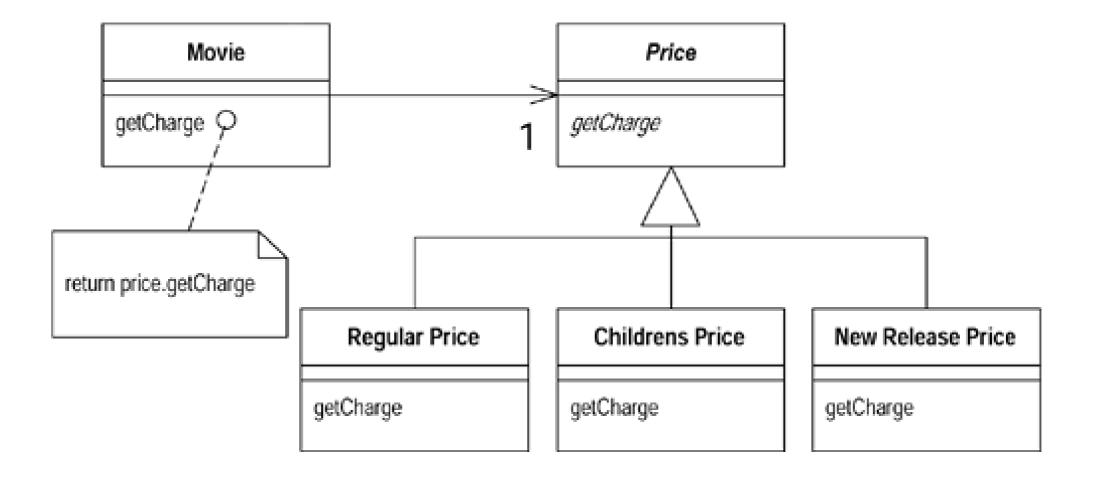
- Allow an object to alter its behavior when its internal state changes. The object will appear to change its class.
- Structure



What is the different between State and Strategy?

- <u>http://www.c-sharpcorner.com/UploadFile/rmcochran/</u> <u>strategy_state01172007114905AM/strategy_state.aspx</u>
- <u>http://stackoverflow.com/questions/1658192/what-is-the-difference-between-strategy-design-pattern-and-state-design-pattern</u>

Using the State pattern on movie



Ninth step : Replace Type Code with State

- Self Encapsulate Field
- Movie Method
- Replace Conditional with Polymorphism

Self Encapsulate Field

class Price(object):
 def get_price_code(self):
 raise NotImplementedError

class ChildrensPrice(Price):
 def get_price_code(self):
 return Movie.CHILDRENS

class RegularPrice(Price):
 def get_price_code(self):
 return Movie.REGULAR

class NewReleasePrice(Price):
 def get_price_code(self):
 return Movie.NEW_RELEASE

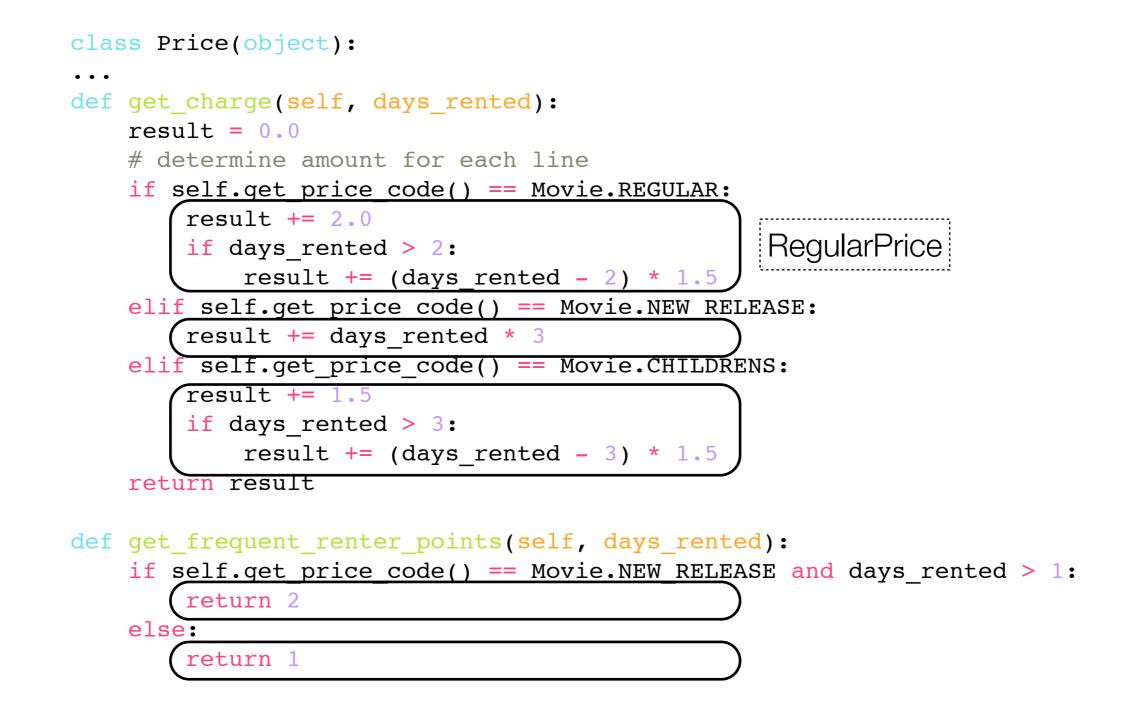
Move Method

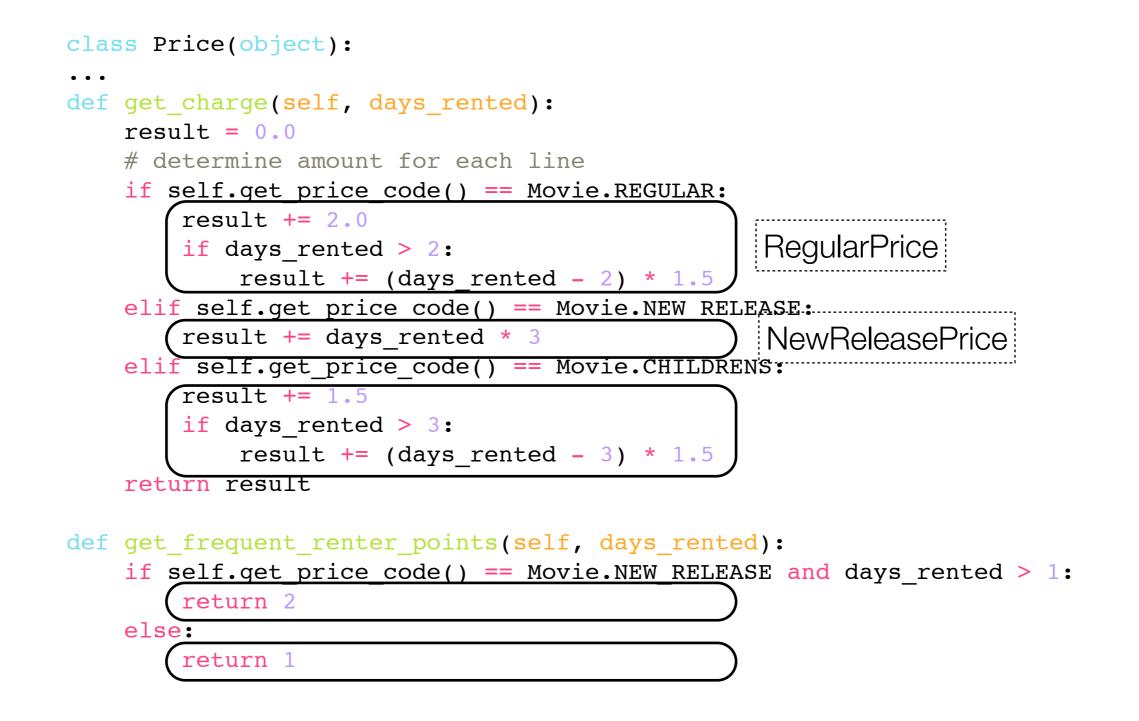
```
class Price(object):
. . .
def get_charge(self, days_rented):
   result = 0.0
   # determine amount for each line
    if self.get_price_code() == Movie.REGULAR:
        result += 2.0
        if days rented > 2:
            result += (days rented - 2) * 1.5
    elif self.get price code() == Movie.NEW RELEASE:
        result += days_rented * 3
    elif self.get price code() == Movie.CHILDRENS:
        result += 1.5
        if days rented > 3:
            result += (days rented - 3) * 1.5
   return result
def get_frequent_renter_points(self, days_rented):
    if self.get_price_code() == Movie.NEW_RELEASE and days_rented > 1:
        return 2
    else:
        return 1
```

Move Method

```
class Movie(object):
. . .
@property
def price(self):
    return self. price
@price.setter
def price(self, price):
    self. price = price
@property
def price code(self):
    return self. price.get price code()
@price code.setter
def price_code(self, price_code):
    if price code == Movie.REGULAR:
        self.price = RegularPrice()
    elif price code == Movie.NEW RELEASE:
        self.price = NewReleasePrice()
    elif price code == Movie.CHILDRENS:
        self.price = ChildrensPrice()
def get charge(self, days rented):
    return self.price.get charge(days rented)
def get frequent renter points(self, days rented):
    return self.price.get frequent renter points(days rented)
```

```
class Price(object):
. . .
def get charge(self, days rented):
   result = 0.0
   # determine amount for each line
    if self.get price code() == Movie.REGULAR:
        result += 2.0
        if days rented > 2:
            result += (days_rented - 2) * 1.5
    elif self.get price code() == Movie.NEW RELEASE:
        result += days rented * 3
    elif self.get price code() == Movie.CHILDRENS:
        result += 1.5
        if days rented > 3:
            result += (days rented -3) * 1.5
    return result
def get_frequent_renter_points(self, days_rented):
    if self.get price code() == Movie.NEW RELEASE and days rented > 1:
        return 2
    else:
        return 1
```





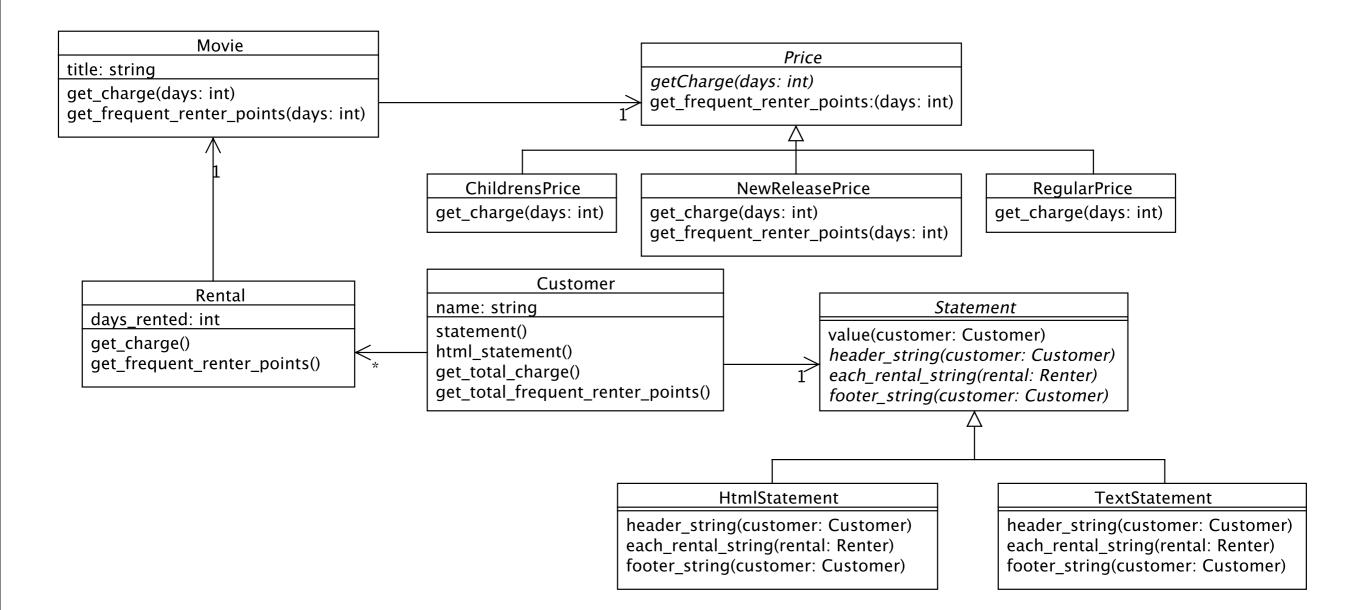




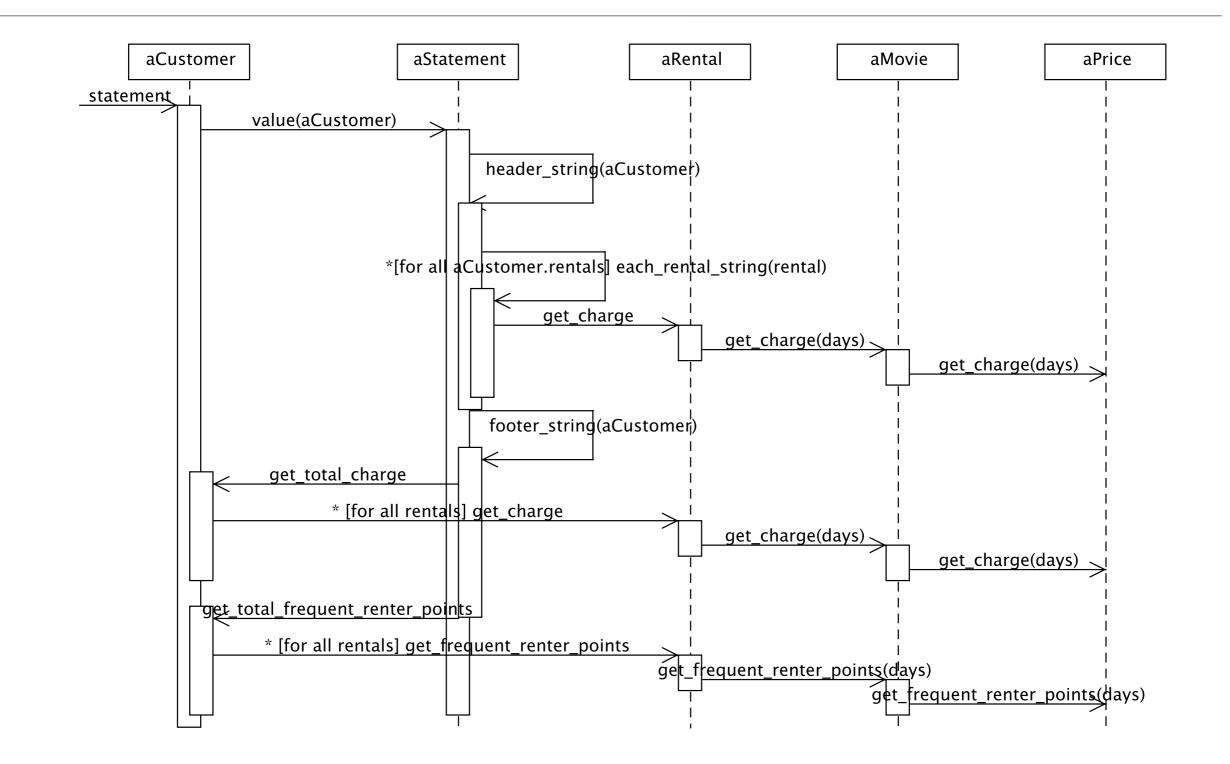


```
class Price(object):
    def get_charge(self, days_rented):
        raise NotImplementedError
    def get frequent renter points(self, days rented):
        return 1
class ChildrensPrice(Price):
    . . .
    def get charge(self, days rented):
        return 1.5+(days rented-3)*1.5 if days_rented > 3 else 1.5
class RegularPrice(Price):
    def get charge(self, days rented):
        return 2.0+(days rented-2)*1.5 if days rented > 2 else 2.0
class NewReleasePrice(Price):
    def get_frequent_renter_points(self, days_rented):
        return 2 if days rented > 1 else 1
    def get charge(self, days rented):
        return days rented * 3
```

Class Diagram



Sequence Diagram



Principles in Refactoring

The Two Hats

- When you use refactoring to develop software, you divide your time between two distinct activities:
 - adding function
 - When you add function, you shouldn't be changing existing code; you are just adding new capabilities.
 - refactoring
 - When you refactor, you make a point of not adding function; you only restructure the code.

Why Should You Refactor?

- Improves the design of software
- makes software easier to understand
- helps you find bugs
- helps you program faster

When Should You Refactor? : The Rule of Three

- Refactor when you add function
 - if you have a hard time to add a new function, you need refactoring
- Refactor when you need to fix a bug
 - if you do get a bug report, you need refactoring because the was not clear enough for you to see there were a bug
- Refactor as you do a code review
 - refactoring also helps the code review have more concrete rseults

When Shouldn't You Refactor?

- When you should rewrite from scratch instead
- When you closed to a deadline
 - unfinished refactoring as going into debt