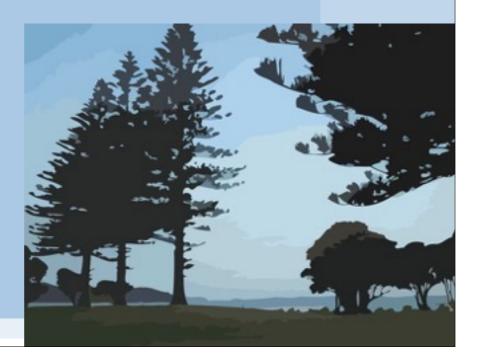


UNIVERSITÄT BERN

# 11. Program Transformation

Oscar Nierstrasz



# Roadmap



- > Program Transformation
- > Refactoring
- > Aspect-Oriented Programming

### Links

#### > Program Transformation:

- http://swerl.tudelft.nl/bin/view/Pt
- http://www.program-transformation.org/

#### > Stratego:

- http://strategoxt.org/
- > TXL:
  - http://www.txl.ca/

#### > Refactoring:

- --- http://www.ibm.com/developerworks/library/os-ecref/
- http://recoder.sourceforge.net/wiki/
- http://www.refactory.com/RefactoringBrowser/

#### > AOP:

— http://www.eclipse.org/aspectj/

## Roadmap



- > Program Transformation
  - -Introduction
  - -Stratego/XT
  - -TXL
- > Refactoring
- > Aspect-Oriented Programming

Thanks to Eelco Visser and Martin Bravenboer for their kind permission to reuse and adapt selected material from their Program Transformation course. http://swerl.tudelft.nl/bin/view/Pt

## What is "program transformation"?

- > <u>Program Transformation</u> is the process of transforming one program to another.
- > Near synonyms:
  - —Metaprogramming
  - —Generative programming
  - —Program synthesis
  - —Program refinement
  - —Program calculation

### **Applications of program transformation**

#### > Translation

- Migration
- Synthesis
  - Refinement
  - Compilation
- Reverse Engineering
  - Decompilation
  - Architecture Extraction
  - Visualization
- Program Analysis
  - Control flow
  - Data flow

Refinement — transform high-level spec down to an implementation that fulfils requirements

Renovation — reengineering

### **Translation** — compilation

```
fact:subu $sp, $sp, 20
function fact(n : int) : int =
 if n < 1 then 1
                                            $fp, 8($sp)
        else (n * fact(n - 1)) \Rightarrow
                                       addiu $fp, $sp, 20
                                       sw $s2, -8($fp)
                                            $ra, -4($fp)
                                       sw
                                            $a0, 0($fp)
                                       move $s2, $a1
                                      li $t0, 1
                                      bge $s2, $t0, c_0
                                      li $v0, 1
                                      b
                                            d_0
                                   c_0: lw $a0, ($fp)
                                      li $t0, 1
                                      subu $a1, $s2, $t0
                                      jal fact_a_0
                                      nul $v0, $s2, $v0
                                  d_0: lw $s2, -8($fp)
                                       lw $ra, -4($fp)
                                       lw
                                            $fp, 8($sp)
                                       addiu $sp, $sp, 20
                                            $ra
                                       jr
```

http://www.cs.uu.nl/docs/vakken/pt/slides/PT05-ProgramTransformation.pdf

### Translation — migration from procedural to OO

```
type tree = {key: int, children: treelist}
type treelist = {hd: tree, tl: treelist}
function treeSize(t : tree) : int =
   if t = nil then 0 else 1 + listSize(t.children)
function listSize(ts : treelist) =
   if ts = nil then 0 else listSize(t.tl)

Tiger

class Tree {
   Int key;
   TreeList children;
   public Int size() {
      return 1 + children.size
   }
   }
   class TreeList { ... }

http://www.cs.uu.n/docs/vakken/pt/slides/PT05-ProgramTransformation.pdf
```

## Rephrasing — desugaring regular expressions

http://www.cs.uu.nl/docs/vakken/pt/slides/PT05-ProgramTransformation.pdf

### **Rephrasing — partial evaluation**

```
function power(x : int, n : int) : int =
  if n = 0 then 1
  else if even(n) then square(power(x, n/2))
  else (x * power(x, n - 1))
```

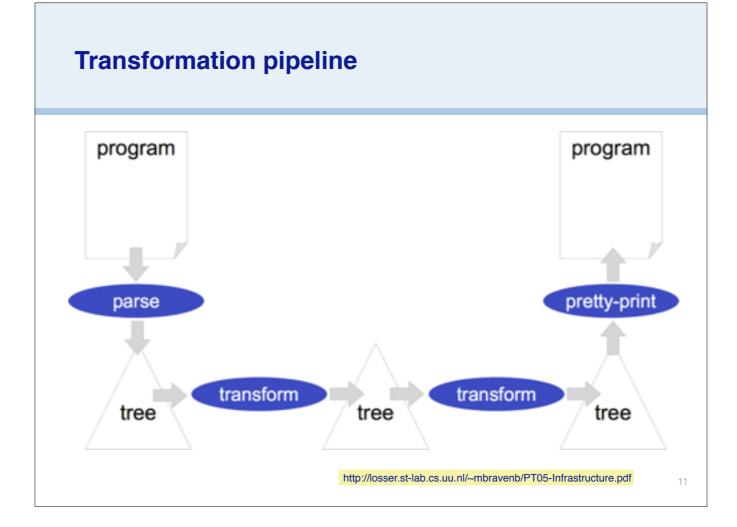
Tiger

$$\Downarrow$$
 n = 5

Tige

```
function power5(x : int) : int =
  x * square(square(x))
```

http://www.cs.uu.nl/docs/vakken/pt/slides/PT05-ProgramTransformation.pdf



This general scheme applies to Stratego, TXL and various other systems. Transformation systems and languages may support or automate different parts of this pipeline.

If the source language is fixed, then a fixed parser and pretty-printer may be used.

If the source and target languages are arbitrary, then there should be support to specify grammars and automatically generate parsers and pretty-printers.

# Roadmap



- > Program Transformation
  - Introduction
  - -Stratego/XT
  - $-\mathsf{TXL}$
- > Refactoring
- > Aspect-Oriented Programming

### Stratego/XT

## > Stratego

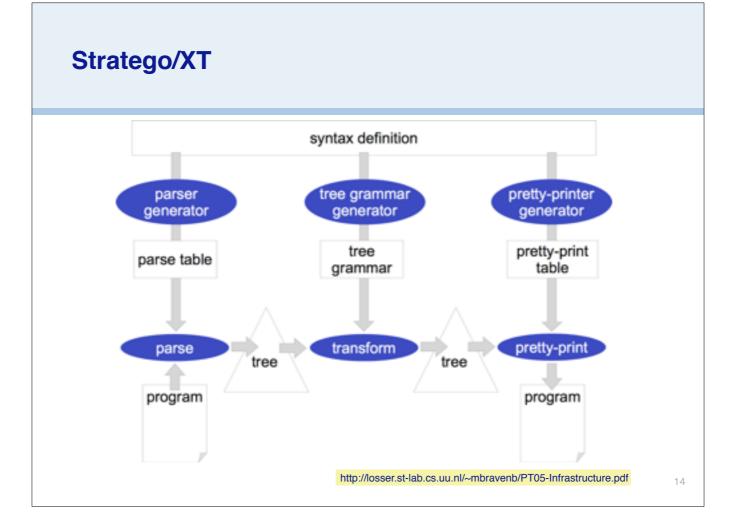
- —A language for specifying program transformations
  - term rewriting rules
  - programmable rewriting strategies
  - pattern-matching against syntax of object language
  - context-sensitive transformations

#### > **XT**

- —A collection of transformation tools
  - parser and pretty printer generators
  - grammar engineering tools

http://strategoxt.org/





Parser and basic pretty-printer 100% generated.

Language specific support for transformations generated.

# **Parsing**

Rules translate terms to terms

Stratego parses any context-free language using Scannerless Generalized LR Parsing

```
module Exp
                                  File: Exp.sdf
exports
 context-free start-symbols Exp
  sorts Id IntConst Exp
 lexical syntax
   [\ \t\n] -> LAYOUT
   [a-zA-Z]+ \rightarrow Id
   [0-9]+ -> IntConst
  context-free syntax
            -> Exp {cons("Var")}
   IntConst -> Exp {cons("Int")}
   "(" Exp ")" -> Exp {bracket}
   Exp "*" Exp -> Exp {left, cons("Mul")}
   Exp "/" Exp -> Exp {left, cons("Div")}
   Exp "%" Exp -> Exp {left, cons("Mod")}
   Exp "+" Exp -> Exp {left, cons("Plus")}
   Exp "-" Exp -> Exp {left, cons("Minus")}
  context-free priorities
     Exp "*" Exp -> Exp
     Exp "/" Exp -> Exp
     Exp "%" Exp -> Exp
     Exp "+" Exp -> Exp
     Exp "-" Exp -> Exp
```

See the Makefile for the steps needed to run this.

GLR parsing essentially does a parallel, breadth-first LR parse to handle ambiguity.

http://en.wikipedia.org/wiki/GLR\_parser

## **Testing**

### **Running tests**

```
pack-sdf -i Exp.sdf -o Exp.def
including ./Exp.sdf

sdf2table -i Exp.def -o Exp.tbl -m Exp
SdfChecker:error: Main module not defined
--- Main

parse-unit -i Exp.testsuite -p Exp.tbl

executing testsuite Exp with 1 tests

* OK : test 1 (eg1 parse)

results testsuite Exp
successes : 1
failures : 0
```

### **Interpretation example**

```
module ExpEval

imports libstratego-lib
imports Exp

rules
   convert : Int(x) -> <string-to-int>(x)
   eval : Plus(m,n) -> <add>(m,n)
   eval : Minus(m,n) -> <subt>(m,n)
   eval : Mul(m,n) -> <mul>(m,n)
   eval : Div(m,n) -> <div>(m,n)
   eval : Mod(m,n) -> <mod>(m,n)
File: ultimate-question.txt

1 + 2 * (3 + 4) * 3 - 1

strategies
main = io-wrap(innermost(convert <+ eval))</pre>
```

Stratego separates the specification of <u>rules</u> (transformations) from <u>strategies</u> (traversals). In principle, both are reusable.

## **Strategies**

A <u>strategy</u> determines how a set of rewrite rules will be used to traverse and transform a term.

- innermost
- top down
- bottom up
- repeat
- ...

-1

### **Running the transformation**

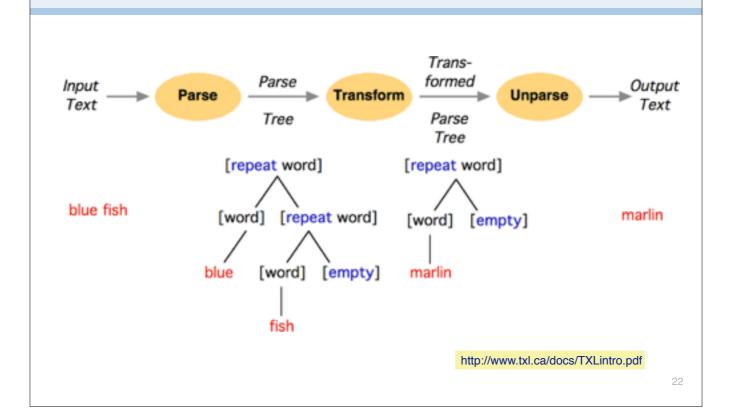
```
sdf2rtg -i Exp.def -o Exp.rtg -m Exp
                                              Generate regular tree grammar
SdfChecker:error: Main module not defined
                                 Generate signature
rtg2sig -i Exp.rtg -o Exp.str
strc -i ExpEval.str -la stratego-lib
                                            Compile to C
[ strc | info ] Compiling 'ExpEval.str'
[ strc | info ] Front-end succeeded
                                           : [user/system] = [0.56s/0.05s]
[ strc | info ] Optimization succeeded -0 2 : [user/system] = [0.00s/0.00s]
[ strc | info ] Back-end succeeded
                                          : [user/system] = [0.16s/0.01s]
 gcc -I /usr/local/strategoxt/include -I /usr/local/strategoxt/include -I /usr/local/strategoxt/
include -Wall -Wno-unused-label -Wno-unused-variable -Wno-unused-function -Wno-unused-parameter -
DSIZEOF VOID P=4 -DSIZEOF LONG=4 -DSIZEOF INT=4 -c ExpEval.c -fno-common -DPIC -o .libs/ExpEval.o
gcc -I /usr/local/strategoxt/include -I /usr/local/strategoxt/include -I /usr/local/strategoxt/
include -Wall -Wno-unused-label -Wno-unused-variable -Wno-unused-function -Wno-unused-parameter -
DSIZEOF VOID P=4 -DSIZEOF LONG=4 -DSIZEOF INT=4 -c ExpEval.c -o ExpEval.o >/dev/null 2>&1
gcc .libs/ExpEval.o -o ExpEval -bind_at_load -L/usr/local/strategoxt/lib /usr/local/strategoxt/lib/
libstratego-lib.dylib /usr/local/strategoxt/lib/libstratego-lib-native.dylib /usr/local/strategoxt/
lib/libstratego-runtime.dylib -lm /usr/local/strategoxt/lib/libATerm.dylib
[ strc | info ] C compilation succeeded
                                         : [user/system] = [0.31s/0.36s]
[ strc | info ] Compilation succeeded
                                           : [user/system] = [1.03s/0.42s]
sglri -p Exp.tbl -i ultimate-question.txt | ./ExpEval
                                                        Parse and transform
```

# Roadmap



- > Program Transformation
  - Introduction
  - -Stratego/XT
  - $-\mathsf{TXL}$
- > Refactoring
- > Aspect-Oriented Programming

## The TXL paradigm: parse, transform, unparse



## **TXL** programs

Base grammar

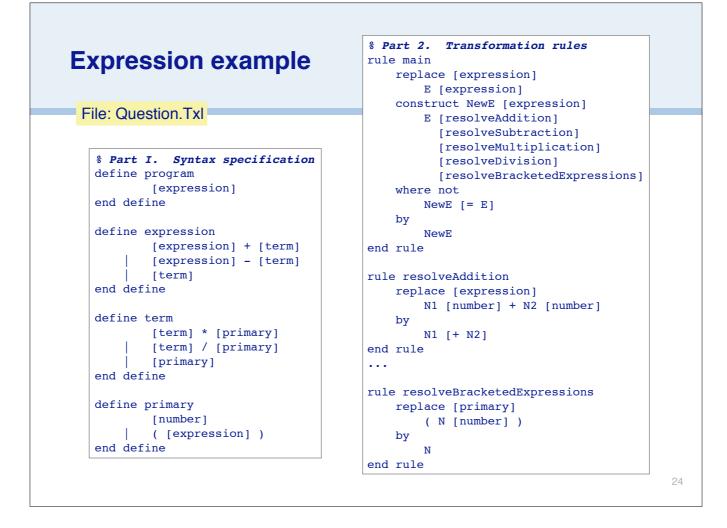
defines tokens and non-terminals

Grammar overrides

extend and modify types from grammar

Transformation rules

rooted set of rules and functions



NB: TXL reverses the usual BNF convention and puts non-terminals in square brackets while interpreting everything else (except special chars) as terminals.

The default lexical scanner can be modified, but is usually fine for first experiments.

## **Running the example**

#### File: Ultimate.Question

```
1 + 2 * (3 + 4) * 3 - 1
```

```
txl Ultimate.Question
TXL v10.5d (1.7.08) (c)1988-2008 Queen's University at Kingston
Compiling Question.Txl ...
Parsing Ultimate.Question ...
Transforming ...
42
```

### Example: TIL — a tiny imperative language

```
// Find all factors of a given input number
var n;
write "Input n please"; File: factors.til
read n;
write "The factors of n are";
var f;
f := 2;
while n != 1 do
    while (n / f) * f = n do
        write f;
        n := n / f;
    end
    f := f + 1;
end
```

http://www.program-transformation.org/Sts/TILChairmarks

#### **TIL Grammar**

```
% Keywords of TIL
keys
    var if then else while
    do for read write
end keys

% Compound tokens
compounds
    := !=
end compounds

% Commenting convention
comments
    //
end comments
```

All TXL parsers are also prettyprinters if the grammar includes formatting cues

```
define program
    [statement*]
end define
define statement
        [declaration]
        [assignment_statement]
        [if_statement]
        [while statement]
        [for_statement]
        [read_statement]
        [write_statement]
end define
% Untyped variables
define declaration
    'var [id] ;
                                  [NL]
end define
define assignment statement
    [id] := [expression];
                                  [NL]
end define
define if statement
    'if [expression] 'then
                                  [IN][NL]
        [statement*]
                                  [EX]
    [opt else statement]
    'end
                                  [NL]
end define
. . .
```

### **Pretty-printing TIL**

```
include "TIL.Grm"
function main
match [program]
_ [program]
end function
```

```
write "Input n please";
read n;
write "The factors of n are";
var f;
f := 2;
while n != 1 do
    while (n / f) * f = n do
    write f;
    n := n / f;
end
f := f + 1;
end
```

### **Generating statistics**

```
include "TIL.Grm"
                                             File: TILstats.Txl
function main
   replace [program]
        Program [program]
    % Count each kind of statement we're interested in
    % by extracting all of each kind from the program
   construct Statements [statement*]
                                                        Total: 11
        _ [^ Program]
                                                        Declarations: 2
                                                        Assignments: 3
   construct StatementCount [number]
                                                        Ifs: 0
        _ [length Statements] [putp "Total: %"]
                                                        Whiles: 2
                                                        Fors: 0
    construct Declarations [declaration*]
                                                        Reads: 1
        _ [^ Program]
                                                        Writes: 3
   construct DeclarationsCount [number]
        _ [length Declarations] [putp "Declarations: %"]
   by
        % nothing
end function
```

### **Tracing**

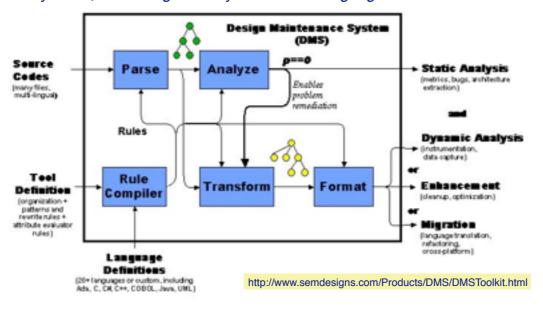
```
include "TIL.Grm"
                       File: TILtrace.Txl
redefine statement
    [traced statement]
end redefine
define traced statement
   [statement] [attr 'TRACED]
end define
                                     write "Trace: var n;";
rule main
                                     var n;
                                     write "Trace: write \"Input n please\";";
replace [repeat statement]
                                     write "Input n please";
        S [statement]
                                     write "Trace: read n;";
        Rest [repeat statement]
                                    read n;
    by
        'write QuotedS;
                            'TRACED
                            'TRACED
        Rest
end rule
                                                                           30
```

# **TXL vs Stratego**

Stratego	TXL
Scannerless GLR parsing	Agile parsing (top-down + bottom-up)
Reusable, generic traversal strategies	Fixed traversals
Separates rewrite rules from traversal strategies	Traversals part of rewrite rules

### **Commercial systems**

"The DMS Software Reengineering Toolkit is a set of tools for automating customized source program analysis, modification or translation or generation of software systems, containing arbitrary mixtures of languages."



32

See also http://www.semdesigns.com/Products/DMS/DMSComparison.html for a comparison to other approaches

## Roadmap



- > Program Transformation
- > Refactoring
  - Refactoring Engine and Code Critics
  - Eclipse refactoring plugins
- > Aspect-Oriented Programming

## What is Refactoring?

> The process of *changing a software system* in such a way that it *does not alter the external behaviour* of the code, yet *improves its internal structure*.

- Fowler, et al., Refactoring, 1999.

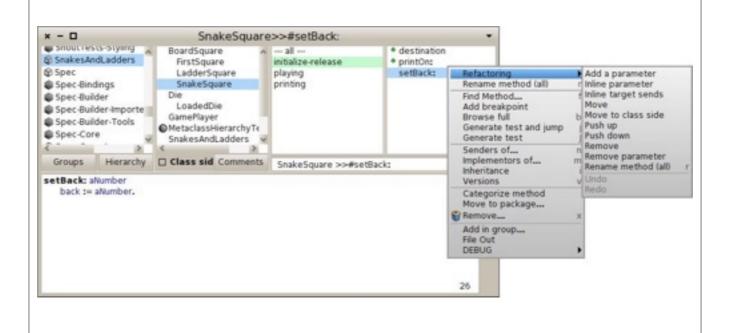
### **Rename Method** — manual steps

- > Do it yourself approach:
  - —Check that no method with the new name already exists in any subclass or superclass.
  - —Browse all the implementers (method definitions)
  - —Browse all the senders (method invocations)
  - —Edit and rename all implementers
  - -Edit and rename all senders
  - —Remove all implementers
  - —Test
- > Automated refactoring is better!

#### **Rename Method**

- > Rename Method (method, new name)
- > Preconditions
  - —No method with the new name already exists in any subclass or superclass.
  - No methods with same signature as method outside the inheritance hierarchy of method
- > PostConditions
  - -method has new name
  - —relevant methods in the inheritance hierarchy have new name
  - —invocations of changed method are updated to new name
- > Other Considerations
  - —Typed/Dynamically Typed Languages => Scope of the renaming

## **The Refactoring Browser**



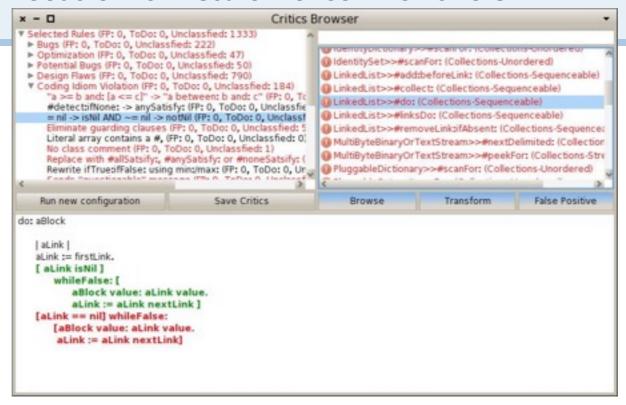
## **Typical Refactorings**

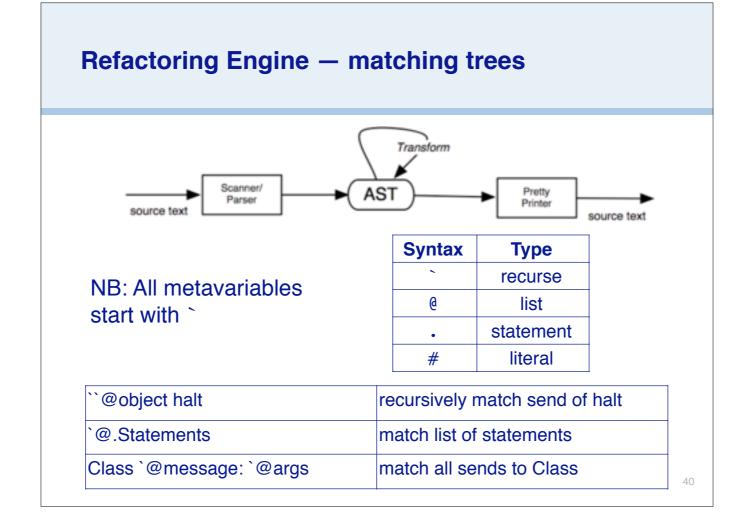
Class Refactorings	Method Refactorings	Attribute Refactorings
add (sub)class to hierarchy	add method to class	add variable to class
rename class	rename method	rename variable
remove class	remove method	remove variable
	push method down	push variable down
	push method up	pull variable up
	add parameter to method	create accessors
	move method to component	abstract variable
	extract code in new method	

Bill Opdyke, "Refactoring Object-Oriented Frameworks," Ph.D. thesis, University of Illinois, 1992.

Don Roberts, "Practical Analysis for Refactoring," Ph.D. thesis, University of Illinois, 1999.

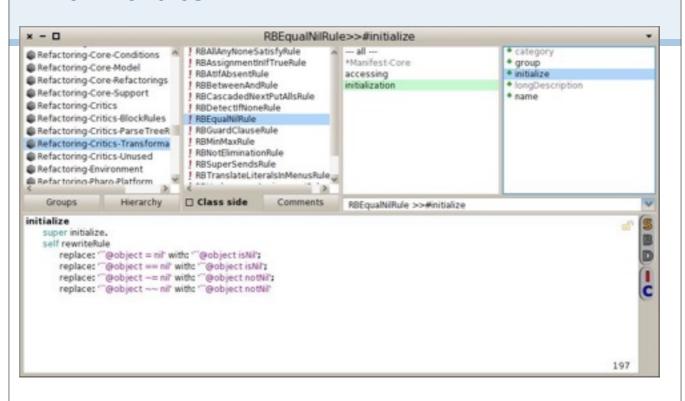
#### **Code Critic** — search for common errors





The first ` is for all meta-variables.

#### **Rewrite rules**



# Roadmap



- > Program Transformation
- > Refactoring
  - Refactoring Engine and Code Critics
  - Eclipse refactoring plugins
- > Aspect-Oriented Programming

### A workbench action delegate

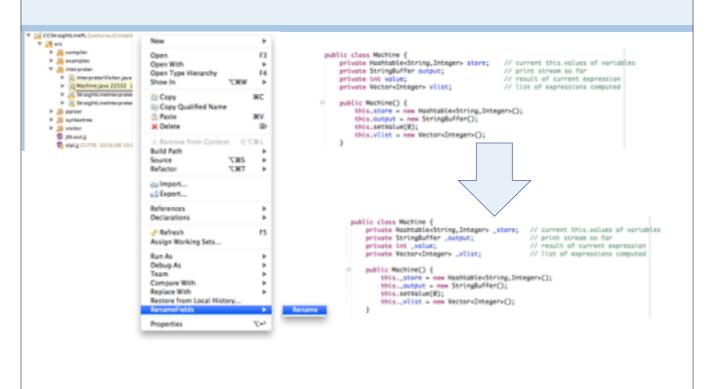
When the workbench action proxy is triggered by the user, it delegates to an instance of this class.

http://help.eclipse.org/ganymede/index.jsp?topic=/org.eclipse.jdt.doc.isv/guide/jdt\_api\_manip.htm

### A field renaming visitor

The visitor simply implements the visit method for field declarations and accesses, and prepends an underscore.

## **Renaming fields**

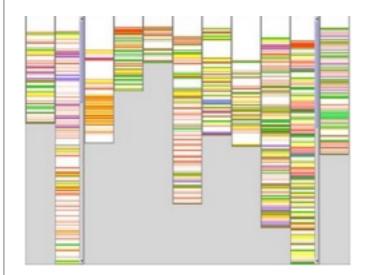


# Roadmap



- > Program Transformation
- > Refactoring
- > Aspect-Oriented Programming

## **Problem: cross-cutting concerns**



Certain features (like logging, persistence and security), cannot usually be encapsulated as classes. They *cross-cut* code of the system.

"Identifying Cross-Cutting Concerns in Embedded C Code", Bruntink, van Deursen, Tourwé

## **Aspect-Oriented Programming**

AOP improves modularity by supporting the separation of cross-cutting concerns.

An *aspect* packages cross-cutting concerns



A *pointcut* specifies a set of *join points* in the target system to be affected

Weaving is the process of applying the aspect to the target system

## Canonical example — logging

```
package tjp;
public class Demo {
static Demo d;
public static void main(String[] args){
 new Demo().go();
void go(){
 d = new Demo();
 d.foo(1,d);
 System.out.println(d.bar(new Integer(3)));
void foo(int i, Object o){
 System.out.println("Demo.foo(" + i + ", " + o + ")\n");
String bar (Integer j){
 System.out.println("Demo.bar(" + j + ")\n");
 return "Demo.bar(" + j + ")";
                                                         Demo.foo(1, tjp.Demo@939b78e)
                                                         Demo.bar(3)
                                                         Demo.bar(3)
http://www.eclipse.org/aspectj/downloads.php
```

## A logging aspect

```
Intercept execution within control flow of Demo.go()

Identify all methods within Demo

aspect GetInfo {

pointcut goCut(): cflow(this(Demo) && execution(void go()));

pointcut demoExecs(): within(Demo) && execution(* *(..));

Object around(): demoExecs() && !execution(* go()) && goCut() {

...
}

Wrap all methods except Demo.go()
```

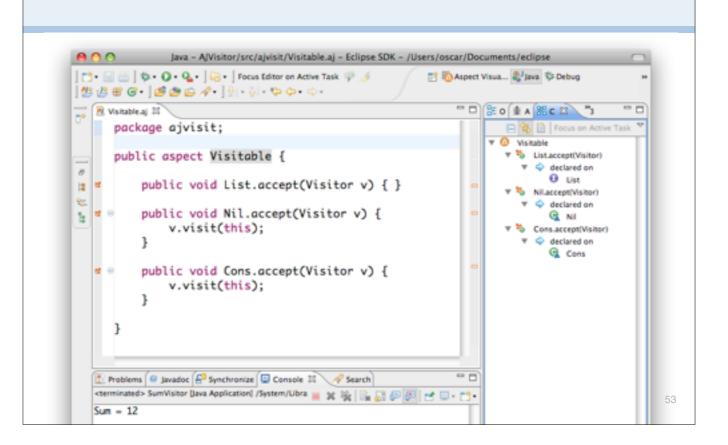
### A logging aspect

```
aspect GetInfo {
Object around(): demoExecs() && !execution(* go()) && goCut() {
 println("Intercepted message: " +
   thisJoinPointStaticPart.getSignature().getName());
 println("in class: " +
 thisJoinPointStaticPart.getSigna in class: tjp.Demo
                                   Intercepted message: foo
 printParameters(thisJoinPoint);
                                   Arguments:
 println("Running original method:
                                     0. i : int = 1
 Object result = proceed();
                                     1. o : java.lang.Object = tjp.Demo@c0b76fa
 println(" result: " + result );
                                   Running original method:
 return result;
}
                                    Demo.foo(1, tjp.Demo@c0b76fa)
. . .
                                     result: null
                                    Intercepted message: bar
                                    in class: tjp.Demo
                                    Arguments:
                                     0. j : java.lang.Integer = 3
                                    Running original method:
                                    Demo.bar(3)
                                      result: Demo.bar(3)
                                    Demo.bar(3)
```

#### Making classes visitable with aspects

```
public class SumVisitor implements Visitor {
int sum = 0;
public void visit(Nil 1) { }
                                    We want to write this
public void visit(Cons 1) {
 sum = sum + 1.head;
                                                 public interface List {}
 1.tail.accept(this);
                                                 public class Nil implements List {}
                                                 public class Cons implements List {
public static void main(String[] args) {
                                                  int head;
 List l = new Cons(5, new Cons(4,
                                                  List tail;
                                                  Cons(int head, List tail) {
      new Cons(3, new Nil()));
 SumVisitor sv = new SumVisitor();
                                                  this.head = head;
                                                  this.tail = tail;
 1.accept(sv);
 System.out.println("Sum = " + sv.sum);
                                                        But we are stuck with this ...
public interface Visitor {
void visit(Nil 1);
void visit(Cons 1);
```

## **AspectJ**



### With aspects, who needs visitors?

```
public class SumList {
  public static void main(String[] args) {
   List l = new Cons(5, new Cons(4, new Cons(3, new Nil())));
   System.out.println("Sum = " + 1.sum());
  }
}
```

The missing method is just an aspect

```
public aspect Summable {
  public int List.sum() {
    return 0;
  }
  public int Nil.sum() {
    return 0;
  }
  public int Cons.sum() {
    return head + tail.sum();
  }
}
```

54

Dunno why List.sum() needs a body – it should just be an interface signature.

#### What you should know!

- What are typical program transformations?
- What is the typical architecture of a PT system?
- What is the role of term rewriting in PT systems?
- How does TXL differ from Stratego/XT?
- How does the Refactoring Engine use metavariables to encode rewrite rules?
- Why can't aspects be encapsulated as classes?
- What is the difference between a pointcut and a join point?

#### Can you answer these questions?

- How does program transformation differ from metaprogramming?
- In what way is optimization a form of PT?
- What special care should be taken when pretty-printing a transformed program?
- How would you encode typical refactorings like "push method up" using a PT system like TXL?
- Mow could you use a PT system to implement AOP?



#### Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)

#### You are free to:

**Share** — copy and redistribute the material in any medium or format

**Adapt** — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

#### **Under the following terms:**



**Attribution** — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.



**ShareAlike** — If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.

No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

http://creativecommons.org/licenses/by-sa/4.0/