

MOTIVATING EXAMPLE

ITERATIVE DIFF (IDIFF)

EVALUATION

CONCLUSIONS





The Brazilian Symposium on Software Engineering (SBES)



Towards a Difference Detection Algorithm aware of Refactoring-related Changes

Fernanda Silva¹ Eral

Eraldo Borel¹

Evandro Lopes²

Leonardo Murta¹

Congresso Brasileiro de Software: Teoria e Prática 28 de Setembro a O3 de Outubro de 2014

Maceió - Alagoas

¹Computing Institute Fluminense Federal University (UFF) Niterói, Rio de Janeiro, Brazil e-mail: {ffloriano,leomurta}@ic.uff.br eraldoborel@id.uff.br ² Department of Statistics Fluminense Federal University (UFF) Niterói, Rio de Janeiro, Brazil e-mail: evandro_dalbem@id.uff.br







MOTIVATING EXAMPLE

Iterative Diff (IDIFF)

EVALUATION

CONCLUSIONS



MOTIVATION

- Refactorings are a usual practice during software development
- At the physical level, refactorings imply file renames and moves and code snippets moves across files
- However, current generic diff tools detect lines additions and deletions within files



Universidade Federal Fluminense

INTRODUCTION

MOTIVATING EXAMPLE

ITERATIVE DIFF (IDIFF)

EVALUATION







Order id: Integer POS registers +getId(): Integer +pay(id: Integer 0..* +register(id:Integer) 1..* 1 Employee **CheckPayment** Payment login : String CashPayment password : String parcelsNumber : Integer -amount : Currency +getLogin(): String +getPassword(): String CardPayment +setLogin(login : String) +setPassword(password : String) parcelsNumber : Integer

Refactoring applied over the base version:

Move Method Rename Method Encapsulate Field







MOTIVATING **EXAMPLE**

ITERATIVE DIFF (IDIFF)

EVALUATION

CONCLUSIONS



POS (point-of-sale) system, implemented in Java

Visualizar Unir Ferramentas Plugins Janela Ajuda	
e.bd CardPayment.bd - Store.bd Establishment.java - Store.java POS.java - POS.java	
H:\Project_1\POS.java	H:\Project_2\POS.java
import java.sql.Connection;	<pre>import java.sql.Connection;</pre>
import java.sql.DriverManager;	<pre>import java.sql.DriverManager;</pre>
import java.sql.ResultSet;	<pre>import java.sql.ResultSet;</pre>
import java.sql.Statement;	<pre>import java.sql.Statement;</pre>
public class POS {	public class POS (
<pre>public void pay(Integer id) throws Exception { Class.forName("org.hsqldb.jdbcDriver"); String url = "jdbc:hsqldb:mem:data/idiff"; Connection conn = DriverManager.getConnection(url, ". Statement stmt = conn.createStatement(); stmt.executeUpdate("update POS set idPago='S' where : stmt.close(); conn.close(); }</pre>	
<pre>public void register(Integer id) { Class.forName("org.hsqldb.jdbcDriver"); String url = "jdbc:hsqldb:mem:data/idiff"; Connection conn = DriverManager.getConnection(url, "); </pre>	<pre>public void register(Integer id) { Class.forName("org.hsqldb.jdbcDriver"); String url = "jdbc:hsqldb:mem:data/idiff"; Connection conn = DriverManager.getConnection(url)</pre>
Visualizar Unir Ferramentas Plugins Janela Ajuda	-
2 == 요 장 장 장 + + + + + + = 2 장 집 2 입	
jão dos Pastas Sem título a esquerda - Sem título a direita Order java - Order java	
* H:\Project_1\Order.java	H:\Project_2\Order.java
<pre>import java.sql.ResultSet;</pre>	<pre>import java.sql.ResultSet;</pre>
<pre>import java.sql.Statement;</pre>	<pre>import java.sql.Statement;</pre>
public class Order {	public class Order {
	<pre>public void pay(Integer 1d) throws Exception { Class.forName("org.hsgldb.jdbcDriver"); String url = "idbc:hsgldb:mem:data/idiff";</pre>

public Integer getId() { return this.id;

private Integer id;

Statement stmt = conn.createStatement(); stmt.executeUpdate("update POS set idPago='S' wher stmt.close(); conn.close(); public Integer getId() { return this.id;

Connection conn = DriverManager.getConnection(url,

private Integer id;

Slide 4



MOTIVATING EXAMPLE

Iterative Diff (IDIFF)

EVALUATION

CONCLUSIONS



GOAL

Conceive a generic diff algorithm that precisely identify refactoring-related changes













Slide 9







Slide 11









































a 🌗 Directory 1	4 퉲 Directory 2
⊿ 퉲 payment	a 퉲 payment
CardPayment	CardPayment
CheckPayment	CheckPayment
Establishment	Store
Order	Order
POS	POS















4 퉲 Directory 1	a 🌗 Directory 2
🛯 퉲 payment	a 퉲 payment
CardPayment	CardPayment
CheckPayment	CheckPayment
Establishment	Store
Order	Order
POS	POS







$$Similarity = \frac{2 \times LCS(F_1, F_2)}{Size(F_1) + Size(F_2)}$$





$$Similarity = \frac{2 \times LCS(F_1, F_2)}{Size(F_1) + Size(F_2)}$$

POS 0% 0%





Hungarian Algorithm







Slide 28





Slide 29

C:\ Directory 1\payment\CardPayment.java

```
import java.util.List;
import java.util.Map;
public class CardPayment extends Payment {
    private Integer installments;
```

C:\ Directory 2\payment\CardPayment.java

```
import java.util.Map;
import java.util.List;
public class CardPayment extends Payment {
   public Integer installments;
```

C:\ Directory 1\payment\CardPayment.java

```
import java.util.List;
import java.util.Map;
public class CardPayment extends Payment {
    private Integer installments;
}
```

C:\ Directory 2\payment\CardPayment.java

```
import java.util.Map;
import java.util.List;
public class CardPayment extends Payment {
    public Integer installments;
```

C:\ Directory 1\payment\CardPayment.java

import java.util.List;
 private Integer installments;

C:\Directory 2\payment\CardPayment.java

import java.util.List;
 public Integer installments;

Slide 32

C:\ Directory 1\payment\CardPayment.java

import java.util.List;
 private Integer installments;

C:\Directory 2\payment\CardPayment.java

import java.util.List;
 public Integer installments;

C:\Directory 1\payment\CardPayment.java

private Integer installments;

C:\Directory 2\payment\CardPayment.java

public Integer installments;







C:\Directory 1\payment\CardPayment.java

private
Integer
installments;

C:\Directory 2\payment\CardPayment.java

public
Integer
installments;







C:\ Directory 1\payment\CardPayment.java

```
import java.util.List;
import java.util.Map;
public class CardPayment extends Payment {
    private Integer installments;
}
```

C:\ Directory 2\payment\CardPayment.java

```
import java.util.Map;
import java.util.List;
public class CardPayment extends Payment
    public Integer installments;
```



DDIFF

FDIFF

Overview vs. Pairwise Comparison vs. Multiple Comparison





DDIFF – Overview



FDIFF – Pairwise Comparison, Differences Perspective (comparing the same file)

Directory (Left)	Left (Establishment.java)		Right (Store.java)
C:USers\Fernanda\Doc	<pre>public class Establishment { private String address; private String name; public void getAddress() { return this.address } public void setAddress(String address) { this.address = address; } public void getName() { return this.name; } public void setName(String name) { this.name = name; } }</pre>		<pre>public class Store { private String address; private String name; public void getAddress() { return this address } public void setAddress(String address) { this address = address; } public void getName() { return this name; } public void setName(String name) { this name = name; } }</pre>
Content	Situation	From (Left)	To (Right)
	REMOVED	Line 1	
public class Establishment {	10000		

Slide 42

FDIFF – Pairwise Comparison, Similarity perspective (comparing different files)

Similarity Perspective		Unchanged	Removed	Added	Moved	Moved Highligh
Left (POS.java)	Right (O	rder.java)				
<pre>mport java.sql.Connection; mport java.sql.DriverManager; mport java.sql.ResultSet; mport java.sql.Statement; public class POS {</pre>	mport j; mport j; mport j; public c	ava.sql.Connection; ava.sql.DriverManage ava.sql.ResultSet; ava.sql.Statement; lass Order { public void pay(Cl St Cd St St st st st cd } public int getIdC rei } private int idOrd	er, int idOrder) ti ass.forName ring url = "jdb onnection cod atement stm mt executeU; mt.close(); onn.close(); Order()(turn this.rdOr fer;	hrows Ex a("org.hsc hc:hsqldb; nn = Drive t = conn.c odate("up der;	ception (aldb.jdbc :mem.da erManage reateSta date PD\	Driver"); ta/idiff"; er.getConnectior tement(); / set idPayment



FDIFF – Multiple Comparison

2	File Over	view	
(3	Show All Similarities	•
_	PDV.java	a	
	import ja import ja import ja import ja public c	ava.sql.Connection; ava.sql.DriverManager; ava.sql.ResultSet; ava.sql.Statement; dass POS {	
	pu	<pre>iblic void pay(Integer id) throws Exception { Class.forName("org.hsqldb.jdbcDriver"); String url = "jdbc:hsqldb:mem:data/idiff"; Connection conn = DriverManager.getConnection(url, "idiff", ""); Statement stmt = conn.createStatement(); stmt.executeUpdate("update POS set idPago="S" where id = " + id); stmt.close(); conn.close();</pre>	
	pu	<pre>iblic void reqister(Integer id) { Class.forName("org.hsqldb.jdbcDriver"); String url = "jdbc:hsqldb:mem:data/idiff"; Connection conn = DriverManager.getConnection(url, "idiff", ""); Statement stmt = conn.createStatement(); stmt.executeUpdate("INSERT INTO POS (id, idPay)VALUES(" + id + ", 'N')"); stmt.close(); Similarities found with C:\Project_2\POS.java </pre>	
	}		



MOTIVATING EXAMPLE

ITERATIVE DIFF (IDIFF)

EVALUATION

CONCLUSIONS



PLANNING AND EXECUTION

- Research questions:
 - Which is the best granularity configuration for IDiff?
 - Does IDiff **increase the precision** (correctness) when compared to a generic Diff tool?
 - Does IDiff **increase the recall** (completeness) when compared to a generic Diff tool?
 - In which situations (refactoring types) IDiff
 performs better than a generic Diff tool? Slide 45



MOTIVATING EXAMPLE

ITERATIVE DIFF (IDIFF)

EVALUATION

CONCLUSIONS



PLANNING AND EXECUTION

- Execution of **76 refactorings** from the **Fowler's book**
- **Comparison** of the expected **results** with the results provided by IDiff and WinMerge
 - WinMerge selected as baseline out of a survey with 63 developers





EXAMPLE

(IDIFF)

DIFF

IDIF F

DIFF

DIFF

DIFF

DIFF

EVALUATION PROCESS







MOTIVATING EXAMPLE

ITERATIVE DIFF (IDIFF)

EVALUATION



PRECISION/RECALL ANALYSIS

COMPARISON OF MEAN – WILCOXON TEST

Precision



F-measure







IDIFF

WinMerge

Summary distributions for IDIFF and WinMerge



Universidade Federal Fluminense

INTRODUCTION

MOTIVATING EXAMPLE

ITERATIVE DIFF (IDIFF)

EVALUATION

CONCLUSIONS



MAIN THREATS TO VALIDITY

- Reliability of measurements
- The use of 76 refactorings described in the Fowler's book
- The absence of experience with large projects leaves doubt whether the result will be satisfactory in these scenarios

Slide 49

• WinMerge as baseline



MOTIVATING EXAMPLE

Iterative Diff (IDIFF)

EVALUATION

CONCLUSIONS



CONTRIBUTIONS

- IDiff provides **results with higher precision** if compared to a generic Diff tool, without drastic reduction of recall
- IDiff employs **efficient algorithms** for detecting the optimal content-based similarity amongst files
- **Different visualizations** (pairwise and multiple) and according to **different perspectives** (similarities and differences)
- The use of iterative granularity reduction to conciliate precision and efficiency





MOTIVATING EXAMPLE

Iterative Diff (IDIFF)

EVALUATION

CONCLUSIONS DIFF DIFF F

DIFF

DIFF

FUTURE WORK

- Consider programming language grammars
- Develop a **merge tool** supported by the foundations of this work
- Exploit **parallel processing** of ever-common multi-core computers and GPU
- Combine with refactoring detection techniques by using regular expressions over the diff results to index a refactoring catalog





MOTIVATING EXAMPLE

ITERATIVE DIFF (IDIFF)

EVALUATION

CONCLUSIONS





The Brazilian Symposium on Software Engineering (SBES)



Towards a Difference Detection Algorithm aware of Refactoring-related Changes

Fernanda Silva¹ Eral

Eraldo Borel¹

Evandro Lopes²

Leonardo Murta¹

¹Computing Institute Fluminense Federal University (UFF) Niterói, Rio de Janeiro, Brazil e-mail: {ffloriano,leomurta}@ic.uff.br eraldoborel@id.uff.br ² Department of Statistics Fluminense Federal University (UFF) Niterói, Rio de Janeiro, Brazil e-mail: evandro_dalbem@id.uff.br





Congresso Brasileiro de Software: Teoria e Prática 28 de Setembro a O3 de Outubro de 2014 Maceió - Alagoas



MOTIVANTING EXAMPLE

IDIFF ITERATIVE DIFF ALGORITHM

IDIFF IMPLEMENTATION

EVALUATION

CONCLUSIONS

DIFF DIFF DIFF DIFF DIFF

PLANNING AND EXECUTION

Inline Method

			1				
OPERAÇÃO	Crão						
	Grao	DE PARA					
DELETED	WORD	moreThanFiveLateDeliveries()	10772		28	0	
DELETED	D WORD boolean moreThanFiveLateDeliveries() {		42	42	0		
MOVED	WORD	WORDnumberOfLateDeliveries > 5numberOfLateDeliveries > 5		25	25	0	
DELETED	WORD	WORD }		1	1	0	
DELETED	ED WORD ;		1	1	0		
	3			97	97	0	
		Código					
OPERAÇÃO	~~	Código Fonte		NC	TP	FP	
	Grao	DE	PARA	1			
DELETED	LINE	(moreThanFiveLateDeliveries())		30	28	2	
DELETED	LINE	}	5777.5	1	1	0	
DELETED	LINE	<pre>boolean moreThanFiveLateDeliveries() { return _numberOfLateDeliveries > 5;</pre>	1 310 1	68	42	26	
ADDED	LINE	575	(_numberOfLateDeliveries > 5)	27	0	27	
	3	9	<u>,</u>	126	71	55	

	Precision	Recall	F-measure
IDIFF	1	1	1
WinMerge	0,56	0,73	0,63



MOTIVATING EXAMPLE

ITERATIVE DIFF (IDIFF)

EVALUATION

CONCLUSIONS



Which is **the best granularity** configuration for IDIFF?

	Precision					
Line (0.58)	- Word (0.84)	p-value	<	0.001		
Line (0.58)	- Character (0.75)	p-value	=	0.002		
Word (0.84)	- Character (0.75)	p-value	=	0.002		
	Word > Character	> Line				
	Recall					
Line (0.87)	- Word (0.78)	p-value	=	0.002		
Line (0.87)	- Character (0.58)	p-value	<	0.001		
Word (0.78)	- Character (0.58)	p-value	<	0.001		
	Line > Word > Cha	racter				
	Harmonic Mea	an				
Line (0.67)	- Word (0.8)	p-value	<	0.001		
Line (0.67)	- Character (0.63)	p-value	=	0.487		
Word (0.8)	p-value	<	0.001			
Word > (Line, Character)						

Answer: Word

(Friedman test \rightarrow Bonferroni corr. \rightarrow Wilcoxon test)



Universidade **Federal Fluminense**

INTRODUCTION

MOTIVATING EXAMPLE

ITERATIVE DIFF (IDIFF)

EVALUATION

CONCLUSIONS



In which situations (refactoring types) IDIFF is more precise than a generic Diff tool?

QualitityIDiffWinMerge=IDiffWinMerge=IDiffWinMerge=Answer:Precision140,000,00100,000,0050,0050,000,0075,0025,00II33,3325,0041,6733,33IV1747,065,8847,0676,4711,7611,7629,4158,8211,76Calls simVMoveVII966,6711,1122,2277,780,0050,0050,0050,0050,0050,0050,0075,0025,00VMoveWWNNN <td< th=""><th></th><th>Quantita</th><th colspan="3">Line</th><th></th><th>Word</th><th></th><th colspan="2">Character</th><th></th></td<>		Quantita	Line				Word		Character			
PrecisionII – Com methodI40,000,00100,000,0050,000,0075,0025,00methodII922,220,0077,7855,5622,2222,2255,5611,1133,33IIIIII128,3325,0066,6733,3333,3333,3325,0041,6733,33IV1747,065,8847,0676,4711,7611,7652,9417,6529,41IVMethodV825,0025,0050,0062,5037,500,0050,0050,0050,000,000,00IVMethodVI1723,5335,2941,1847,0641,1811,7629,4158,8211,76VMovVII966,6711,1122,2277,780,0022,2255,5622,2222,2222,22RecallII40,000,00100,000,0050,0050,0058,3341,671,761,761,761,761,761,761,761,761,112,225,5622,2222,2222,2222,2222,2222,2222,2222,2222,221,7664,7123,5325,0050,0025,0012,5062,5025,0025,0025,0025,0025,0025,0025,0025,0025,0025,0025,0025,0025,0025,0025		Quantity	IDiff	WinMerge	=	IDiff	WinMerge	=	IDiff	WinMerge	=	Answer:
I40,000,00100,000,0050,0050,000,0075,0025,00methodII922,220,0077,7855,5622,2222,2255,5611,1133,33IV33,3333,3333,3333,3325,0041,6733,33IV1747,065,8847,0676,4711,7611,7652,9417,6529,4129,4123,5325,0050,0062,5037,500,0050,0050,000,000,00100,000,00100,0022,2255,5622,22<	Precis	ion										II – Com
II922,220,0077,7855,5622,2222,2255,5611,1133,33IIIIII128,3325,0066,6733,3333,3333,3325,0041,6733,33IV1747,065,8847,0676,4711,7611,7652,9417,6529,41IV- MailV825,0025,0050,0062,5037,500,0050,0050,000,000,00V- MovVI1723,5335,2941,1847,0641,1811,7629,4158,8211,76V- MovVII966,6711,1122,2277,780,0022,2255,5622,2222,2222,22RecallI40,000,00100,000,0050,0050,0050,0058,3341,6723,53V825,0012,5062,5025,0050,0025,0012,5062,5025,0025,0025,00VI1711,7635,2952,9417,6552,9429,4111,7670,5917,6511,11VII944,4411,1144,4422,2244,4433,3311,1166,6722,22Harmonic meanI40,000,00100,000,0050,0050,0075,0025,0011,1112,83335,3333,33IV1747,065,8847,06	1	4	0,00	0,00	100,00	0,00	50,00	50,00	0,00	75,00	25,00	mothod
III128,3325,0066,6733,3333,3333,3325,0041,6733,33IVIVMain methodIV1747,065,8847,0676,4711,7611,7652,9417,6529,41V825,0025,0050,0062,5037,500,0050,0050,000,000VI1723,5335,2941,1847,0641,1811,7629,4158,8211,76VII966,6711,1122,2277,780,0022,2255,5622,2222,22RecallI40,000,00100,000,0050,0050,0075,0025,00VII - Sin conditionII911,110,0088,8922,2244,4433,3311,1177,7811,11III120,0025,0075,008,3350,0041,670,0058,3341,67V825,0012,5062,5025,0050,0025,0012,5062,5025,00VI1711,7635,2952,9417,6552,9429,4111,7670,5917,65VI944,4411,1144,4422,2244,4433,3311,1166,6722,22Harmonic meanI40,000,0077,7844,4433,3322,2033,3355,5611,1111128,3325,00 <th>П</th> <th>9</th> <th>22,22</th> <th>0,00</th> <th>77,78</th> <th>55,56</th> <th>22,22</th> <th>22,22</th> <th>55,56</th> <th>11,11</th> <th>33,33</th> <th>methou</th>	П	9	22,22	0,00	77,78	55,56	22,22	22,22	55,56	11,11	33,33	methou
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	Ш	12	8,33	25,00	66,67	33,33	33,33	33,33	25,00	41,67	33,33	IV – Ma
V 8 25,00 25,00 50,00 62,50 37,50 0,00 50,00 50,00 0,00 VI 17 23,53 35,29 41,18 47,06 41,18 11,76 29,41 58,82 11,76 VI 9 66,67 11,11 22,22 77,78 0,00 22,22 55,56 22,22 22,22 Recall I 4 0,00 0,00 100,00 0,00 50,00 0,00 75,00 25,00 V Move II 9 11,11 0,00 88,89 22,22 44,44 33,33 11,11 77,78 11,11 III 12 0,00 25,00 75,00 8,33 50,00 25,00 25,00 25,00 25,00 25,00 29,41 13,76 64,71 23,53 24,44 23,33 11,11 66,67 22,22 24,44 33,33 11,11 66,67 22,22 VII 9 44,44	IV	17	47,06	5,88	47,06	76,47	11,76	11,76	52,94	17,65	29,41	
VI 17 23,53 35,29 41,18 47,06 41,18 11,76 29,41 58,82 11,76 VI 9 66,67 11,11 22,22 77,78 0,00 22,22 55,56 22,22 24,44 33,33 11,11 77,78 11,11 II 9 11,11 0,00 25,00 75,00 8,33 50,00 41,67 0,00 58,33 41,67 0,00 58,33 41,67 0,00 58,33 41,67 0,00 58,33 41,67 0,00 50,00 25,00 52,00 50,00 25,00 <th>V</th> <th>8</th> <th>25,00</th> <th>25,00</th> <th>50,00</th> <th>62,50</th> <th>37,50</th> <th>0,00</th> <th>50,00</th> <th>50,00</th> <th>0,00</th> <th>calls sim</th>	V	8	25,00	25,00	50,00	62,50	37,50	0,00	50,00	50,00	0,00	calls sim
VII 9 66,67 11,11 22,22 77,78 0,00 22,22 55,56 22,22 22,22 between I 4 0,00 0,00 100,00 0,00 50,00 50,00 75,00 25,00 VII 1 9 11,11 0,00 88,89 22,22 44,44 33,33 11,11 77,78 11,11 III 9 11,11 0,00 25,00 75,00 8,33 50,00 41,67 0,00 58,33 41,67 IV 17 23,53 5,88 70,59 29,41 35,29 35,29 11,76 64,71 23,53 V 8 25,00 12,50 62,50 25,00 50,00 25,00 12,50 62,50 25,00 12,50 62,50 25,00 12,50 62,50 25,00 12,50 62,50 25,00 12,50 62,50 25,00 12,50 62,50 25,00 11 66,67 12,50 62,50 <td< th=""><th>VI</th><th>17</th><th>23,53</th><th>35,29</th><th>41,18</th><th>47,06</th><th>41,18</th><th>11,76</th><th>29,41</th><th>58,82</th><th>11,76</th><th>V - Mov</th></td<>	VI	17	23,53	35,29	41,18	47,06	41,18	11,76	29,41	58,82	11,76	V - Mov
Recall between I 4 0,00 0,00 100,00 0,00 50,00 50,00 75,00 25,00 VIII 9 11,11 0,00 88,89 22,22 44,44 33,33 11,11 77,78 11,11 III 12 0,00 25,00 75,00 8,33 50,00 41,67 0,00 58,33 41,67 23,53 41,67 23,53 41,67 23,53 5,88 70,59 24,1 35,29 35,29 11,76 64,71 23,53 0,00 25,00 50,00 25,00 </th <th>VII</th> <th>9</th> <th>66,67</th> <th>11,11</th> <th>22,22</th> <th>77,78</th> <th>0,00</th> <th>22,22</th> <th>55,56</th> <th>22,22</th> <th>22,22</th> <th>V IVIOV</th>	VII	9	66,67	11,11	22,22	77,78	0,00	22,22	55,56	22,22	22,22	V IVIOV
I 4 0,00 0,00 100,00 0,00 50,00 50,00 75,00 25,00 V 25,00 VII 11,11 0,00 88,89 22,22 44,44 33,33 11,11 77,78 11,11 III 12 0,00 25,00 75,00 8,33 50,00 41,67 0,00 58,33 41,67 IV 17 23,53 5,88 70,59 29,41 35,29 35,29 11,76 64,71 23,53 V 8 25,00 12,50 62,50 25,00 50,00 25,00 12,50 62,50 25,00 25,00 17,65 21,00 25,00 17,65 21,01 24,44 33,33 11,11 66,67 22,22 44,44 33,33 11,11 66,67 22,22 22,22 44,44 33,33 11,11 66,67 22,22 22,22 17,65 11,76 70,59 17,65 11,11 111 111 111 14 0,00 0,00 50,00 50,00 50,00 25,00 25,00 11,11 111	Recall	l										betwee
II 9 11,11 0,00 88,89 22,22 44,44 33,33 11,11 77,78 11,11 III 12 0,00 25,00 75,00 8,33 50,00 41,67 0,00 58,33 41,67 IV 17 23,53 5,88 70,59 29,41 35,29 35,29 11,76 64,71 23,53 V 8 25,00 12,50 62,50 25,00 50,00 25,00 12,50 62,50 25,00 VI 17 11,76 35,29 52,94 29,41 11,76 70,59 17,65 VII 9 44,44 11,11 44,44 22,22 44,44 33,33 11,11 66,67 22,22 Harmonic mean I 4 0,00 0,00 100,00 0,00 50,00 50,00 0,00 75,00 25,00 II 9 22,22 0,00 77,78 44,44 33,33 22,22 33,33 55,56 11,11 III 9 22,20 66,67 16,67	1	4	0,00	0,00	100,00	0,00	50,00	50,00	0,00	75,00	25,00	VII – Sin
III 12 0,00 25,00 75,00 8,33 50,00 41,67 0,00 58,33 41,67 Conditio IV 17 23,53 5,88 70,59 29,41 35,29 35,29 11,76 64,71 23,53 70,59 29,41 35,29 12,50 62,50 25,00 25,00 12,50 62,50 25,00 12,50 62,50 25,00 12,50 62,50 25,00 12,50 62,50 25,00 12,50 62,50 25,00 12,50 62,50 25,00 12,50 62,50 25,00 12,50 62,50 25,00 12,50 62,50 25,00 12,50 62,50 25,00 12,50 62,50 25,00 12,50 62,50 25,00 14,44 33,33 11,11 66,67 22,22 44,44 33,33 22,22 33,33 55,56 11,11 11 12 8,33 25,00 77,78 44,44 33,33 22,22 33,33 58,33 33,33 33,33 33,33 33,33 33,33 33,33 33,33 33,33 33,33 <td< th=""><th>П</th><th>9</th><th>11,11</th><th>0,00</th><th>88,89</th><th>22,22</th><th>44,44</th><th>33,33</th><th>11,11</th><th>77,78</th><th>11,11</th><th></th></td<>	П	9	11,11	0,00	88,89	22,22	44,44	33,33	11,11	77,78	11,11	
IV 17 23,53 5,88 70,59 29,41 35,29 35,29 11,76 64,71 23,53 V 8 25,00 12,50 62,50 25,00 50,00 25,00 12,50 62,50 25,00 VI 17 11,76 35,29 52,94 17,65 52,94 29,41 11,76 70,59 17,65 VII 9 44,44 11,11 44,44 22,22 44,44 33,33 11,11 66,67 22,22 Harmonic mean I 4 0,00 0,00 100,00 0,00 50,00 50,00 0,00 75,00 25,00 11,11 III 9 22,22 0,00 77,78 44,44 33,33 22,22 33,33 55,56 11,11 III 12 8,33 25,00 62,50 25,00 33,33 8,33 58,33 33,33 33,33 IV 17 47,06 5,88 47,06 70,59 17,65 11,76 35,29 47,06 17,65 Making method Moving Featu </th <th>Ш</th> <th>12</th> <th>0,00</th> <th>25,00</th> <th>75,00</th> <th>8,33</th> <th>50,00</th> <th>41,67</th> <th>0,00</th> <th>58,33</th> <th>41,67</th> <th>conditio</th>	Ш	12	0,00	25,00	75,00	8,33	50,00	41,67	0,00	58,33	41,67	conditio
V 8 25,00 12,50 62,50 25,00 50,00 25,00 12,50 62,50 25,00 VI 17 11,76 35,29 52,94 17,65 52,94 29,41 11,76 70,59 17,65 VI 9 44,44 11,11 44,44 22,22 44,44 33,33 11,11 66,67 22,22 Harmonic mean V 4 0,00 0,00 100,00 0,00 50,00 0,00 75,00 25,00 11,11 11 9 22,22 0,00 77,78 44,44 33,33 22,22 33,33 55,56 11,11 111 12 8,33 25,00 66,67 16,67 50,00 33,33 8,33 58,33 33,33 1V 17 47,06 5,88 47,06 70,59 17,65 11,76 35,29 47,06 17,65 V 8 25,00 12,50 62,50 25,00 50,00 25,00 <th>IV</th> <th>17</th> <th>23,53</th> <th>5,88</th> <th>70,59</th> <th>29,41</th> <th>35,29</th> <th>35,29</th> <th>11,76</th> <th>64,71</th> <th>23,53</th> <th>overocci</th>	IV	17	23,53	5,88	70,59	29,41	35,29	35,29	11,76	64,71	23,53	overocci
VI 17 11,76 35,29 52,94 17,65 52,94 29,41 11,76 70,59 17,65 VII 9 44,44 11,11 44,44 22,22 44,44 33,33 11,11 66,67 22,22 Harmonic mean Legend: II 4 0,00 0,00 100,00 0,00 50,00 0,00 75,00 25,00 III 9 22,22 0,00 77,78 44,44 33,33 22,22 33,33 55,56 11,11 III 12 8,33 25,00 66,67 16,67 50,00 33,33 8,33 58,33 33,33 IV 17 47,06 5,88 47,06 70,59 17,65 11,76 35,29 47,06 17,65 V 8 25,00 12,50 62,50 25,00 50,00 25,00 12,50 62,50 25,00	V	8	25,00	12,50	62,50	25,00	50,00	25,00	12,50	62,50	25,00	expressi
VII 9 44,44 11,11 44,44 22,22 44,44 33,33 11,11 66,67 22,22 Harmonic mean - <	VI	17	11,76	35,29	52,94	17,65	52,94	29,41	11,76	70,59	17,65	
Harmonic mean I 4 0,00 0,00 100,00 0,00 50,00 0,00 75,00 25,00 Big Refactorin II 9 22,22 0,00 77,78 44,44 33,33 22,22 33,33 55,56 11,11 III 12 8,33 25,00 66,67 16,67 50,00 33,33 8,33 58,33 33,33 IV 17 47,06 5,88 47,06 70,59 17,65 11,76 35,29 47,06 17,65 V 8 25,00 12,50 62,50 25,00 50,00 25,00 12,50 62,50 25,00 VI 17 23,53 35,29 41,18 52,94 35,29 11,76 11,76 76,47 11,76 VII 9 77,78 0,00 22,22 66,67 11,11 22,22 55,56 22,22 22,22	VII	9	44,44	11,11	44,44	22,22	44,44	33,33	11,11	66,67	22,22	
I 4 0,00 0,00 100,00 0,00 50,00 0,00 75,00 25,00 Big Refactorin II 9 22,22 0,00 77,78 44,44 33,33 22,22 33,33 55,56 11,11 III 12 8,33 25,00 66,67 16,67 50,00 33,33 8,33 58,33 33,33 IV 17 47,06 5,88 47,06 70,59 17,65 11,76 35,29 47,06 17,65 V 8 25,00 12,50 62,50 25,00 50,00 25,00 12,50 62,50 25,00 VI 17 23,53 35,29 41,18 52,94 35,29 11,76 11,76 76,47 11,76 VII 9 77,78 0,00 22,22 66,67 11,11 22,22 55,56 22,22 22,22	Harm	onic mean	i.									
II 9 22,22 0,00 77,78 44,44 33,33 22,22 33,33 55,56 11,11 III 12 8,33 25,00 66,67 16,67 50,00 33,33 8,33 58,33 33,33 IV 17 47,06 5,88 47,06 70,59 17,65 11,76 35,29 47,06 17,65 V 8 25,00 12,50 62,50 25,00 50,00 25,00 12,50 62,50 25,00 VI 17 23,53 35,29 41,18 52,94 35,29 11,76 11,76 76,47 11,76 VII 9 77,78 0,00 22,22 66,67 11,11 22,22 55,56 22,22 22,22 Simplifying col	1	4	0,00	0,00	100,00	0,00	50,00	50,00	0,00	75,00	25,00	Legend:
III 12 8,33 25,00 66,67 16,67 50,00 33,33 8,33 58,33 33,33 IC Composing Milling IV 17 47,06 5,88 47,06 70,59 17,65 11,76 35,29 47,06 17,65 Making method V 8 25,00 12,50 62,50 25,00 50,00 25,00 12,50 62,50 25,00 Making method Moving Featu VI 17 23,53 35,29 41,18 52,94 35,29 11,76 11,76 76,47 11,76 VII 9 77,78 0,00 22,22 66,67 11,11 22,22 55,56 22,22 22,22 Simplifying col	П	9	22,22	0,00	77,78	44,44	33,33	22,22	33,33	55,56	11,11	Big Refactorin
IV 17 47,06 5,88 47,06 70,59 17,65 11,76 35,29 47,06 17,65 Making metho V 8 25,00 12,50 62,50 25,00 50,00 25,00 12,50 62,50 25,00 Making metho VI 17 23,53 35,29 41,18 52,94 35,29 11,76 11,76 76,47 11,76 VII 9 77,78 0,00 22,22 66,67 11,11 22,22 55,56 22,22 22,22 Simplifying col	Ш	12	8,33	25,00	66,67	16,67	50,00	33,33	8,33	58,33	33,33	Composing M
V 8 25,00 12,50 62,50 25,00 50,00 25,00 12,50 62,50 25,00 VI 17 23,53 35,29 41,18 52,94 35,29 11,76 11,76 76,47 11,76 VII 9 77,78 0,00 22,22 66,67 11,11 22,22 55,56 22,22 22,22 Simplifying col	IV	17	47,06	5,88	47,06	70,59	17,65	11,76	35,29	47,06	17,65	Dealing with g
VI 17 23,53 35,29 41,18 52,94 35,29 11,76 11,76 76,47 11,76 Organizing da VII 9 77,78 0,00 22,22 66,67 11,11 22,22 55,56 22,22 22,22 Simplifying corr	V	8	25,00	12,50	62,50	25,00	50,00	25,00	12,50	62,50	25,00	Moving Feature
VII 9 77,78 0,00 22,22 66,67 11,11 22,22 55,56 22,22 22,22 Simplifying col	VI	17	23,53	35,29	41,18	52,94	35,29	11,76	11,76	76,47	11,76	Organizing da
	VII	9	77,78	0,00	22,22	66,67	11,11	22,22	55,56	22,22	22,22	Simplifying cor

- Composing nethods V – Making method alls simpler – Moving features etween objects /II – Simplifying onditional xpression

.egena:	
Big Refactoring	
and the second se	

Big Refactoring	J.
Composing Methods	11
Dealing with generalization	Ш
Making method calls simpler	IV
Moving Features between objects	v
Organizing data	VI
Simplifying conditional expression	VII



Precision – word grain



Recall – word grain



F-measure – word grain



Related Work

- Clone detection
- Refactoring detection
- Diff
 - Malpohl (2003): rename detection, language specific
 - Canfora et al. (2009): improvements over Unix
 Diff, line grain
 - Antoniol et al. (2004): evolution discontinuities, language specific