Lecture 11

LSdiff evaluation / Focus group study
Mining Software Repositories, Part 1
eRose

Announcement

- Project Checkpoint Due on this thursday.
  - I won't grade them.
  - It is not mandatory.
  - You are encouraged to submit to seek my feedback.
  - Available for both research project, literature survey, and tool evaluation

Today's Agenda

- LSdiff evaluation
- LSdiff focus group study
- Presentation: Tileli (advocate), Gaurav (skeptic)
- eRose
- Quiz

Question: What kinds of rules can LSdiff find?

Horn Clause: `A(x) ∧ B(x,y) ∧ C(y) ⇒ D(x,y)`

Rule Styles	High-level Change Patterns	Example
`past_* => deleted_*`	dependency removal, feature deletion, etc.	`past_calls(m, "DB.exec") ⇒` `deleted_calls(m, "DB.exec")`
`past_* => added_*`	consistent updates to clones, etc.	`past_accesses( "Log.on", m)⇒` `added_calls(m, "Log.trace")`

`current_* => added_*`	dependency addition, feature addition, etc.	`current_method(m, "getHost", t)∧`
		`current_subtype("Svc", t) ⇒`
		`added_calls(m, "Log.trace")`

`deleted_* => added_*` `added_* => deleted_*`	related code change, API replacement, etc.	`deleted_method(m, "getHost", t) ⇒` `added_inheritedfield("getHost", "Svc",t)`

LSdiff Evaluation

- Quantitative Evaluation
- Qualitative Evaluation
  - **Focus Group Study**
  - Comparison with *diff*
  - Comparison with *check-in comments*
- Impact of Input Parameters

LSdiff Evaluation: Research Questions

1. How often do individual changes form systematic change patterns?
2. How concisely does LSdiff describe structural differences in comparison to existing differencing approach at the same abstraction level?
3. How much contextual information does LSdiff find from unchanged code fragments?

LSdiff Evaluation: Research Questions

1. How often do individual changes form systematic change patterns? Measure *coverage*, # of facts in ΔFB matched by inferred rules
2. How concisely does LSdiff describe structural differences in comparison to existing differencing approach at the same abstraction level? Measure *conciseness*, ΔFB / (# rules + # facts)
3. How much contextual information does LSdiff find from unchanged code fragments? Measure *the number of facts mentioned by rules but* are not contained in ΔFB

Comparison with Δ FB

	FB_o	FB_n	ΔFB	Rule	Fact	Cvrg.	Csc.	Ad'l.
10 revision pairs in carol (carol.objectweb.org)								
Min	3080	3452	15	1	3	59%	2.3	0.0
Max	10746	10610	1812	36	71	98%	27.5	19.0
Med	9615	9635	97	5	16	87%	5.8	4.0
Avg	8913	8959	426	10	20	85%	9.9	5.5
29 release pairs in dnsjava (www.dnsjava.org)								
Min	3109	3159	4	0	2	0%	1.0	0.0
Max	7200	7204	1500	36	201	98%	36.1	91.0
Med	4817	5096	168	3	24	88%	4.8	0.0
Avg	5144	5287	340	8	37	73%	8.4	14.9
10 version pairs in LSdiff								
Min	8315	8500	2	0	2	0%	1.0	0.0
Max	9042	9042	396	6	54	97%	28.9	12.0
Med	8732	8756	142	1	11	91%	9.8	0.0
Avg	8712	8783	172	2	17	68%	11.2	2.3
three data sets above								
Med	6650	6712	132	2	17	89%	7.3	0.0
Avg	6632	6732	302	7	27	75%	9.3	9.7

(m=3, a=0.75, k=2)

Focus Group Study

- **Why** would you conduct a focus group study?
- **When** do you you conduct one?
- **What** can you learn from a focus group?

Focus Group Study

- **Why** would you conduct a focus group study?
  - to explore how customers will respond to a new idea
  - testing new concepts, products, and messages
- **What** can you learn from a focus group?
  - exploratory qualitative research: "thermometer" that allows you to test the "temperature" of consumers' reactions to your research topics
  - no statistical sampling of the target population
  - less formal than a survey
  - in-depth understanding of the target's perspectives or opinions

When to Use and When to Avoid

- When the concept or idea you wish to evaluate is new and when the best evaluation comes from letting the target customer view the concept directly.
  - e.g. new advertising campaign
- When not to do this
  - testing consumer reactions when there was no budget to accommodate changes
  - when you ask "how many...?" and "how much...?" questions or need graphs, tables, etc.
  - testing personally sensitive issues: medical conditions, politics, sex, etc.

How to conduct a focus group study

- research objectives
- recruiting profile
- screener questionnaire
- invitation to participate & co-op fee
- discussion guide
- moderator
- audio or video taping
- transcript ==> quotes

Focus Group Study

- Screener questionnaire
- Participants: five professional software engineers
  - industry experience ranging from 6 to over 30 years
  - use *diff* and *diff*-based version control system daily
  - review code changes daily except one who did weekly
- One hour structured discussion
  - I worked as a moderator. We also had a note-taker transcribe the discussion. Discussion was audio-taped and transcribed.

Focus Group Hands-On Trial

Carol Revision 430.
SVN check-in message: Common methods go in an abstract class. Easier to extend/maintain/fix
Author: benoif @ Thu Mar 10 12:21:46 2005 UTC
723 lines of changes across 9 files (2 new files and 7 modified files).

Overview

Generated based on LSDiff output.

Inferred Rules		
1	(50/50)	By this change, six classes inherit many methods from AbsRegistry class.
2	(32/32)	By this change, six classes implement NameService interface.
3	(6/8)	All methods that are included in JacORBCosNaming class and NameService interface are deleted except start and stop methods.
4	(5/6)	All host fields in the classes that implement NameService interface got deleted except LmiRegistry class.
5	(5/6)	All port fields in the classes that implement NameService interface got deleted except LmiRegistry class.
6	(5/6)	All getHost methods in the classes that implement NameService interface got deleted except LmiRegistry class.

http://www.cs.washington.edu/homes/miryung/LSDiff/carol429-430.htm

Focus Group Hands-On Trial

```
46: public class IIOPCosNaming extends AbsRegistry implements NameService {
47:
48:     /**
49:      * Default port number ( 12350 for default)
50:      */
All DEFAULT PORT NUMBER fields are added fields except JacORBCosNaming class.
51:     private static final int DEFAUL_PORT DEFAULT_PORT_NUMBER = 12350;
52:
53:     /**
54:      * Sleep time to wait
55:      */
56:     private static final int SLEEP_TIME = 2000;
57:
58:     /**
59:      * port number
60:      */
All port fields in the classes that implement NameService interface got deleted except LmiRegistry
 61:     private int port = DEFAUL_PORT;
62:
63:     /**
64:      * Hostname to use
65:      */
All host fields in the classes that implement NameService interface got deleted except LmiRegistry
 66:     private String host = null;
```

Show related changes

http://www.cs.washington.edu/homes/miryung/LSDiff/carol429-430.htm

Focus-Group Participants' Comments

"You can't infer the intent of a programmer, but this is pretty close."

"This 'except' thing is great!"

"This is cool. I'd use it if we had one."

"This is a definitely winner tool."

Focus-Group Participants' Comments

"This looks great for big architectural changes, but I wonder what it would give you if you had lots of random changes."

"This wouldn't be used if you were just working with one file."

"This will look for relationships that do not exist."

Recap

- Many differencing techniques individually compare code elements at particular granularities using similarity measures.
  - Hard to comprehend as a long list of matches
  - Difficult to identify exceptions that violate systematic patterns
- LSdiff uses rule-based change representations to explicitly capture systematic changes and automatically infers these rules.

Presentation on eRose

- Tileli
- Guarav

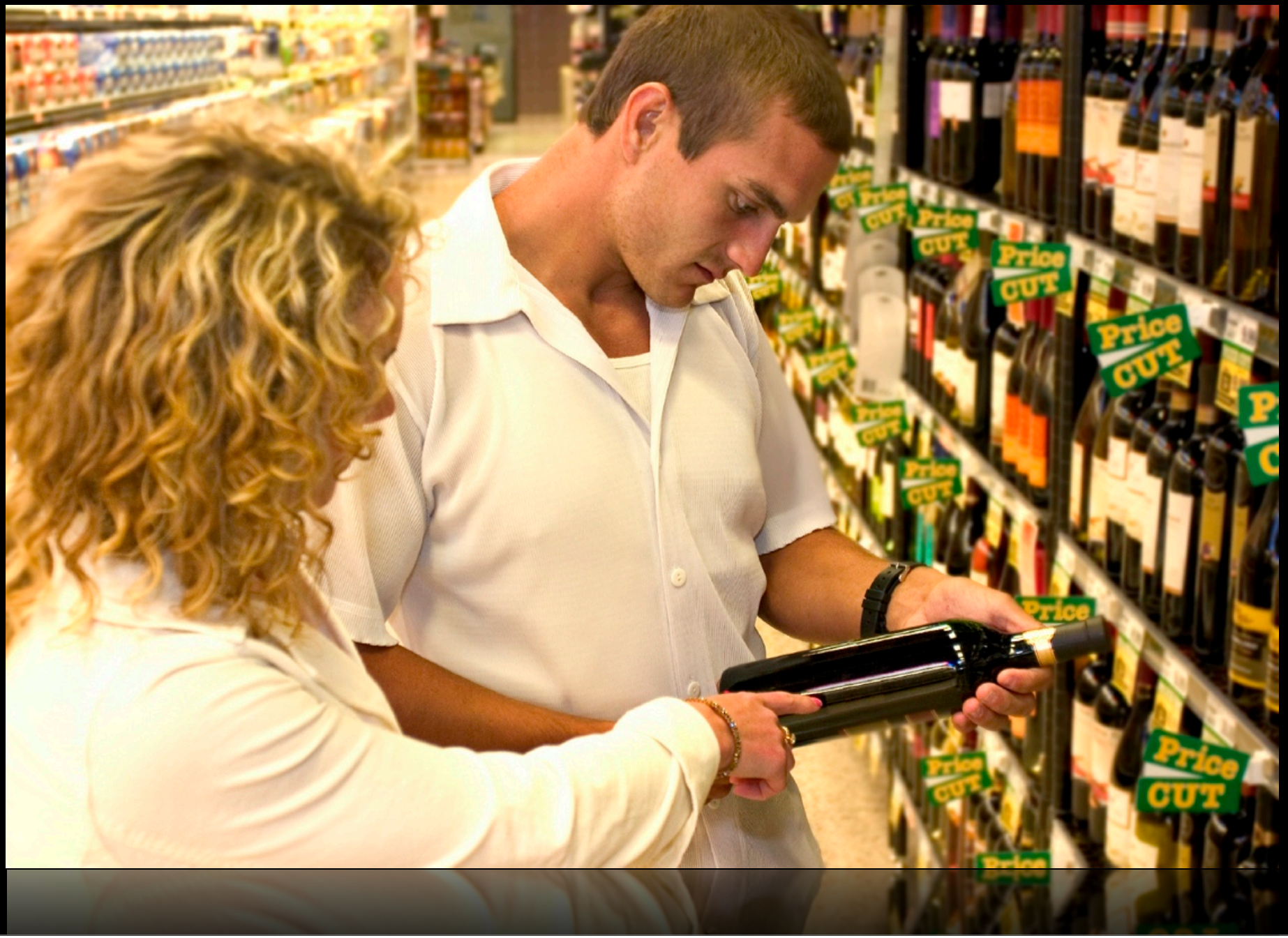
eROSE

Related Changes

(ICSE 2004, TSE 2005)

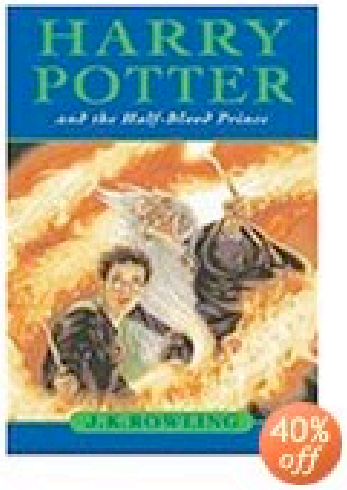
Tom Zimmermann • Saarland University
Peter Weißgerber • University of Trier
Stephan Diehl • University of Trier
Andreas Zeller • Saarland University

Price
CUT


Price
CUT

Harry Potter and the Half-Blood Prince (Book 6) (Hardcover)

by J. K. Rowling



See larger image and other views

List Price: ~~CDN$ 41.00~~
Our Price: **CDN$ 24.60** & eligible for **FREE Super Saver Shipping** on orders over CDN$ 39. Details
You Save: CDN$ 16.40 (40%)

Availability: In Stock. Ships from and sold by Amazon.ca.

8 used & new available from CDN$ 12.00

Find all things Harry in our Harry Potter Store.

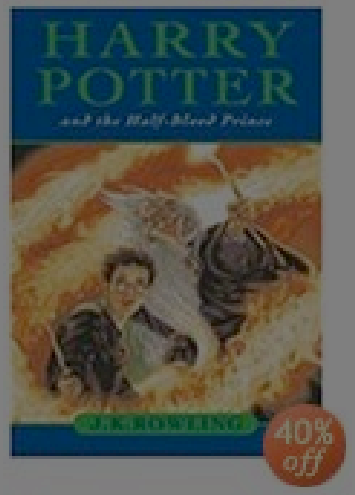
Customers who bought this item also bought

- J.K. Rowling: Classic Books from the Library of Hogwarts School Of... by J. K. Rowling
- Harry Potter and the Order of the Phoenix by J.K. Rowling
- Harry Potter and the Order of the Phoenix (Book 5) by J.K. Rowling
- Harry Potter and the Prisoner of Azkaban (Widescreen) **DVD** ~ Alfonso Cuarón
- Harry Potter and the Goblet of Fire (Book 4) by J.K. Rowling
- Eldest (Inheritance, Book 2) by Christopher Paolini

Explore similar items

Harry Potter and the Half-Blood Prince (Book 6) (Hardcover)

by J. K. Rowling



See larger image and other views

List Price: ~~CDN$ 41.00~~
Our Price: **CDN$ 24.60** & eligible for **FREE Super Saver Shipping** on orders over CDN$ 39. Details
You Save: CDN$ 16.40 (40%)

Availability: In Stock. Ships from and sold by Amazon.ca.

8 used & new available from CDN$ 12.00

Find all things Harry in our Harry Potter Store.

Customers who bought this item also bought

- J.K. Rowling: Classic Books from the Library of Hogwarts School Of... by J. K. Rowling
- Harry Potter and the Order of the Phoenix by J.K. Rowling
- Harry Potter and the Order of the Phoenix (Book 5) by J.K. Rowling
- Harry Potter and the Prisoner of Azkaban (Widescreen) **DVD** ~ Alfonso Cuarón
- Harry Potter and the Goblet of Fire (Book 4) by J.K. Rowling
- Eldest (Inheritance, Book 2) by Christopher Paolini

Explore similar items

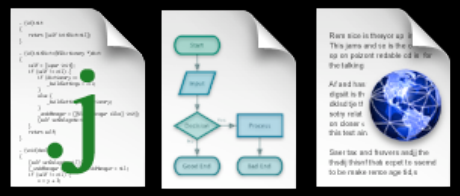
eROSE: Guiding Developers

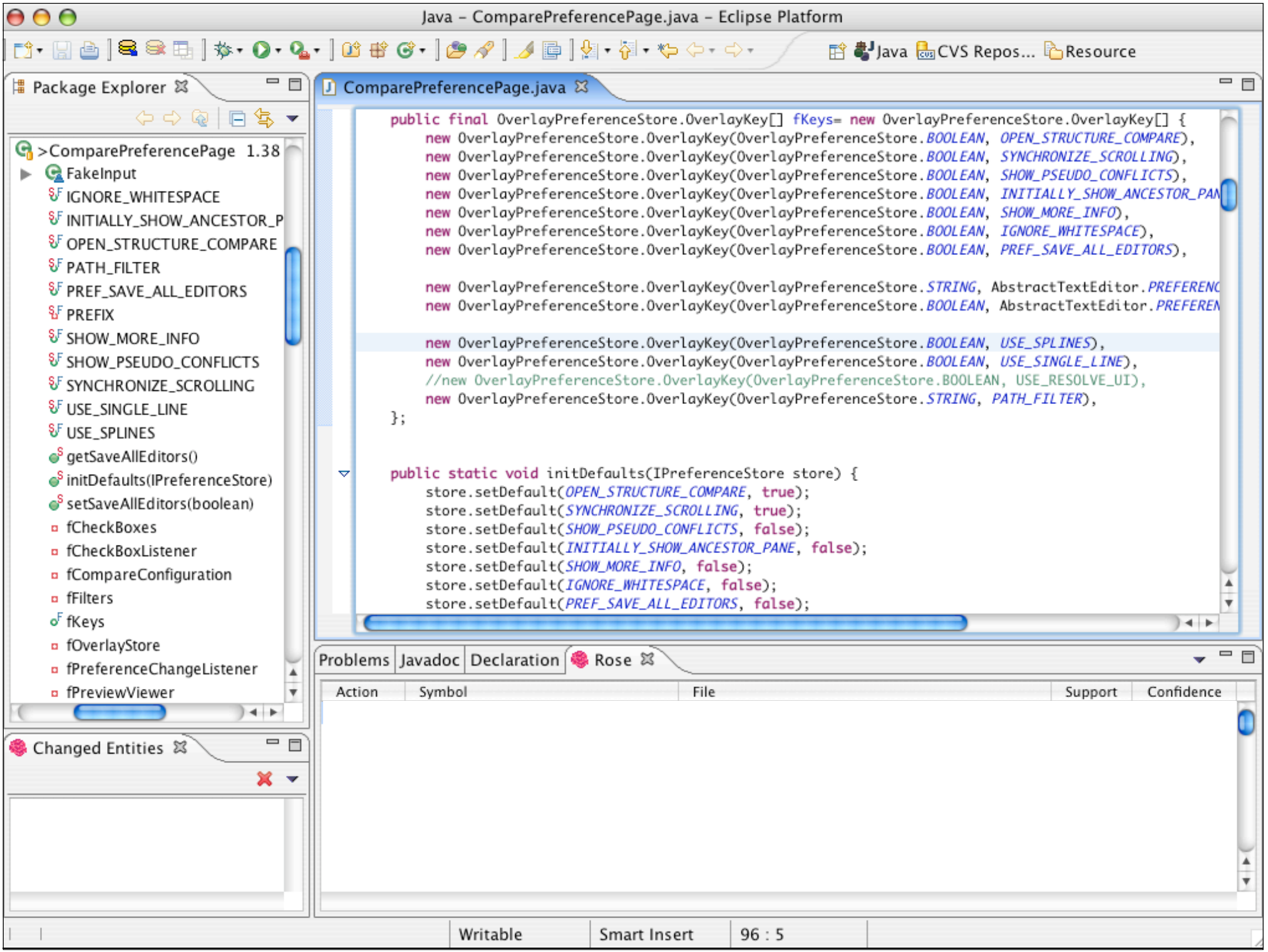
Customers who bought this item also bought...

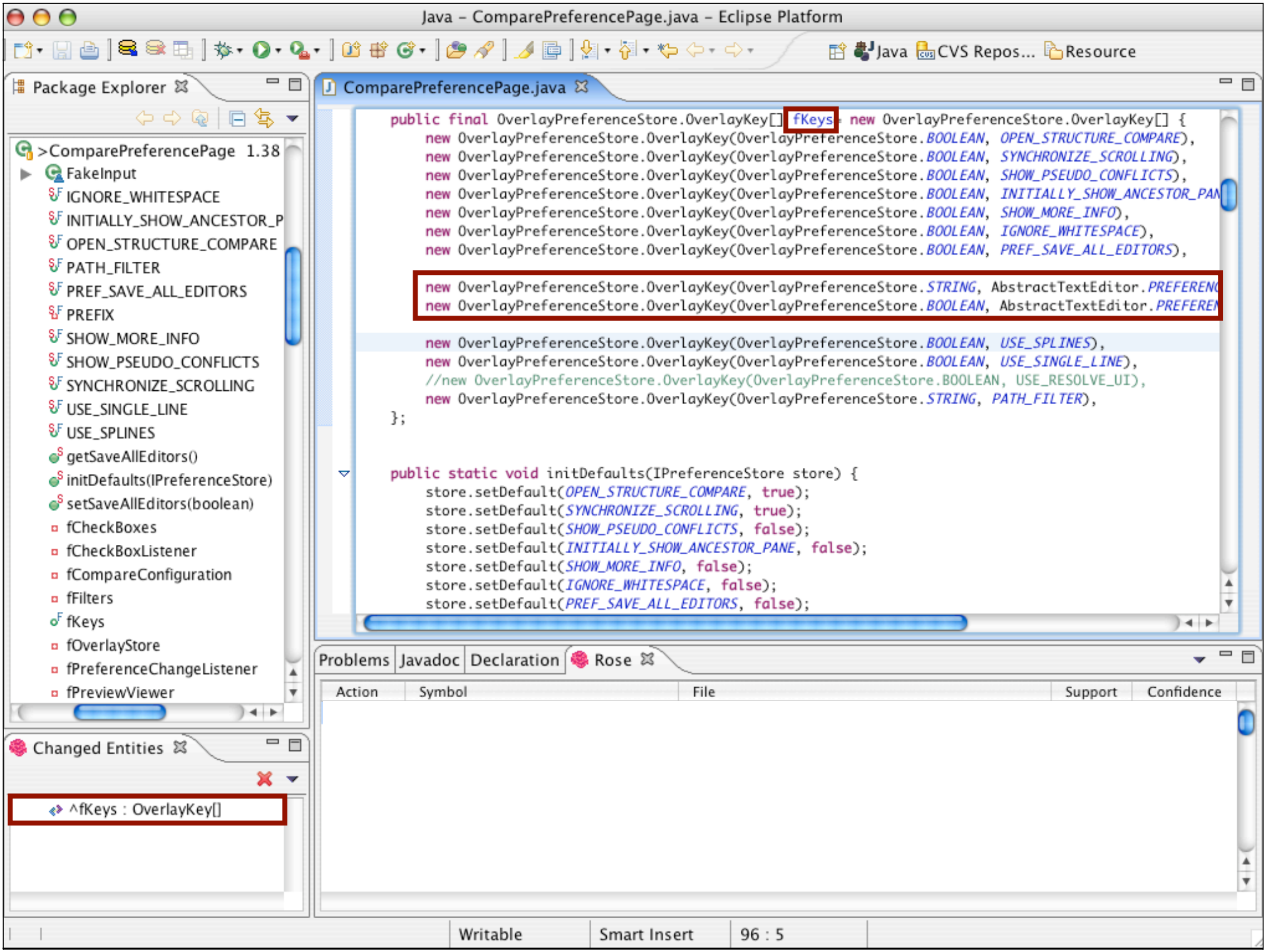
Purchase History

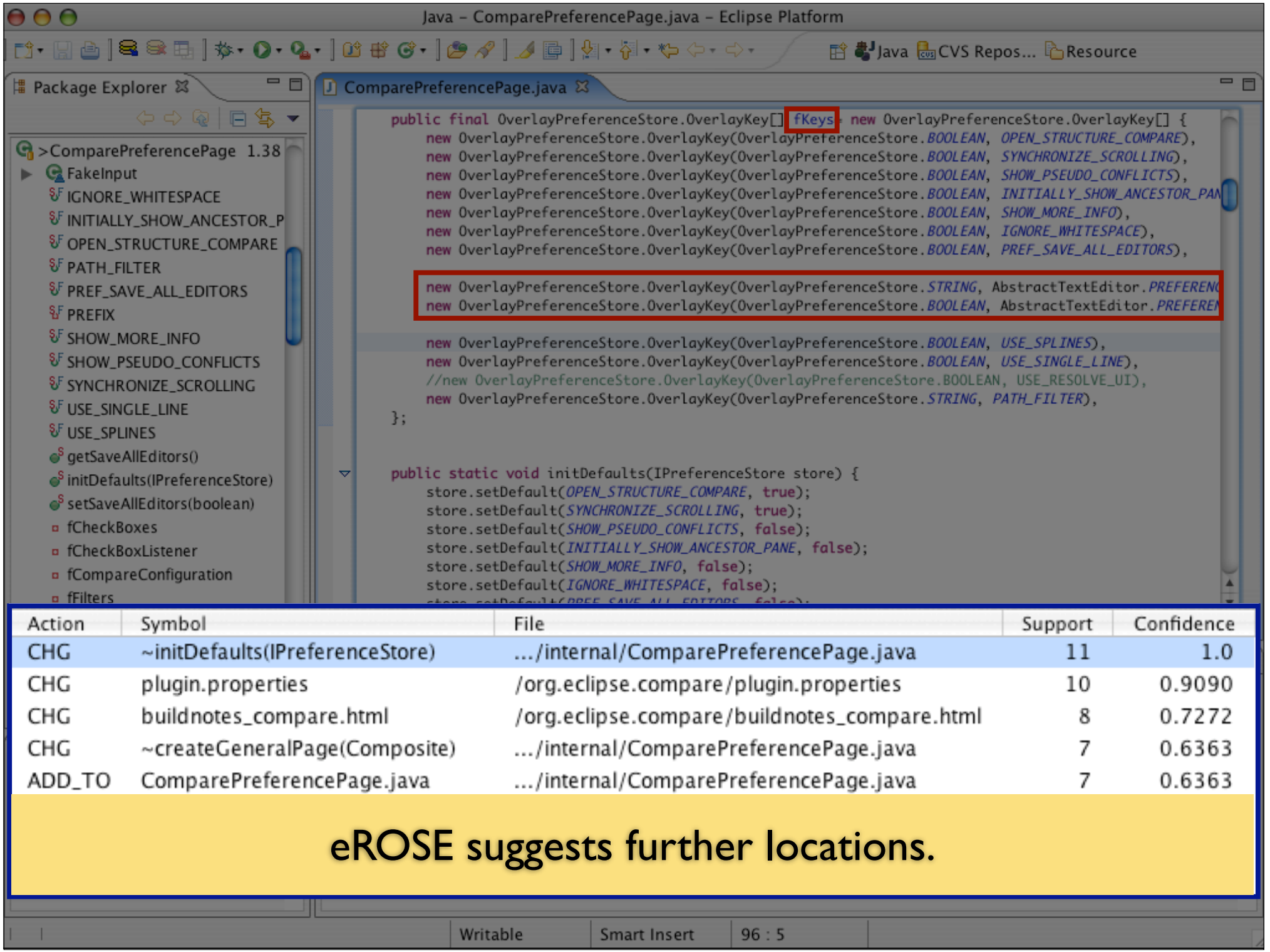
Developers who changed this function also changed...

Version Archive



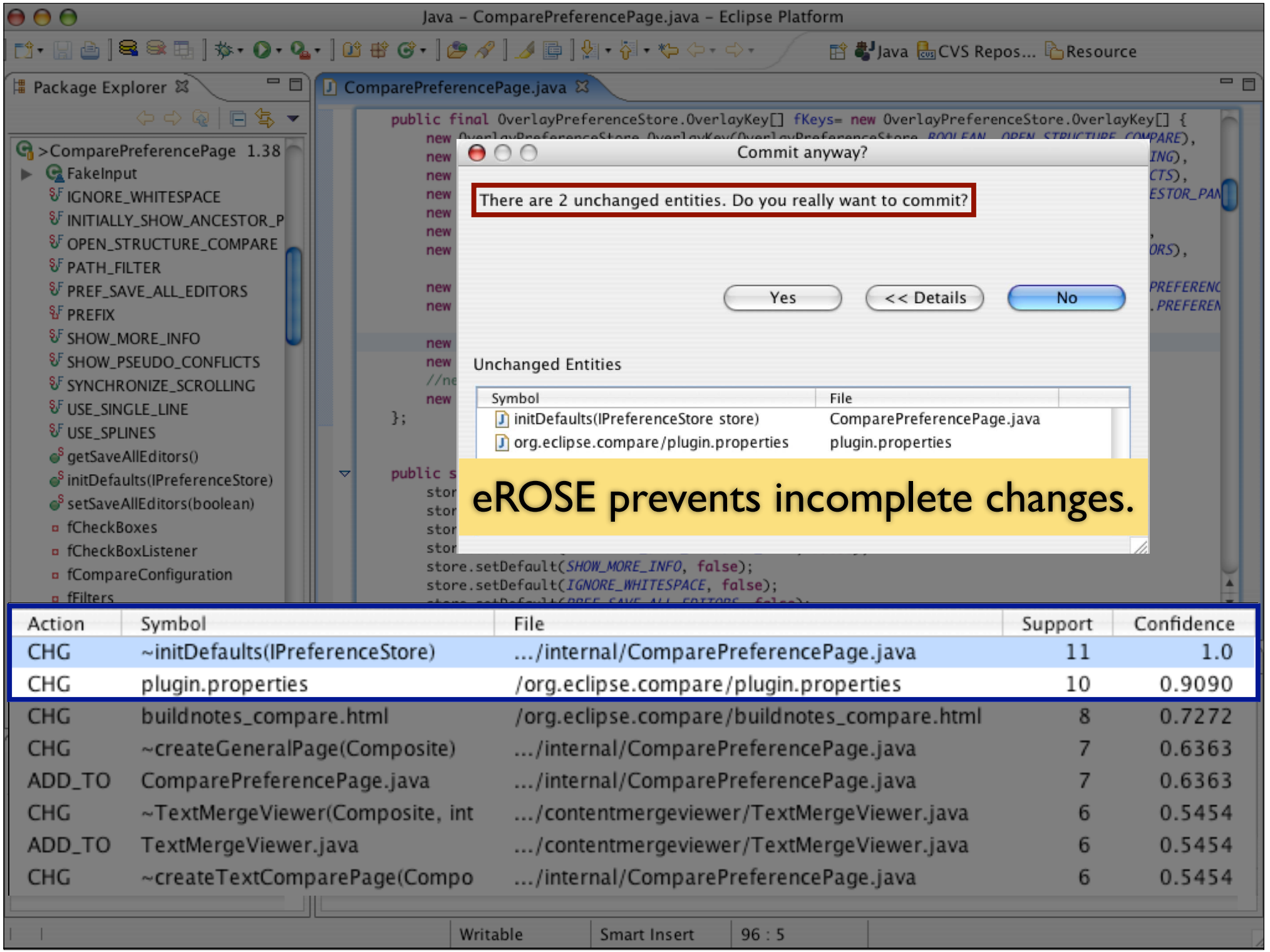

Java – ComparePreferencePage.java – Eclipse Platform
Java
CVS Repos...
Resource
Package Explorer
>ComparePreferencePage 1.38
FakeInput
IGNORE_WHITESPACE
INITIALLY_SHOW_ANCESTOR_P
OPEN_STRUCTURE_COMPARE
PATH_FILTER
PREF_SAVE_ALL_EDITORS
PREFIX
SHOW_MORE_INFO
SHOW_PSEUDO_CONFLICTS
SYNCHRONIZE_SCROLLING
USE_SINGLE_LINE
USE_SPLINES
getSaveAllEditors()
initDefaults(IPreferenceStore)
setSaveAllEditors(boolean)
fCheckBoxes
fCheckBoxListener
fCompareConfiguration
fFilters
fKeys
fOverlayStore
fPreferenceChangeListener
fPreviewViewer
ComparePreferencePage.java
public final OverlayPreferenceStore.OverlayKey[] fKeys= new OverlayPreferenceStore.OverlayKey[] {
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.BOOLEAN, OPEN_STRUCTURE_COMPARE),
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.BOOLEAN, SYNCHRONIZE_SCROLLING),
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.BOOLEAN, SHOW_PSEUDO_CONFLICTS),
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.BOOLEAN, INITIALLY_SHOW_ANCESTOR_PAN
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.BOOLEAN, SHOW_MORE_INFO),
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.BOOLEAN, IGNORE_WHITESPACE),
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.BOOLEAN, PREF_SAVE_ALL_EDITORS),
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.STRING, AbstractTextEditor.PREFERENC
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.BOOLEAN, AbstractTextEditor.PREFEREN
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.BOOLEAN, USE_SPLINES),
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.BOOLEAN, USE_SINGLE_LINE),
//new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.BOOLEAN, USE_RESOLVE_UI),
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.STRING, PATH_FILTER),
};
public static void initDefaults(IPreferenceStore store) {
store.setDefault(OPEN_STRUCTURE_COMPARE, true);
store.setDefault(SYNCHRONIZE_SCROLLING, true);
store.setDefault(SHOW_PSEUDO_CONFLICTS, false);
store.setDefault(INITIALLY_SHOW_ANCESTOR_PANE, false);
store.setDefault(SHOW_MORE_INFO, false);
store.setDefault(IGNORE_WHITESPACE, false);
store.setDefault(PREF_SAVE_ALL_EDITORS, false);
Changed Entities
Problems
Javadoc
Declaration
Rose
Action
Symbol
File
Support
Confidence
Writable
Smart Insert
96 : 5


Java – ComparePreferencePage.java – Eclipse Platform
Java
CVS Repos...
Resource
Package Explorer
>ComparePreferencePage 1.38
FakeInput
IGNORE_WHITESPACE
INITIALLY_SHOW_ANCESTOR_P
OPEN_STRUCTURE_COMPARE
PATH_FILTER
PREF_SAVE_ALL_EDITORS
PREFIX
SHOW_MORE_INFO
SHOW_PSEUDO_CONFLICTS
SYNCHRONIZE_SCROLLING
USE_SINGLE_LINE
USE_SPLINES
getSaveAllEditors()
initDefaults(IPreferenceStore)
setSaveAllEditors(boolean)
fCheckBoxes
fCheckBoxListener
fCompareConfiguration
fFilters
fKeys
fOverlayStore
fPreferenceChangeListener
fPreviewViewer
Changed Entities
^fKeys : OverlayKey[]
ComparePreferencePage.java
public final OverlayPreferenceStore.OverlayKey[] fKeys = new OverlayPreferenceStore.OverlayKey[] {
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.BOOLEAN, OPEN_STRUCTURE_COMPARE),
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.BOOLEAN, SYNCHRONIZE_SCROLLING),
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.BOOLEAN, SHOW_PSEUDO_CONFLICTS),
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.BOOLEAN, INITIALLY_SHOW_ANCESTOR_PAN
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.BOOLEAN, SHOW_MORE_INFO),
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.BOOLEAN, IGNORE_WHITESPACE),
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.BOOLEAN, PREF_SAVE_ALL_EDITORS),
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.STRING, AbstractTextEditor.PREFEREN
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.BOOLEAN, AbstractTextEditor.PREFERE
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.BOOLEAN, USE_SPLINES),
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.BOOLEAN, USE_SINGLE_LINE),
//new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.BOOLEAN, USE_RESOLVE_UI),
new OverlayPreferenceStore.OverlayKey(OverlayPreferenceStore.STRING, PATH_FILTER),
};
public static void initDefaults(IPreferenceStore store) {
store.setDefault(OPEN_STRUCTURE_COMPARE, true);
store.setDefault(SYNCHRONIZE_SCROLLING, true);
store.setDefault(SHOW_PSEUDO_CONFLICTS, false);
store.setDefault(INITIALLY_SHOW_ANCESTOR_PANE, false);
store.setDefault(SHOW_MORE_INFO, false);
store.setDefault(IGNORE_WHITESPACE, false);
store.setDefault(PREF_SAVE_ALL_EDITORS, false);
Problems
Javadoc
Declaration
Rose
Action
Symbol
File
Support
Confidence
Writable
Smart Insert
96 : 5

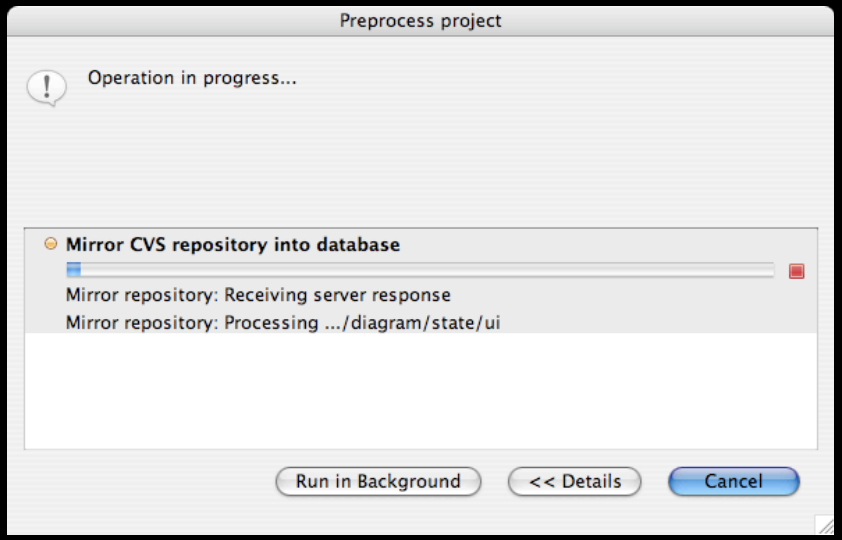


Action	Symbol	File	Support	Confidence
CHG	~initDefaults(IPreferenceStore)	.../internal/ComparePreferencePage.java	11	1.0
CHG	plugin.properties	/org.eclipse.compare/plugin.properties	10	0.9090
CHG	buildnotes_compare.html	/org.eclipse.compare/buildnotes_compare.html	8	0.7272
CHG	~createGeneralPage(Composite)	.../internal/ComparePreferencePage.java	7	0.6363
ADD_TO	ComparePreferencePage.java	.../internal/ComparePreferencePage.java	7	0.6363

eROSE suggests further locations.

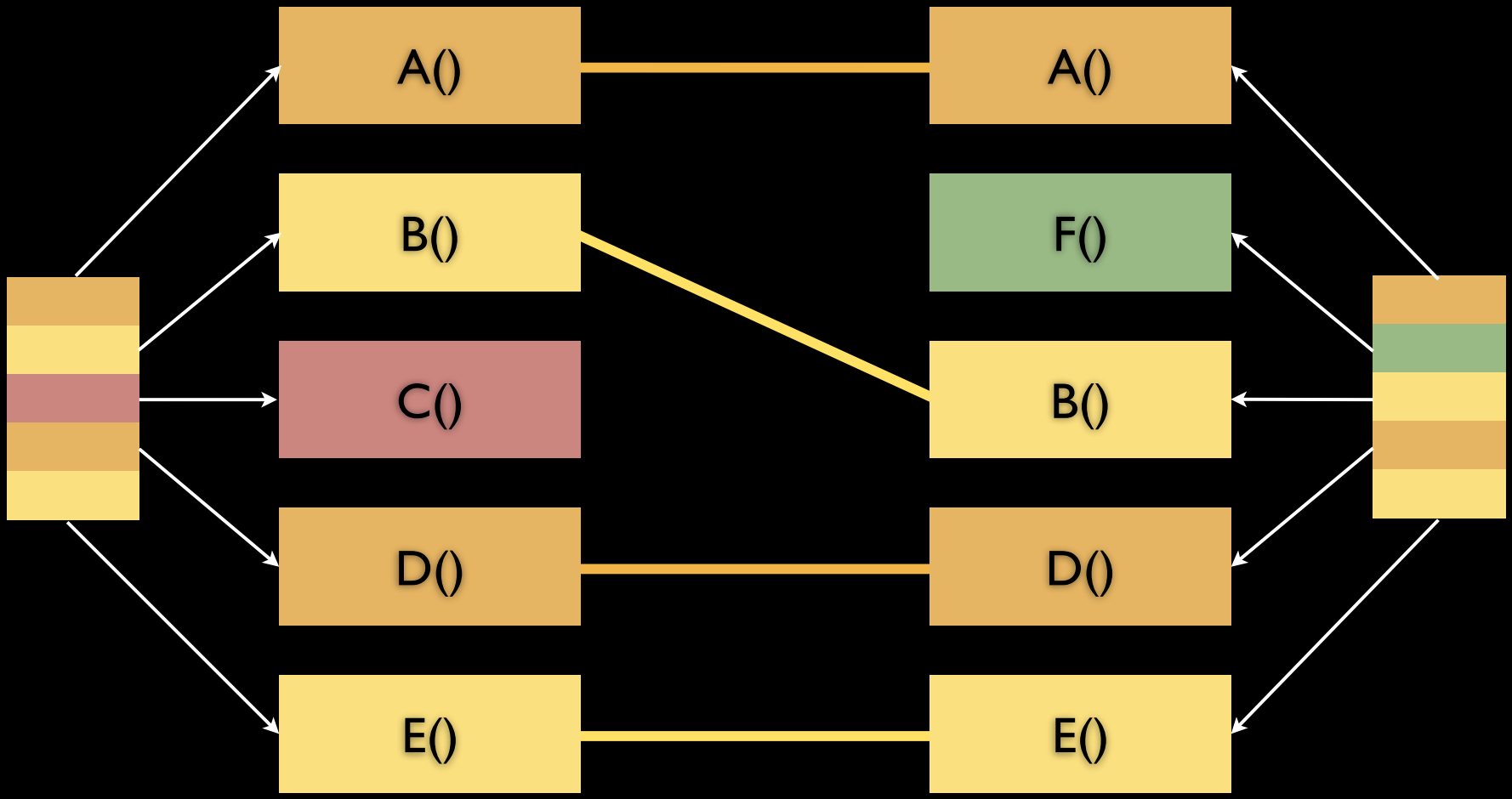

Java – ComparePreferencePage.java – Eclipse Platform
Package Explorer
ComparePreferencePage.java
Commit anyway?
There are 2 unchanged entities. Do you really want to commit?
Yes
<< Details
No
Unchanged Entities
Symbol
File
initDefaults(IPreferenceStore store)
ComparePreferencePage.java
org.eclipse.compare/plugin.properties
plugin.properties
eROSE prevents incomplete changes.
Action
Symbol
File
Support
Confidence
CHG ~initDefaults(IPreferenceStore) .../internal/ComparePreferencePage.java 11 1.0
CHG plugin.properties /org.eclipse.compare/plugin.properties 10 0.9090
CHG buildnotes_compare.html /org.eclipse.compare/buildnotes_compare.html 8 0.7272
CHG ~createGeneralPage(Composite) .../internal/ComparePreferencePage.java 7 0.6363
ADD_TO ComparePreferencePage.java .../internal/ComparePreferencePage.java 7 0.6363
CHG ~TextMergeViewer(Composite, int .../contentmergeviewer/TextMergeViewer.java 6 0.5454
ADD_TO TextMergeViewer.java .../contentmergeviewer/TextMergeViewer.java 6 0.5454
CHG ~createTextComparePage(Compo .../internal/ComparePreferencePage.java 6 0.5454
Writable
Smart Insert
96 : 5

Processing CVS data

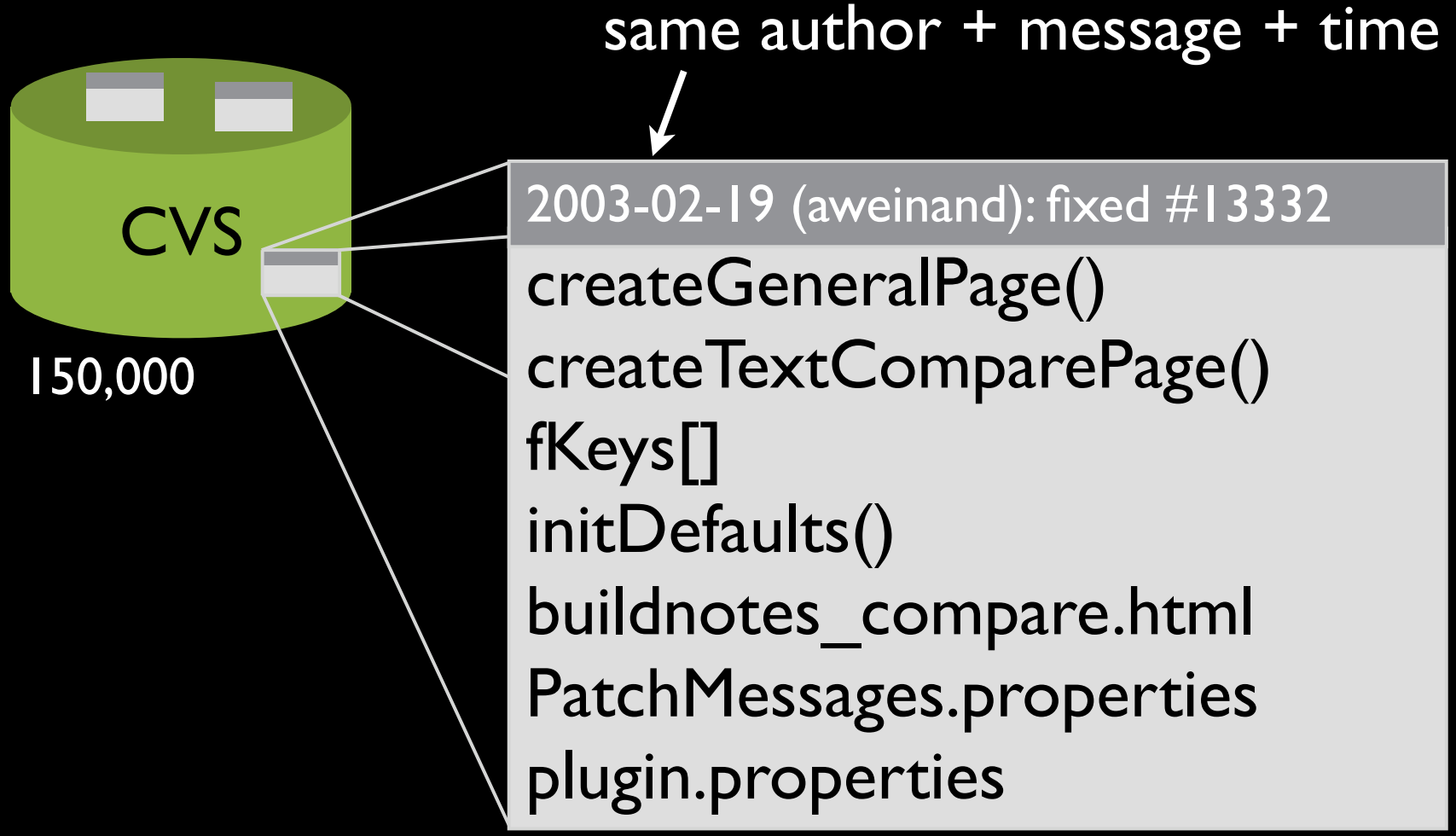


1. Comparing files
2. Building transactions

Comparing Files



Building Transactions



Mining Associations

User changes fKeys[] and initDefaults()

Mining Associations

EROSE finds past transactions

#756 fKeys[] initDefaults() ... plugin.properties	#6721 fKeys[] initDefaults() ... plugin.properties	#21078 fKeys[] initDefaults() ... plugin.properties	
#42432 fKeys[] initDefaults() ... plugin.properties	#51345 fKeys[] initDefaults() ... plugin.properties	#59998 fKeys[] initDefaults() ... plugin.properties	#71003 fKeys[] initDefaults() ... plugin.properties
#87264 fKeys[] initDefaults() ...	#91220 fKeys[] initDefaults() ... plugin.properties	#101823 fKeys[] initDefaults() ... plugin.properties	#104223 fKeys[] initDefaults() ... plugin.properties

Mining Associations

EROSE finds past transactions

#756
fKeys[]
initDefaults()
...
plugin.properties

#6721
fKeys[]
initDefaults()
...
plugin.properties

#21078
fKeys[]
initDefaults()
...
plugin.properties

{fKeys[], initDefaults()} ⇒ {plugin.properties}

Support 10, Confidence 10/11 = 0.909

#87264
fKeys[]
initDefaults()
...

#91220
fKeys[]
initDefaults()
...
plugin.properties

#101823
fKeys[]
initDefaults()
...
plugin.properties

#104223
fKeys[]
initDefaults()
...
plugin.properties

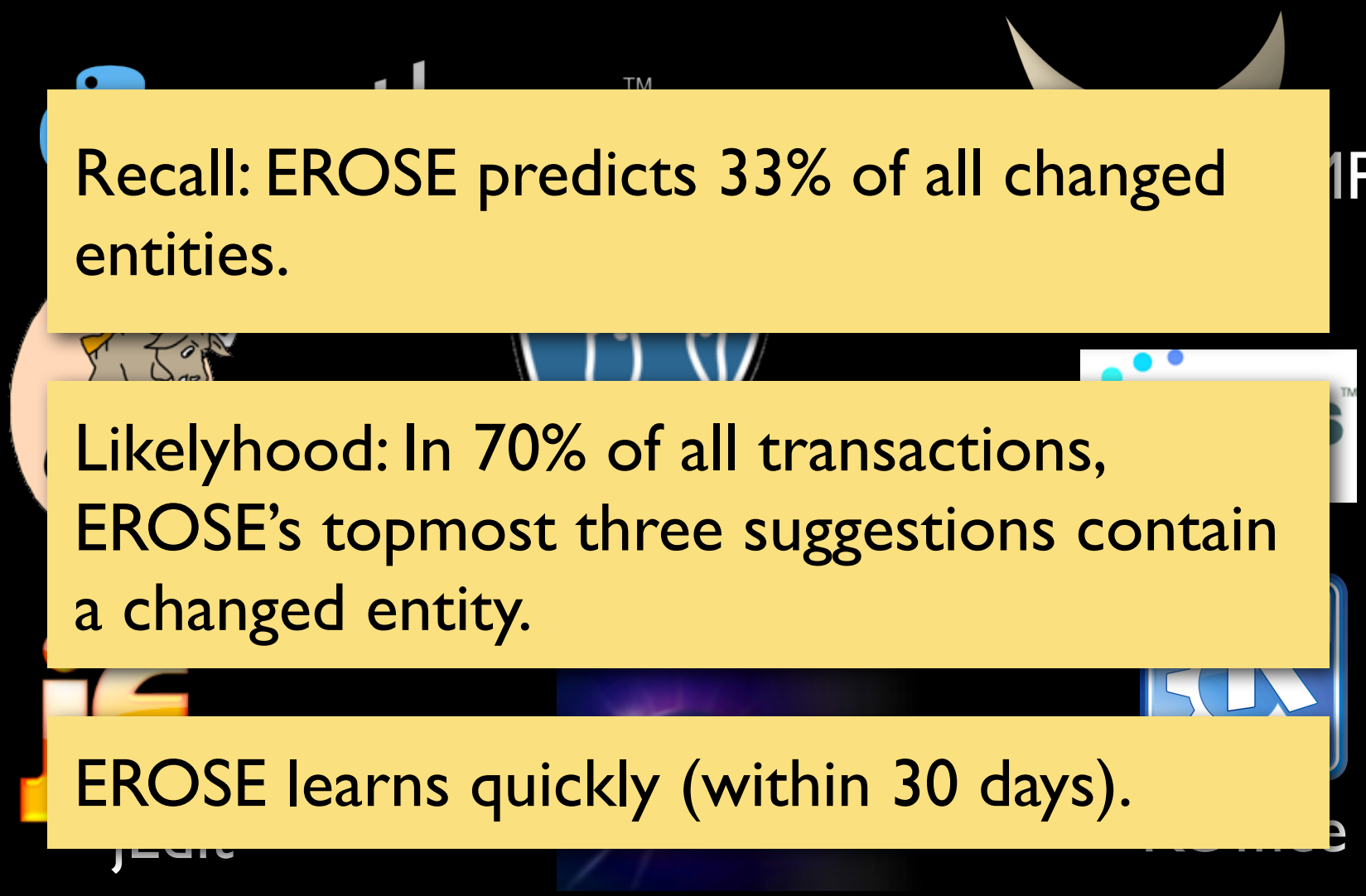
Evaluation: Research Questions

- Given a single change, can ROSE point programmers to entities that should typically be changed, too?
- Does ROSE find the missing change?
- Suppose a transaction is finished, how often does ROSE erroneously suggest that a change is missing?

Evaluation Questions:

- What are differences between course-grained vs. fine-grained suggestions?
- How well does ROSE perform if it is applied to changes without add and delete?
- What are the actual benefit of "add_to" and "del_from" items?
- How much of the version history does ROSE need?
- Would focusing on recent changes improve the quality of recommendations?

Evaluation



Recall: EROSE predicts 33% of all changed entities.

Likelyhood: In 70% of all transactions, EROSE's topmost three suggestions contain a changed entity.

EROSE learns quickly (within 30 days).

Evaluation Measure

Precision:
a fraction of the returned items that were expected

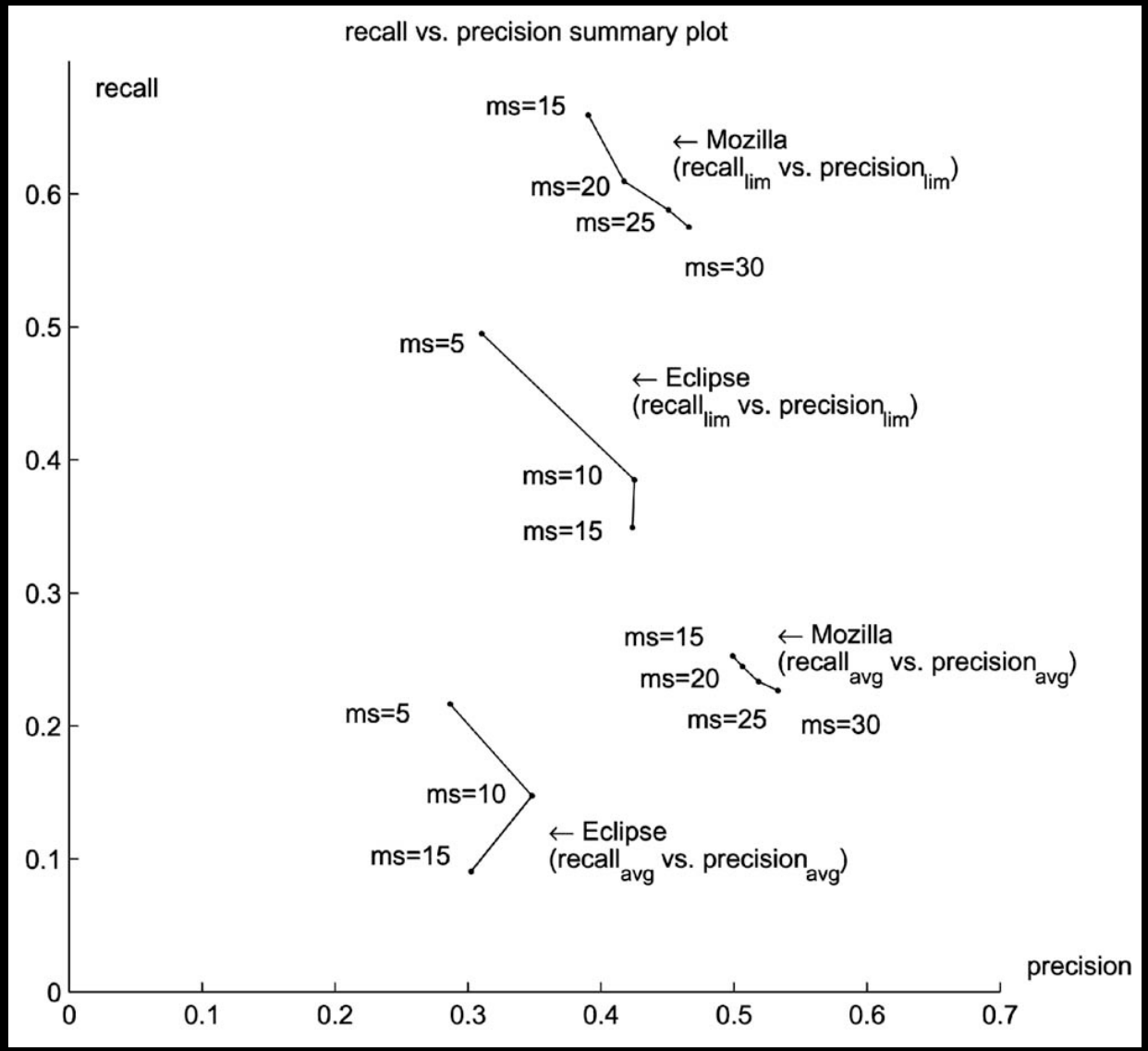
$$P_q = \frac{|A_q \cap E|}{|A_q|},$$

Recall:
a fraction of expected items that were returned

$$R_q = \frac{|A_q \cap E|}{|E|}.$$

Where Aq is a set of items recommended by querying with q.
E is a set of items in the evaluation data (ground truth)

Precision vs. Recall



Evaluation Measure

Feedback: |Z*|/|Z|

the percentage of queries where eRose makes at least one recommendation

$$Z^* = \{q \mid q = (Q, E) \in Z, \texttt{apply}_{R_1}(Q) \neq \emptyset\},$$

Likelihood:

a probability that at least one of the top k recommendations for a query is correct

$$L_k = \frac{\left|\{q \mid q = (Q, E) \in Z, \texttt{apply}_{R_k}(Q) \cap E \neq \emptyset\}\right|}{\left|\{q \mid q = (Q, E) \in Z, \texttt{apply}_{R_k}(Q) \neq \emptyset\}\right|}$$

Where Aq is a set of items recommended by querying with q.
E is a set of items in the evaluation data (ground truth)

Quiz

- You cannot discuss your solution with your classmates
- It will be graded (scale of 0-3)

My general thoughts on eRose & Recap

- eRose uses association rule mining to identify related code elements from a version history data.
- The approach & idea is very novel, though the results are not very impressive.
- One of the first practical system that recovers institutional knowledge from history data
- Trade offs between precision vs. recall is thoroughly investigated.

Announcement

- Project Checkpoint Due on this thursday.
  - I won't grade them.
  - It is not mandatory.
  - You are encouraged to submit to seek my feedback.
  - Available for both research project, literature survey, and tool evaluation

Preview for Next Monday

- Davor Cubranic and Gail C. Murphy. "Hipikat: recommending pertinent software development artifacts". In ICSE '03: Proceedings of the 25th International Conference on Software Engineering, pages 408–418, Washington, DC, USA, 2003. IEEE Computer Society.
- Focus on how they integrated heterogeneous software artifact repositories
- Look at their user study design: what else would you have done to evaluate this system?
- If time permits, briefly go over BugCache (S.Kim et al. ICSE 2007) & Social Structure Mining (C. Bird et al. FSE 2008) papers.