A Dataset from Change History to Support Evaluation of Software Maintenance Tasks

Bogdan Dit
Andrew Holtzhauer
Denys Poshyvanyk
Huzefa Kagdi

MSR 2013
San Francisco, CA
Our Datasets...

• Support the **evaluation** of software maintenance tasks
  - Textual description of maintenance task
  - Changed methods (gold sets)
  - Execution traces
Our Datasets...

- Have already been evaluated in research papers:
  - Feature location
    - [ICSE'13][EMSE'13][EMSE'12][WCRE'12][TR'12][ICPC’11]
  - Change impact analysis
    - [ICSE'12]
  - Developer recommendation
    - [ICSM’12]
  - Traceability link recovery
    - [TSE'12]
Our Datasets...

+ Tools

http://www.cs.wm.edu/semeru/data/msr13/
Choose systems...
Choose systems...
Choose systems...
Choose systems...

Link

SUBVERSION

Bugzilla

trac
Choose systems...

“Link”

- Subversion
- Bugzilla
- Trac
- Java

Allows to collect feature/bug specific execution traces
Jump to revision: 14897  
Author: shlomy  
Date: Wed Apr 8 12:46:23 2009 UTC (4 years, 1 month ago)  
Changed paths: 2  
Log Message: Fixed bug #1593576: Autoindenting should copy exact whitespaces.

<table>
<thead>
<tr>
<th>Path</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="https://example.com">JEdit/trunk/org/jt/sp/jedit/buffer/JEditBuffer.java</a></td>
<td>modified, text changed</td>
</tr>
<tr>
<td><a href="https://example.com">JEdit/trunk/org/jt/sp/util/StandardUtilities.java</a></td>
<td>modified, text changed</td>
</tr>
</tbody>
</table>
Commit log message

Fixed bug #1593576: Autoindenting should copy exact whitespaces.

<table>
<thead>
<tr>
<th>Path</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>jEdit/trunk/org/jgit/sp/edit/buffer/JEditBuffer.java</td>
<td>modified, text changed</td>
</tr>
<tr>
<td>jEdit/trunk/org/jgit/sp/util/StandardUtilities.java</td>
<td>modified, text changed</td>
</tr>
</tbody>
</table>
Jump to revision: 14897
Author: shlomy
Date: Wed Apr 8 12:46:23 2009 UTC (4 years, 1 month ago)
Changed paths: 2
Log Message: Fixed bug #1593576: Autoindenting should copy exact whitespaces.

<table>
<thead>
<tr>
<th>Path</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>jEdit/trunk/org/jt/sp/jedit/buffer/JEditBuffer.java</td>
<td>modified, text changed</td>
</tr>
<tr>
<td>jEdit/trunk/org/jt/sp/util/StandardUtilities.java</td>
<td>modified, text changed</td>
</tr>
<tr>
<td>Path</td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>jEdit/trunk/org/jgit/sp/jedit/buffer/JEditBuffer.java</td>
<td></td>
</tr>
<tr>
<td>jEdit/trunk/org/jgit/sp/util/StandardUtilities.javassist</td>
<td></td>
</tr>
</tbody>
</table>

**Details:**
If autoindenting copies the indentation of last line the indentation should be done the same way it was copied from, e.g. currently if Soft tabs are enabled, always spaces are inserted, even if the previous line is indented with literally tabs or tabs and spaces mixed. The whitespaces should simply be copied from the line that the auto-indenting is taken from.

Sorry if this is a dupe, but I have many bugs to post and am too lazy to check them all for dupes currently. :-(
Link to issue tracking system

Manually verified for correct links

Jump to revision: 14897
Author: shlomy
Date: Wed Apr 8 12:46:23 2009 UTC (4 years, 1 month ago)
Changed paths: 2
Log Message: Fixed bug #1593576: Autoindenting should copy exact whitespaces.

Details: If autoindenting copies the indentation of last line the indentation should be done the same way it was copied from, e.g. currently if Soft tabs are

Sorry if this is a dupe, but I have many bugs to post and am too lazy to check them all for dupes currently. :-}
Autoindenting should copy exact whitespaces  ID: 1593576

Details:

If autoindenting copies the indentation of last line the indentation should be done the same way it was copied from, e.g. currently if Soft tabs are enabled, always spaces are inserted, even if the previous line is indented with literally tabs or tabs and spaces mixed. The whitespaces should simply be copied from the line that the auto-indenting is taken from.

Sorry if this is a dupe, but I have many bugs to post and am too lazy to check them all for dupes currently. :-(
Description used to collect execution traces

Textual description of maintenance task (query)
Jump to revision: 14897
Author: shlomy
Date: Wed Apr 8 12:46:23 2009 UTC (4 years, 1 month ago)
Changed paths: 2
Log Message: Fixed bug #1593576: Autoindenting should copy exact whitespaces.

<table>
<thead>
<tr>
<th>Path</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>modified, text changed</td>
</tr>
<tr>
<td>jEdit/trunk/org/jgit/sp/jedit/buffer/JEditBuffer.java</td>
<td>modified, text changed</td>
</tr>
<tr>
<td>jEdit/trunk/org/jgit/sp/util/StandardUtilities.java</td>
<td>modified, text changed</td>
</tr>
</tbody>
</table>
### Changed files

<table>
<thead>
<tr>
<th>Path</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>JEdit/trunk/org/jgit/sp/jedit-buffer/JEditBuffer.java</td>
<td>modified, text changed</td>
</tr>
<tr>
<td>JEdit/trunk/org/jgit/sp/util/StandardUtilities.java</td>
<td>modified, text changed</td>
</tr>
</tbody>
</table>

**Author:** shlomy

**Date:** Wed Apr 8 12:46:23 2009 UTC (4 years, 1 month ago)

**Log Message:** Fixed bug #1593576: Autoindenting should copy exact whitespaces.
Examine file changes

<table>
<thead>
<tr>
<th>Path</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>jEdit/trunk/org/git/sp/jeditbuffer/JEditBuffer.java</td>
<td>modified, text changed</td>
</tr>
<tr>
<td>jEdit/trunk/org/git/sp/util/StandardUtilities.java</td>
<td>modified, text changed</td>
</tr>
</tbody>
</table>
/**
 * Indents the specified line.
 * @param lineIndex The line number to indent
 * @param canDecreaseIndent If true, the indent can be decreased as a result of this. Set this to false for Tab key.
 * @return true if indentation took place, false otherwise.
 * @since jEdit 4.2pre2
 */

public boolean indentLine(int lineIndex, boolean canDecreaseIndent)
{
    int[] whitespaceChars = new int[1];
    int currentIndent = getCurrentIndentForLine(lineIndex, whitespaceChars);

    int realIndent = getRealIndentForLine(lineIndex);
    int idealIndent = getCurrentIndentForLine(lineIndex);

    // Do it
    return false;
}
/**
 * Indents the specified line.
 * @param lineIndex The line number to indent
 * @param canDecreaseIndent If true, the indent can be decreased as a result of this. Set this to false for Tab key.
 * @return true If indentation took place, false otherwise.
 * @since jEdit 4.2pre2
 */
public boolean indentLine(int lineIndex, boolean canDecreaseIndent) {
    int[] whitespaceChars = new int[1];
    int currentIndent = getCurrentIndentForLine(lineIndex, whitespaceChars);

    int idealIndent = getIdealIndentForLine(lineIndex);

    if(idealIndent == -1 || idealIndent == currentIndent 
        || (!canDecreaseIndent && idealIndent < currentIndent))
        return false;

    // Do it
}

/**
 * Indents the specified line.
 * @param lineIndex The line number to indent
 * @param canDecreaseIndent If true, the indent can be decreased as a result of this. Set this to false for Tab key.
 * @return true If indentation took place, false otherwise.
 * @since jEdit 4.2pre2
 */
public boolean indentLine(int lineIndex, boolean canDecreaseIndent) {
    int[] whitespaceChars = new int[1];
    int currentIndent = getCurrentIndentForLine(lineIndex, whitespaceChars);

    int prevLine = PriorNonEmptyLine(lineIndex);
    int prevPrevLine = (prevLine <= 0) ? prevLine : PriorNonEmptyLine(prevLine);

    if (prevLine == (lineIndex - 1) 
        || prevPrevLine == (lineIndex - 2))
        return false;

    // Do it
Our Datasets...

ArgoUML

jEdit

muCommander

JabRef reference manager
Our Datasets...

- 633 textual descriptions of issues
- 633 execution traces (manually collected)
- >4K gold set methods
Thank you!
Questions?

http://www.cs.wm.edu/semeru/data/msr13/