# VisCad User Guide

# 1 Introducing VisCad

	1.1	Introduction	4	
	1.2	System Requirement	5	
	1.3	Obtaining and Running VisCad	6	
	1.4	Importing clone detection result	7	
	1.5	User Interface Components	18	
	1.6	Analyze Clone Fragments	23	
	1.7	Analyze Clone Files	27	
2	Visu	/isualization		
	2.1	Introduction	31	
	2.2	Scatter Plot	32	
	2.3	Treemap	36	
	2.4	Hierarchical Dependency Graph	41	
3	Code	Code Clone Metrics		
	3.1	Introduction	48	
	3.2	Obtaining Metrics For Files	49	
	3.3	Obtaining Metrics For Directories	55	
	3.4	Obtaining Metrics For Clone Classes	58	
4	Filte	Filtering		
	4.1	Introduction	62	
	4.2	Overlapping Clone filterng	63	
	4.3	Textual filtering	68	

# **Introducing VisCad**

#### Introduction

Detection and analysis of similar code fragments ("code clones") has become an integral part of software maintenance. In response, over the last decade a great many clone detection techniques and tools have been proposed. However, identifying useful cloning information from the large volume of textual data produced by these detectors is challenging. VisCad is a tool with which a user can visualize and analyze large volumes of raw cloning data in an interactive fashion. Users can analyze and identify distinctive code clones through a set of visualization techniques, metrics and data filtering operations. The loosely coupled architecture of VisCad allows users to work with the clones of any clone detection tools that report source co-ordinates of the found clones. This yields the opportunity to work with the clone detectors of choice, which is important for clone analysis since clone detectors have their own strengths and weaknesses.

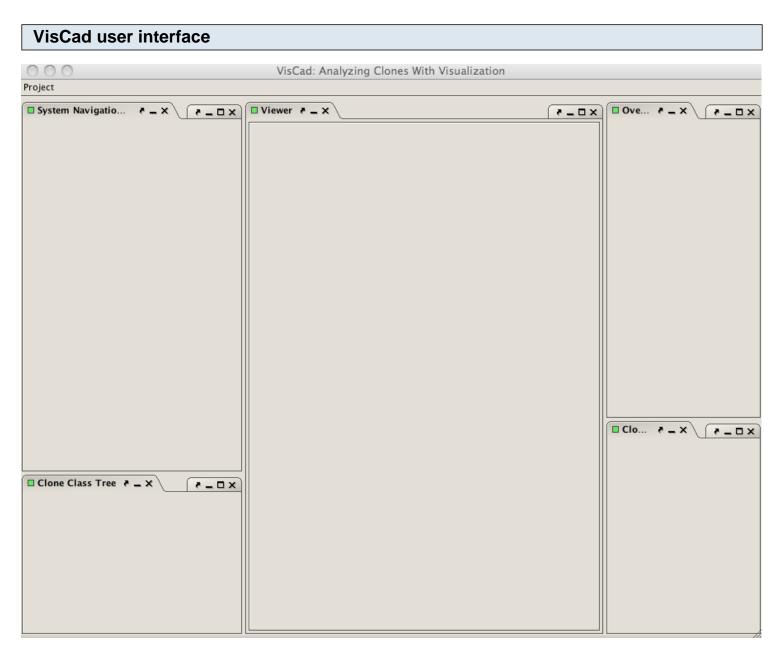
# **System Requirement**

- 1. VisCad requires Java Runtime Environment (JRE) 6 or later. You can download the recent JRE from here .
- 2. We have successfully tested VisCad on Windows XP, Windows 7, Max OS X( version 10.6) and on some Linux distribution (such as Ubuntu).
- 4. The more RAM your computer has, the better performance you gain from VisCad. However, 2.0 GB or more is recommended.
- 5. Display dimension of 1024 x 768 or greater is recommended.

# **Obtaining and Running VisCad**

Follow the steps listed below:

- 1. Download the VisCad\_beta.zip file from here . You can obtain the most recent version of VisCad, documentation, source code from this location.
- 2. Extract the contents of the archive.
- 3. Double click on the VisCad.jar file to run the program.



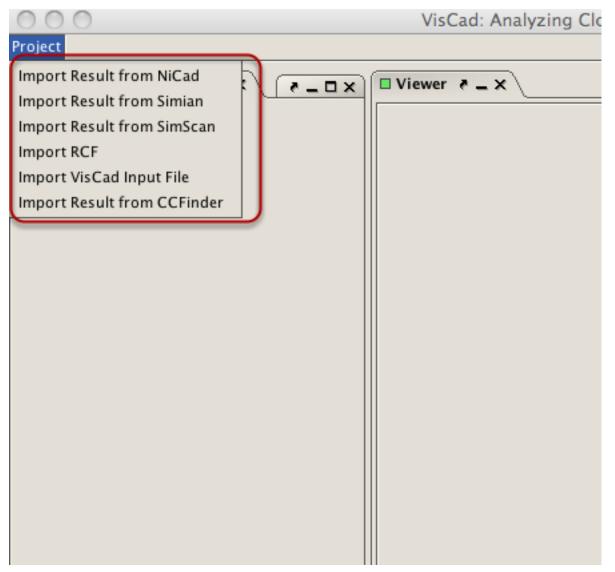
Running the viscad.jar file opens the above window.

### Importing clone detection result

VisCad requires the subject system and the clone detection result you obtained by running clone detection on the subject system using the supported clone detectors.

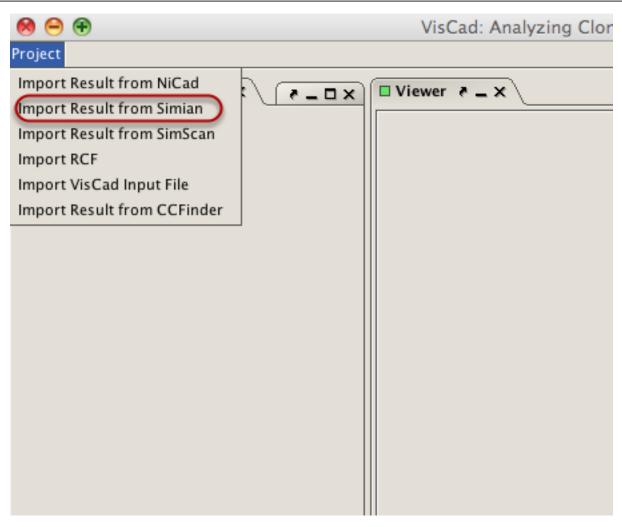
Current you can directly import clone detection result of CCFinder, Simian, SimScan, NiCad. If you have clone detection result in RCF format, you can also import and anlyze the data in VisCad. For other clone detection tools, you need to convert the result into VisCad input file format. VisCadBeta.zip file contains an example of VisCad input file.

#### Make the selection



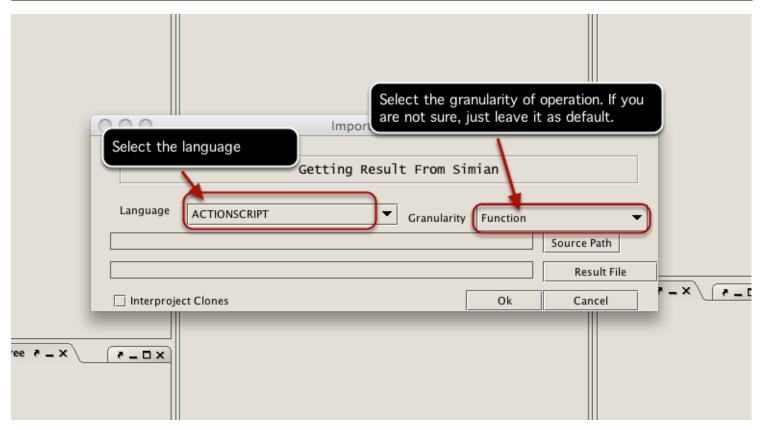
To import clone detection result or RCF data file, click on the *Project* menu. This opens a list of importing operations supported by VisCad.

# Import clone detection result of Simian



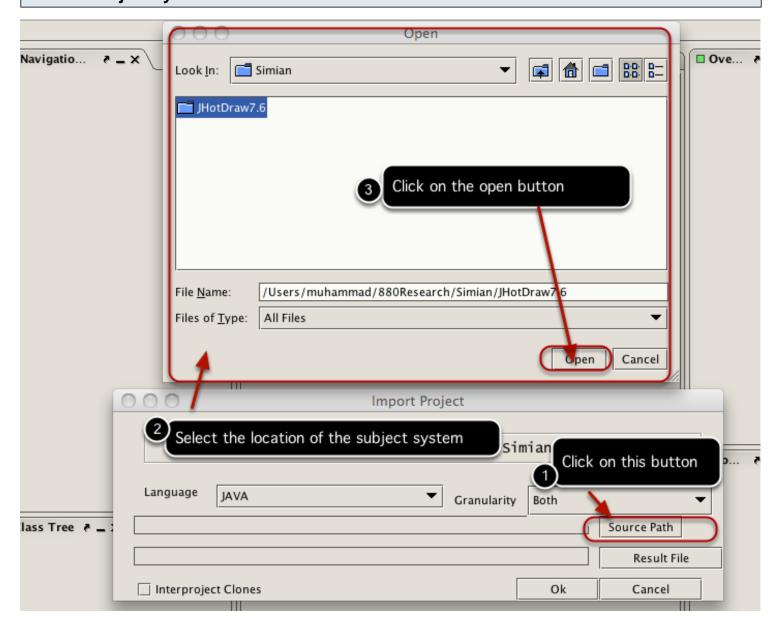
Suppose your computer contians the source code of JHotDraw(version 7.6) and located in the JHotDraw7.6 directory. You detect the clones using Simian on that directory and stores the result in jhotdrawResultSimian.txt file. To import the result, you need to click on the Project menu and select the *Import Result from Simian* menu item.

# Importing clone detection result of Simian( Continued)



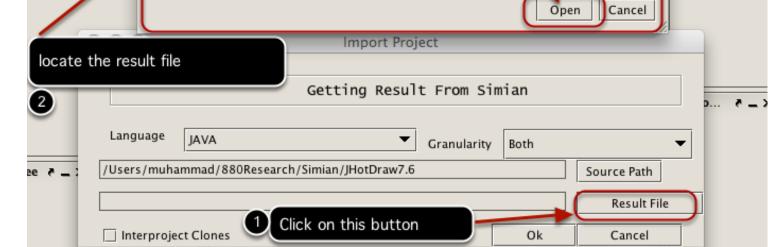
This opens the above dialog box.

# Select subject system



# Select the clone detection result file Open Look In: Simian JHotDraw7.6 jhotdrawResultSimian.txt

Click on the open button



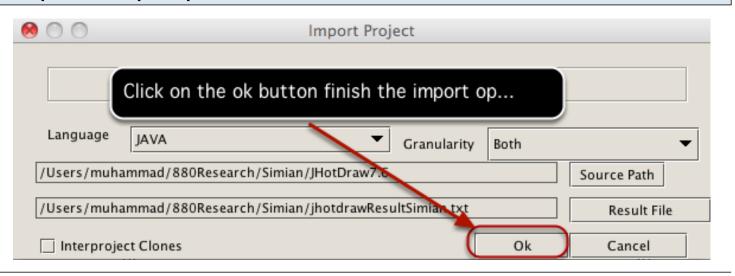
jhotdrawResultSimian.txt

All Files

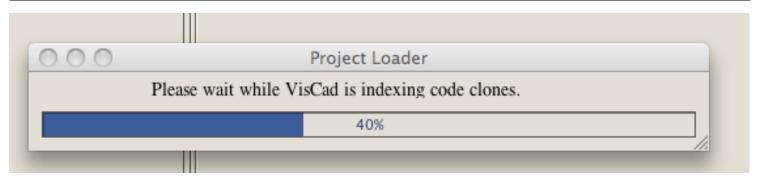
# Complete the import opeation

File Name:

Files of Type:

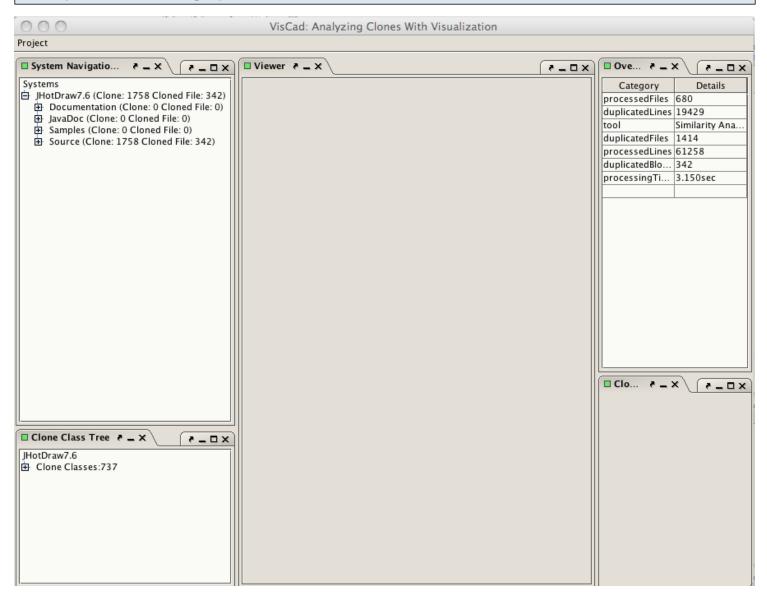


# Loading clone detection result



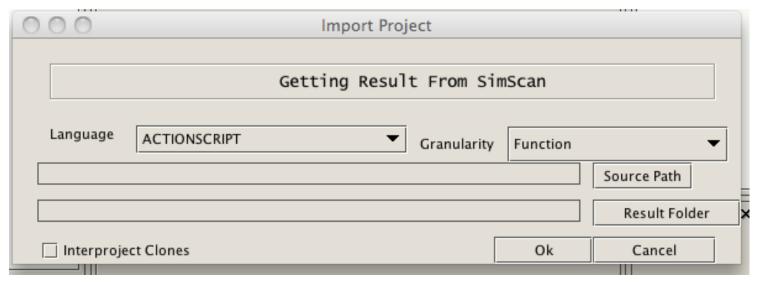
VisCad is now loading the clone detection result.

# Complete the loading operation



The result is loaded into VisCad.

#### Import result from SimScan

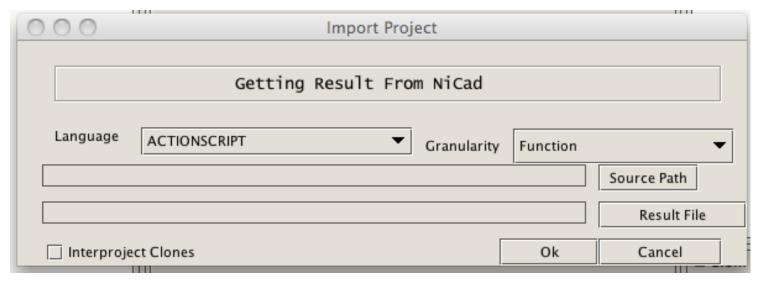


SimScan creates a folder that contains the output of the clone detection. A sample clone detection result (jhotdrawReportSimScan.zip) for JHotDraw (Version 7.6) is included in the VisCadBeta.zip file. You can also find the source code of the subject system in a zip file (jhotdraw-7.6.nested.zip).

#### Follow the steps listed below:

- 1. Unzip the jhotdraw-7.6.nested.zip file and rename the folder to JHotDraw7.6.
- 2. Unzip the jhotdrawReportSimScan.zip file. This will create a folder (jhotdrawReportSimScan) that contains the clone detection results. Put the folder in the same place where JHotDraw7.6 folder resides.
- 3. Select the 'Import Result from SimScan' menu item from the 'Project' menu (located on the top left of the VisCad interface). This opens a dialog to import the result.
- 4. Click on the Source Path button and select the source code directory (In this case, JHotDraw7.6 folder). This should be the folder on which you apply clone detection.
- 5. Now, click on the Result Folder button and select the clone detection result folder (In this case, jhotdrawReportSimScan).
- 6. Select the language of the subject system and also the granularity of clone detection. If you are not sure, you can leave them as default.
- 7. Now, click on the Ok button to load the subject system by VisCad. The loading time may vary depending on the size of the system.
- 8. You can click on the Cancel button to cancel the import operation.

#### Import result from NiCad

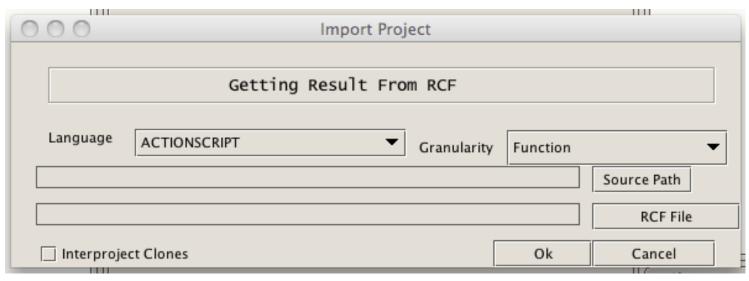


A sample clone detection result (JHotDraw7.6\_functions-clones-0.3.xml) for JHotDraw (Version 7.6) is included in the zip file. You can also find the source code of the subject system in a zip file (jhotdraw-7.6.nested.zip). Clones were detected using NiCad.

Follow the steps listed below:

- 1. Unzip the jhotdraw-7.6.nested.zip file and rename the folder to JHotDraw7.6.
- 2. Put the clone detection result file in the same directory where JHotDraw7.6 folder resides.
- 3. Select the 'Import Result from NiCad' menu item from the 'Project' menu (located on the top left of the VisCad interface). This opens a dialog to import the result.
- 4. Click on the Source Path button and select the source code directory (In this case, JHotDraw7.6 folder). This should be the folder on which you apply clone detection.
- 5. Now, Click on the Result File button and select the clone detection result file (In this case, JHotDraw7.6\_functions-clones-0.3.xml).
- 6. Select the language of the subject system and also the granularity of clone detection. If you are not sure, you can leave them as default.
- 7. Now, click on the Ok button to load the subject system by VisCad. The loading time may vary depending on the size of the system.
- 8. You can click on the Cancel button to cancel the import operation.

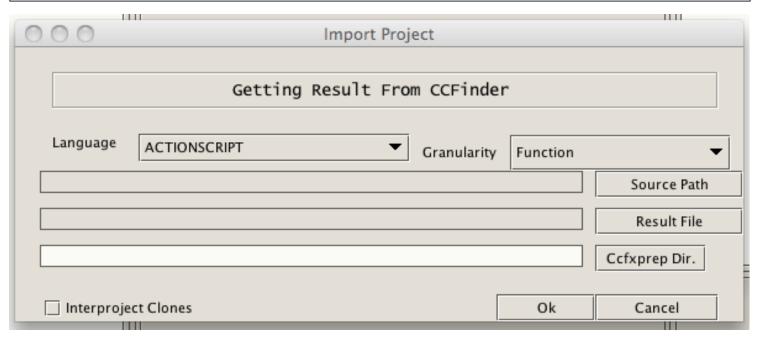
#### Import RCF file



RCF is a data format that can store clone data for a number of versions of a subject system. For this example, we have used GNU Wget. Download both the subject system and the RCF file from here. The RCF file is packaged with the source code of the subject system. Follow the steps listed below:

- 1. Extract the content of the archive and rename the folder to rcfResult. You now have two folders inside the rcfResult folder. One is the wget that contains the source code of seven different versions of the system. Another one is the wget.rcf that contains the clone detection results of those systems in the RCF format.
- 2. Select the 'Import RCF' menu item from the 'Project' menu (located on the top left of the VisCad interface). This opens a dialog to import the result.
- 3. Click on the Source Path button and select the wget folder located inside the rcfResult folder.
- 4. Now, click on the RCF File button and select the wget.rcf file.
- 5. Select the language of the subject system and also the granularity of clone detection. If you are not sure, you can leave them as default.
- 6. Now, click on the Ok button to load all seven-subject subjects by VisCad. The loading time may vary depending on the number of subject systems.
- 7. You can click on the Cancel button to cancel the import operation.

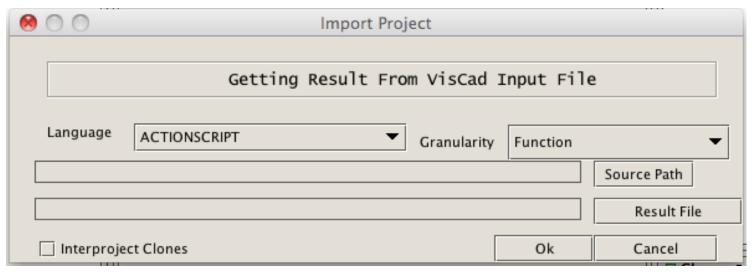
#### Import result from CCFinder



During the clone detection, CCFinder creates a directory (.ccfxprepdir). If you apply clone detection on JHotDraw7.6 directory, then the .ccfxprepdir will be created right inside that directory. Follow the steps listed below:

- 1. Unzip the jhotdraw-7.6.nested.zip file and rename the folder to JHotDraw7.6. Detect clones with CCFinder and then rename the .ccfxprepdir directory to ccfxprepdir.
- 2. The output file is a binary one (such as a.ccfxd). You need to convert that into a text file (see here). From now, we will refer the text file as the result file.
- 3. Put the clone detection result file in the same directory where JHotDraw7.6 folder resides.
- 4. Select the 'Import CCFinder Project' menu item from the 'Project' menu (located on the top left of the VisCad interface). This opens a dialog to import the result.
- 5. Click on the Source Path button and select the source code directory (In this case, JHotDraw7.6 folder). This should be the folder on which you apply clone detection.
- 6. Now, click on the Result File button and select the clone detection result file.
- 7. Click on the Ccfxprep Dir. button and select the ccfxprepdir directory.
- 8. Select the language of the subject system and also the granularity of clone detection. If you are not sure, you can leave them as default.
- 9. Now, click on the Ok button to load the subject system by VisCad. The loading time may vary depending on the size of the system.
- 10. You can click on the Cancel button to cancel the import operation.

#### Import VisCad input file



For any other clone detection tools, the result file needs to be converted to the VisCad input file format. A sample clone detection result (resultViscadInputFile.txt) for JHotDraw (Version 7.6) is included in the VisCadBeta.zip file. Clones were detected using Simian.

Follow the steps listed below:

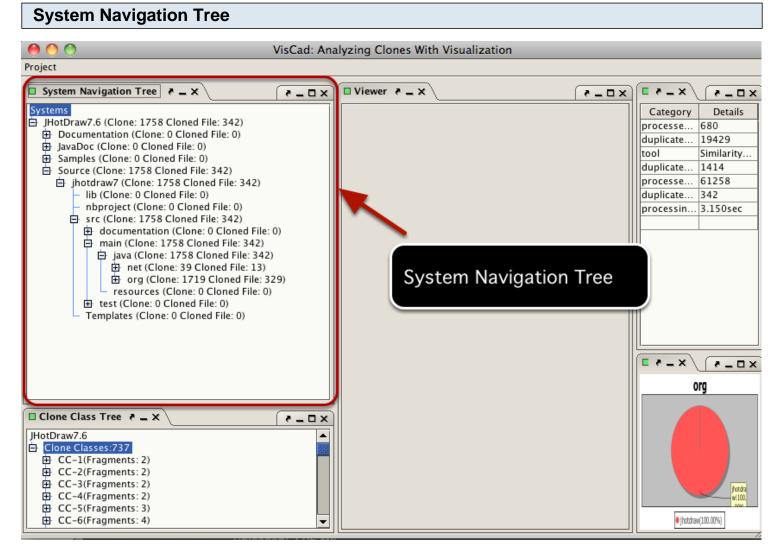
- 1. Unzip the jhotdraw-7.6.nested.zip file and rename the folder to JHotDraw7.6.
- 2. Put the clone detection result file in the same directory where JHotDraw7.6 folder resides.
- 3. Select the 'Import VisCad Input File' menu item from the 'Project' menu (located on the top left of the VisCad interface). This opens a dialog to import the result.
- 4. Click on the Source Path button and select the source code directory (In this case, JHotDraw7.6 folder). This should be the folder on which you apply clone detection.
- 5. Now, click on the Result File button and select the clone detection result file (In this case, resultViscadInputFile.txt).
- 6. Select the language of the subject system and also the granularity of clone detection. If you are not sure, you can leave them as default.
- 7. Now, click on the Ok button to load the subject system by VisCad. The loading time may vary depending on the size of the system.
- 8. You can click on the Cancel button to cancel the import operation.

# **User Interface Components**

The main user interface of VisCad can be divided into three parts.

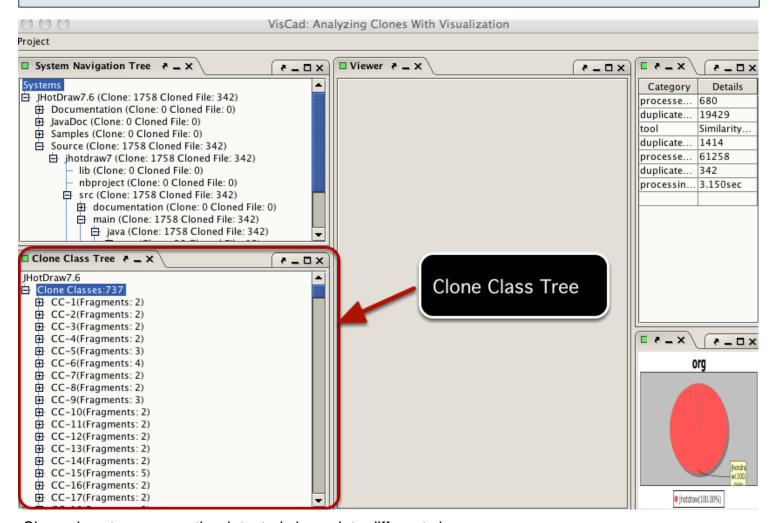
- 1. Left Part: The left part accommodates the *clone browser*. The clone browser has two parts, one of which displays the distribution of clones over the directories and sub- directories in the subject system, known as *System Navigation Tree*. The other part, located on the bottom of the clone browser, lists all *clone classes* and the number of clone snippets in each class, called *Clone Class Tree*.
- 2. Middle Part: The middle part of VisCad accommodates different views in separate tabs. We will refers this part as *Viewer*.
- 3. Right Part: The top- right window shows the clone detection specific information VisCad obtained while parsing the result file for the selected

subject system. For any selected directory in the *system navigation tree*, the bottom-right window shows the distribution of clones in its sub-directories through a pie chart.



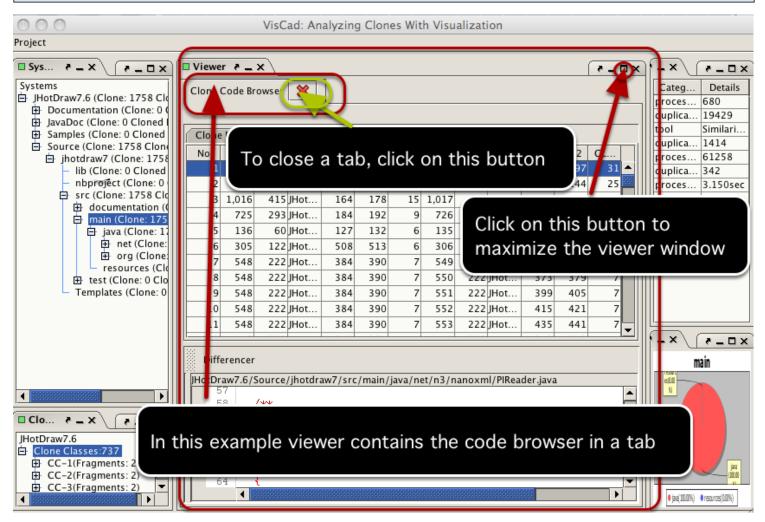
It maps clones to files and directories

#### **Clone Class Tree**



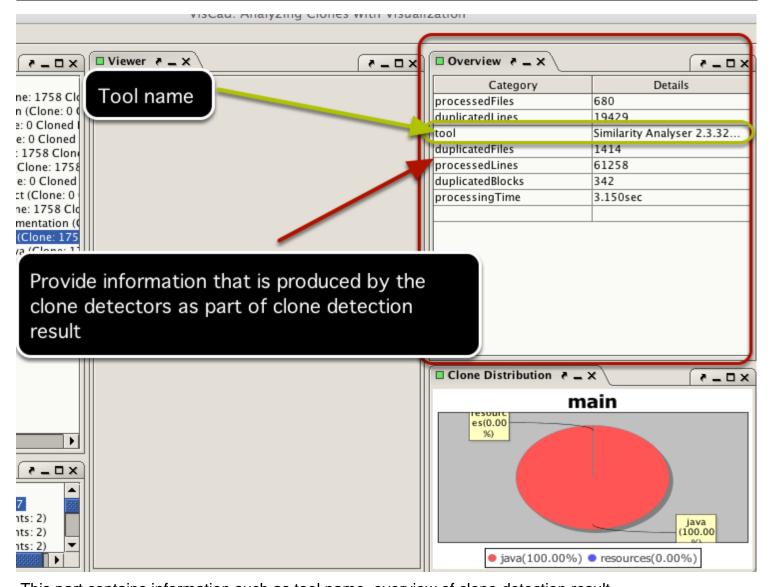
Clone class tree groups the detected clones into different classes.

#### **Viewer**



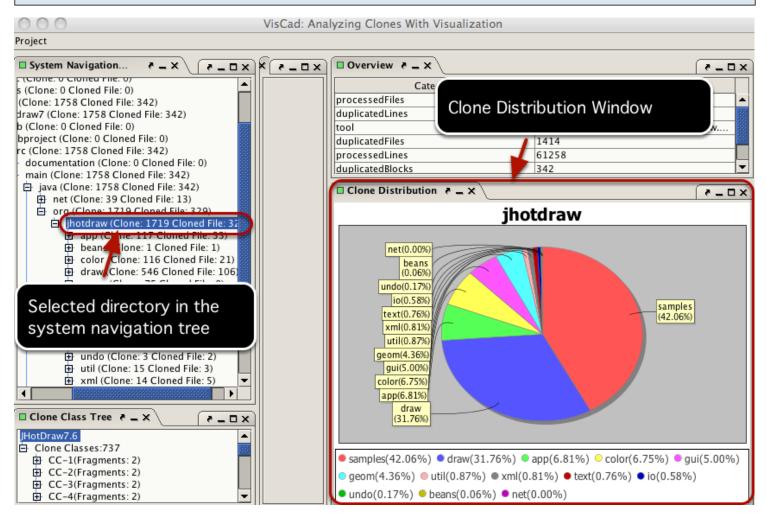
Depending on the user's selection, it accommodates different views(such as the scatter plot, treemap, hierarchical dependency graph, source code browser etc.)

# Obtaining information specific to clone detection



This part contains information such as tool name, overview of clone detection result.

#### Clone distribution window

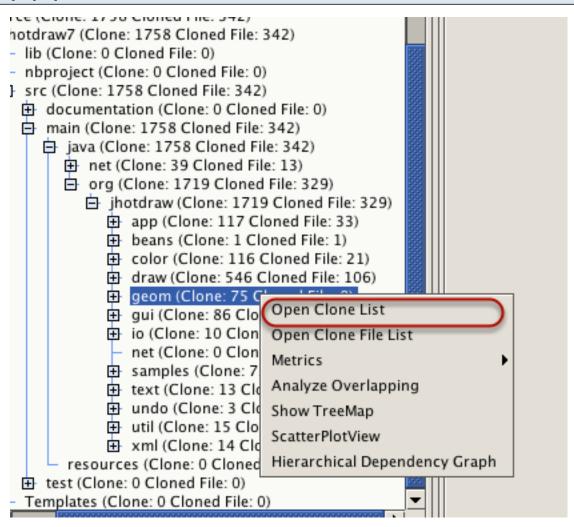


Click on a directory in the system navigation tree to select it. Based on the selection, the clone distribution window updates and shows the distribution of clones in its subdirectories through a pie chart.

# **Analyze Clone Fragments**

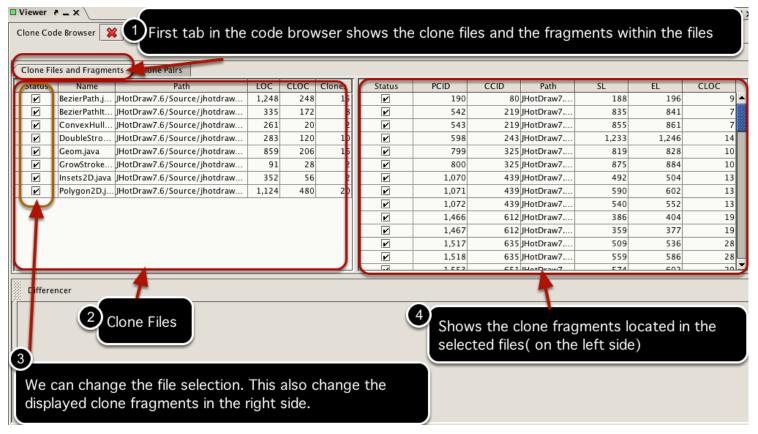
You can analyze the source code of the clone fragments using the *code browser*. The same component is used for analyzing clone code fragments in other places also.

### Open the popup menu and make selection



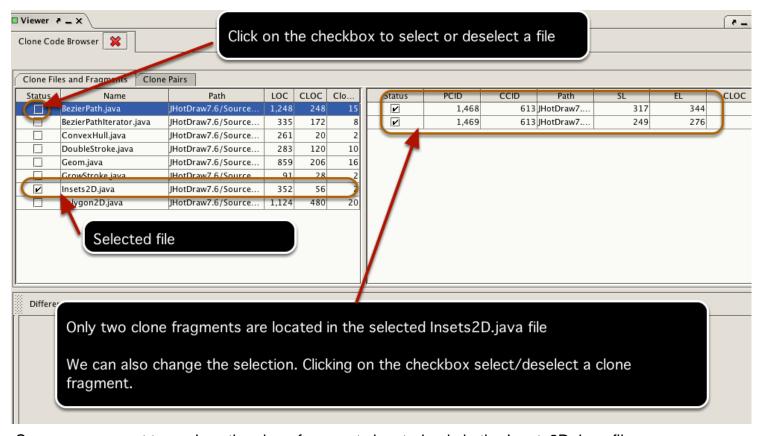
If we want to analyze the clone fragments on the *geom directory*, we need to select the directory, right click on the mouse button to bring up the popup menu and select the *Open clone list* menu item.

#### The code browser



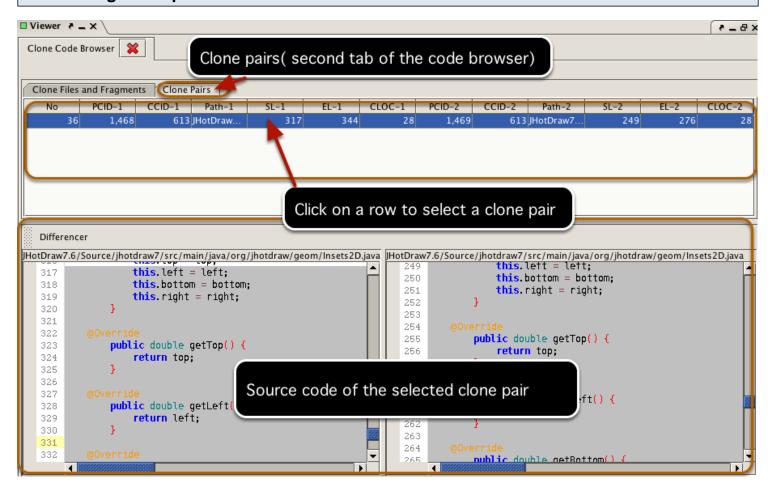
The previous selection open the code browser in the *viewer* window in a new tab.

## **An Example**



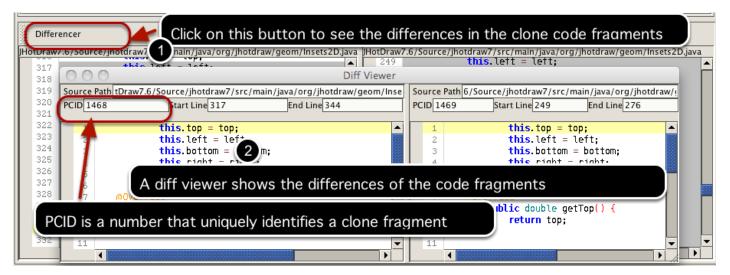
Suppose, we want to analyze the clone fragments located only in the Insets2D. java file. We need to deselect all files except Insets2D.java file by clicking on the checkboxes. From the right side, we can see that only two code fragments are located in this file.

#### **Examining clone pairs**



For the selected clone fragments, the next tab in the code browser shows the clone pairs. Selecting a clone pair also displays the code fragments in the bottom panel.

# Source code difference analysis



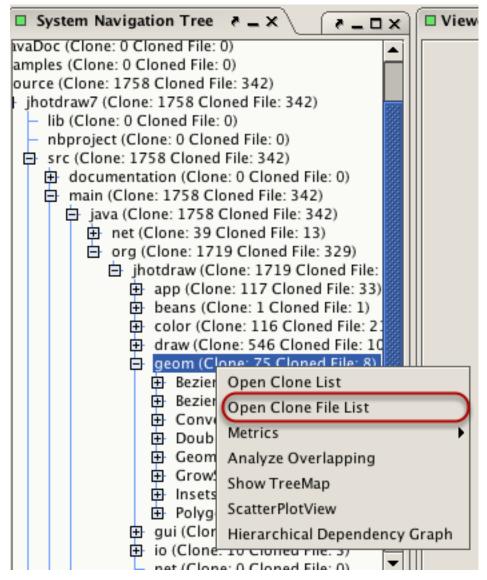
We can understand the source code differences using the diff viewer. In the above figure, the two clone fragments are exact copy of each other.

Similar to PCID, another number(CCID) is used to uniquely identify each clone class.

# **Analyze Clone Files**

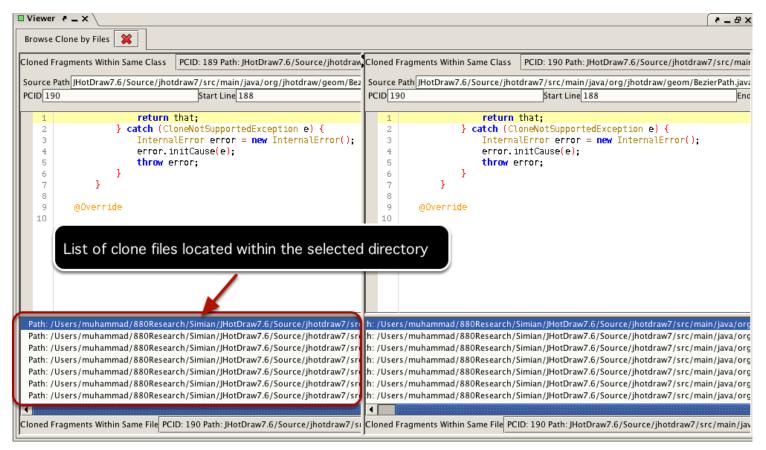
You can use this view to analyze and compare clone files with grouping and selection features.

#### Make selection



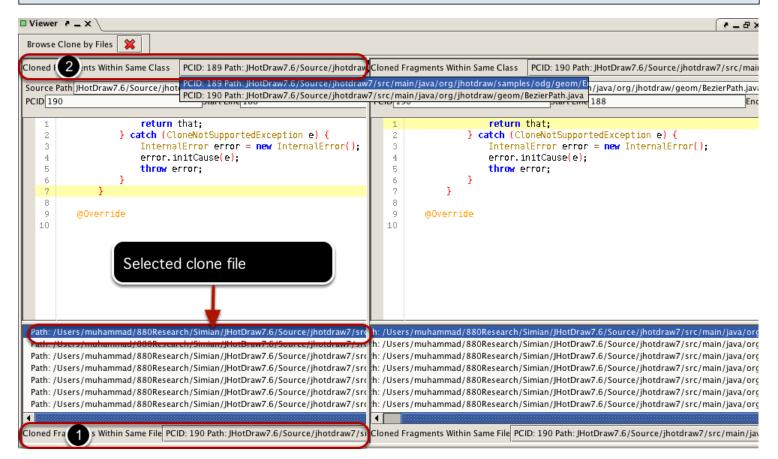
Select a target directory. Right click on the selected directory to open the popup menu and select *Open Clone File List* menu item.

#### List of clone files



We can select a clone file from the list of files.

#### **Features**



This view has some advantages. It groups clone fragments located within the selected file (labelled with 1). For the selected clone fragment, it also groups all clone fragments that falls within the same clone class (labelled with 2).

The left and right part shows the same list of clone files. We can change the selection and compare the related clone fragments side by side.

# **Visualization**

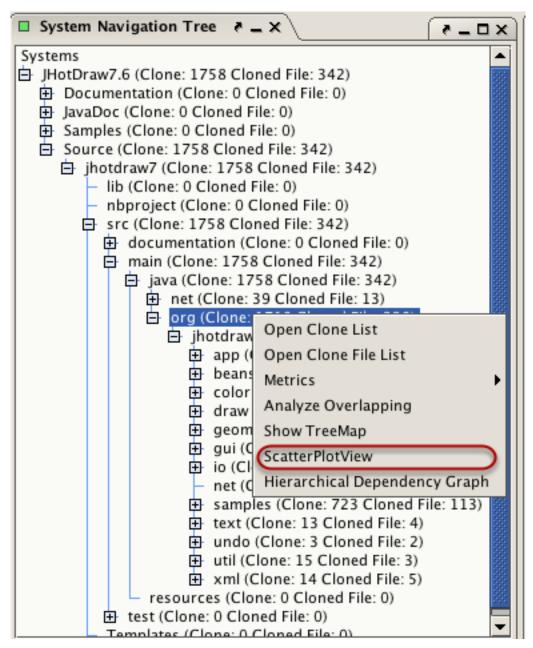
## Introduction

inti oddetion		
Visualization plays an important role in code clone analysis since it can provide high level overview of cloning in a system. At present, VisCad supports three different visualizations which are scatter plot, treemap and hierarchical dependency graph.		

#### **Scatter Plot**

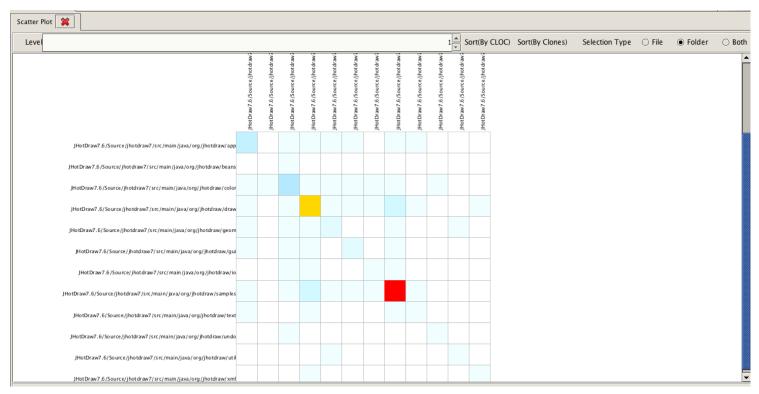
A scatter plot can be viewed as a two dimensional matrix where each cell represents the cloning status between a pair of files or directories. In VisCad, cells render the clone pairs distributed between a pair of files or directories using a color heatmap. Cells are also labelled in the horizontal and vertical axes.

### Opening a scatter plot

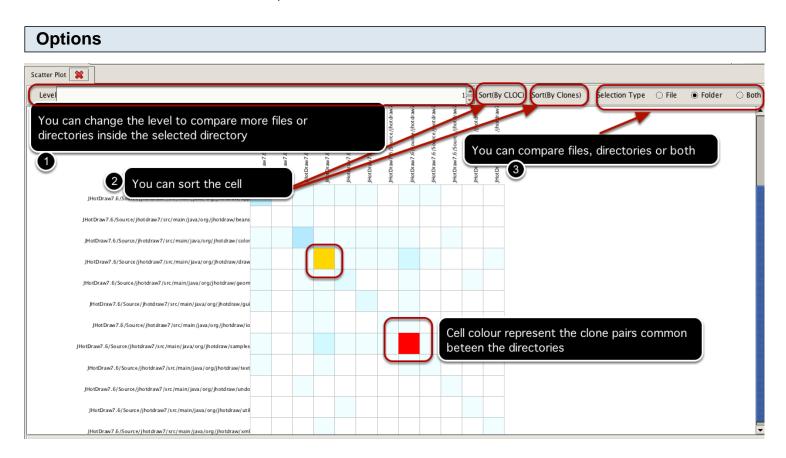


Select a directory from the system navigation tree and right click on it to open a popup menu. Click on the *Scatter Plot View* menu item.

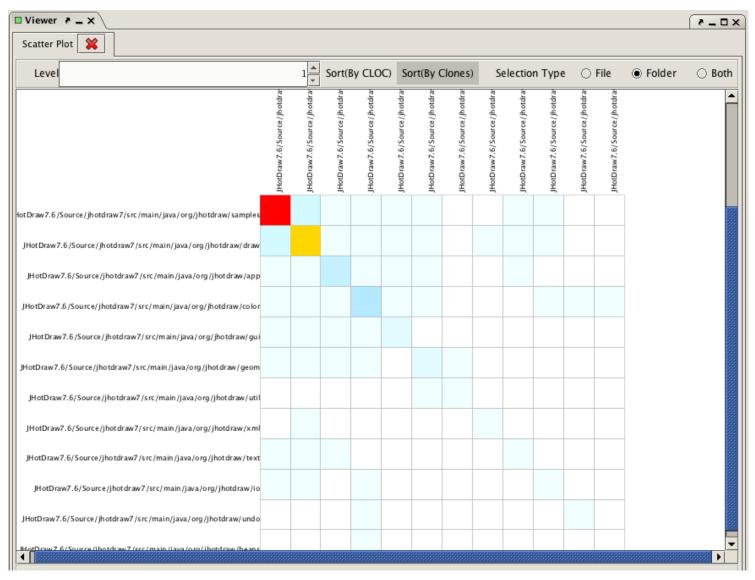
## An example



The previous selection opens the above scatter plot on a new tab in the *viewer* view panel (in the middle of the VisCad user interface).

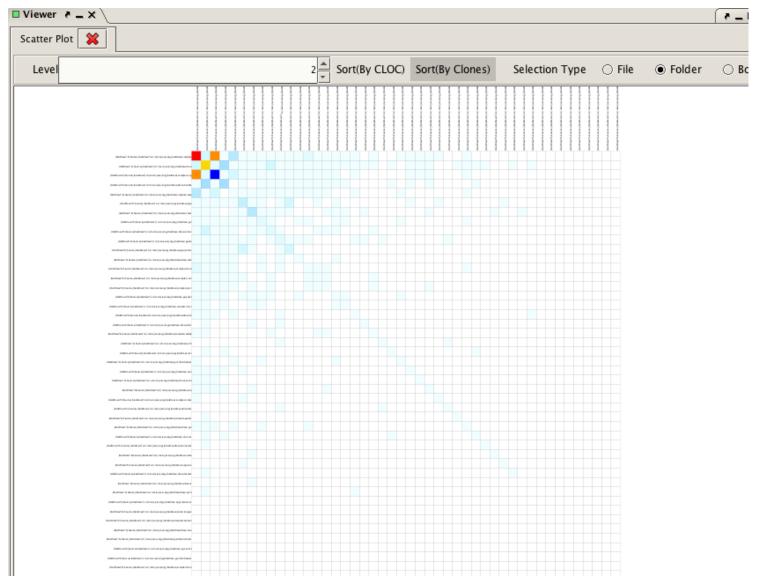


# **Example of Sorting**



Sorting allows to identify cloning patterns easily.

#### Zoom in or zoom out

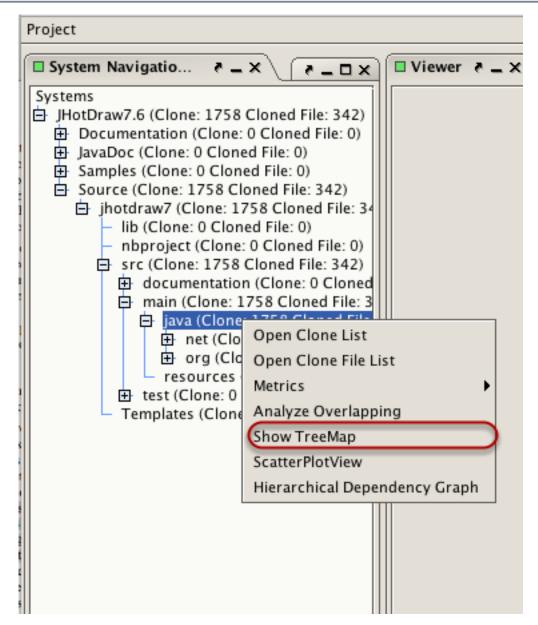


You can hold the right mouse button and move it in the inward or outward direction to perform zoom in or zoom out operations. The above figure shows an example of zoom out operation. You can also identify the files or directories involved with a cell with tooltip by holding the mouse pointer on the cell.

#### **Treemap**

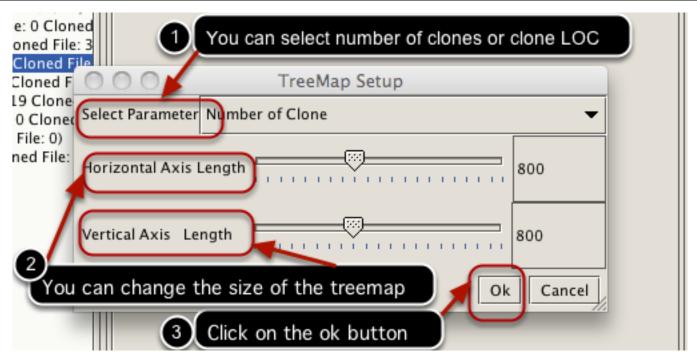
Treemap preserves the hierarchical structure of subject systems where each rectangle represents a file or directory. The rectangles representing the files are aggregated to indicate the cloning status of a directory in the system hierarchy.

### **Opening treemap**



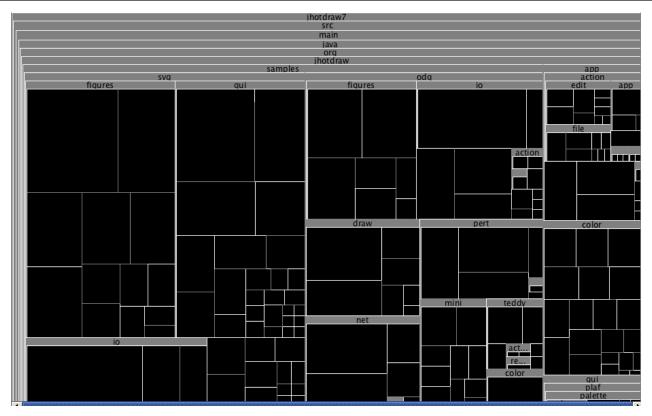
Select a directory and right clink on it to open the popup menu. Click on the Show TreeMap menu item.

## Configure the treemap



A dialog box appears that can be used to configure thre treemap. Click on the ok button to open the treemap.

## An Example



An example of a treemap for the JHotDraw system(Version 7.6). Putting the mouse over any rectangle will provide details information about that using a tooltip.

## Zoom in or zoom out



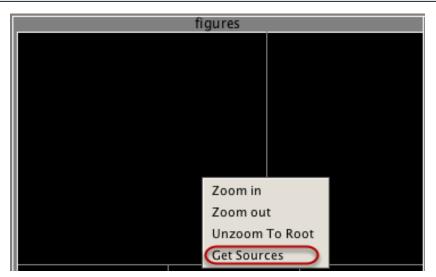
Right click on rectangle representing a directory opens a popup menu. You can then perform zoom in or zoom out operation.

## An example of zoom in operation



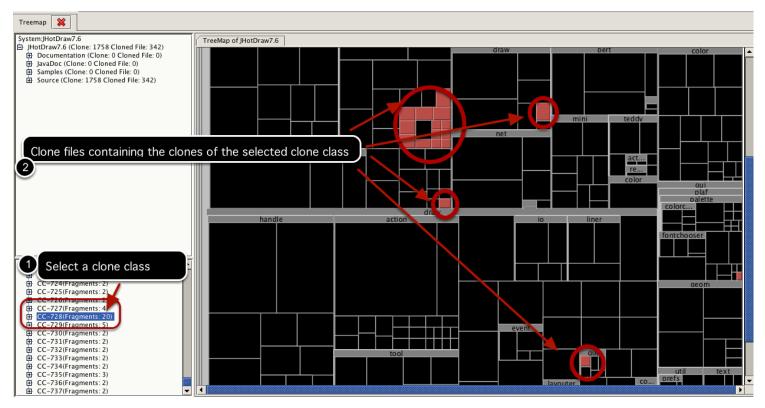
An example of zoom in operation on the *samples* directory. We can select the zoom out operation to move one step back or can select the *unzoom to root* option to move back to the beginning state.

## Obtaining source code of the clone fragment



Right click on a rectangle and select the *Get Source* menu item. This will open the clone code fragments of the file(s) in a new tab using the clone code browser. For details about the code browser, see Analyze Clone Fragments.

## Obtaining clone distribution information

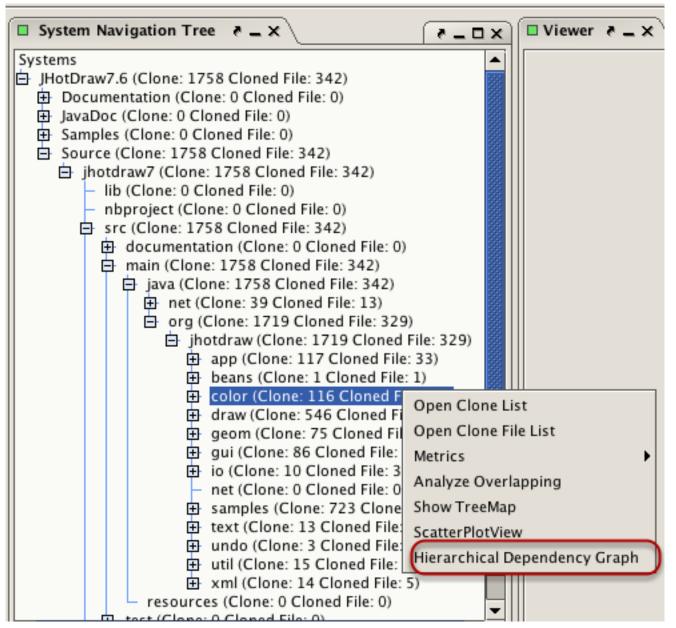


You can understand how the clone fragments of a clone class are distributed with the help of a treemap. Select a clone class from the bottom-left panel and the files containing those clones will be marked with red color in the treemap.

## **Hierarchical Dependency Graph**

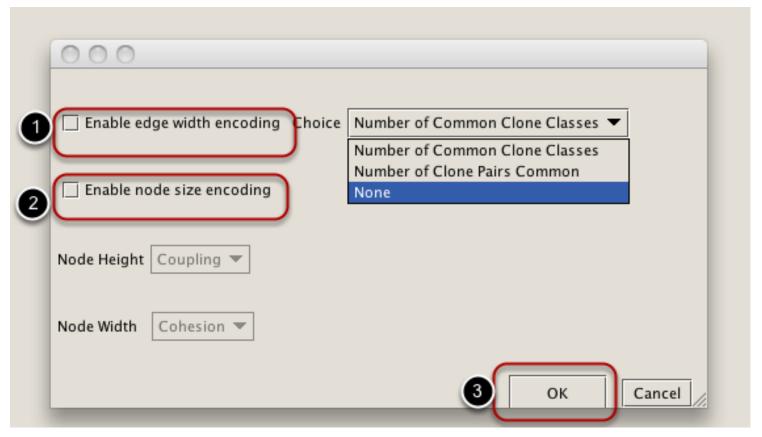
Clones are more problematic when members of a clone class scattered in different parts of a software system because this requires changes need to be made in different parts of the system. Thus, it is required to discover how clone fragments are distributed across subsystems/directories. Moreover, understanding cloning relationships among different subsystems can also reveal their dependencies. VisCad can render the hierarchical organization of a software system along with the distribution of clones using a hierarchical dependency graph.

## **Opening the graph**



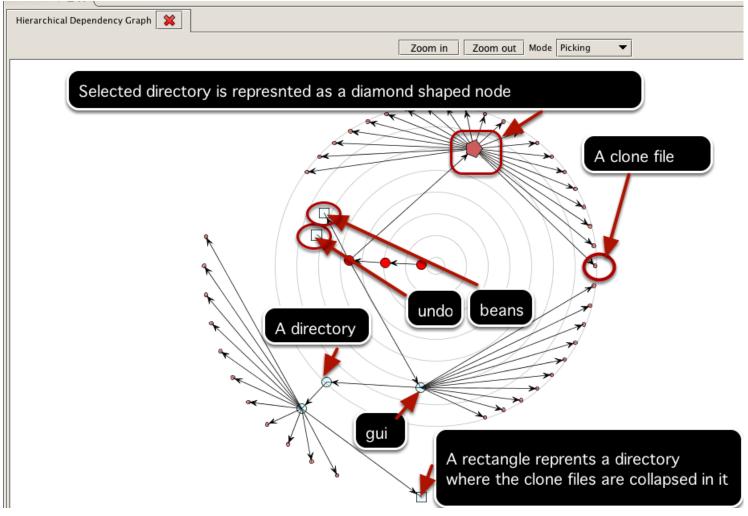
Select a directory from the system navigation tree and right click on it to open a popup menu. Click on the *Hierarchical Dependency Graph* menu item.

## Configure the graph



Edge width can represent number clone classes/clone pairs distribued between a pair of directories. Nodes width and height can represent cohesion and coupling value. Click on the ok button when you are done.

### An example

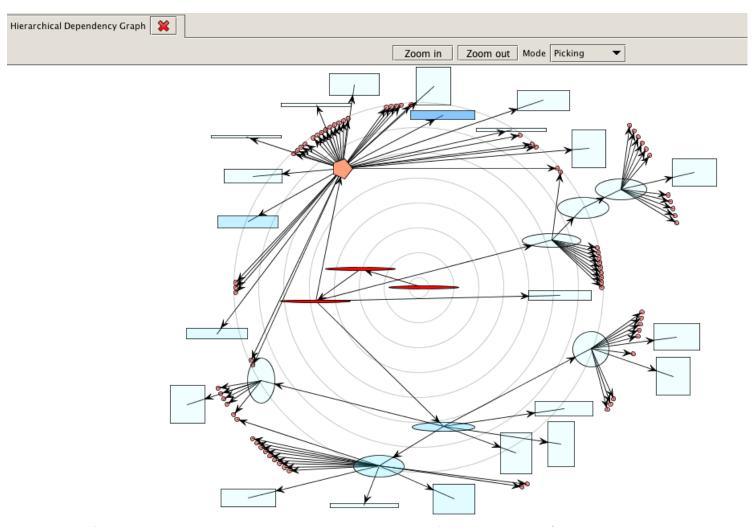


Suppose we want to understand which directories in the system contain the external clones of the color directory. We select the *color* directory from the system navigation tree, open the popup menu and select the *Hierarchical Dependency Graph*.

With the default settings, we get a graph as shown in the above figure. From the figure we can understand that the color directory has external cloning relationship with the *gui*, *undo* and *beans* directories.

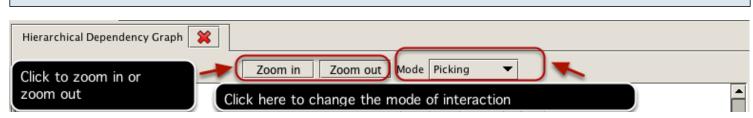
VisCad visualizes the graph by filtering those nodes that do not have any cloning relation with the selected directory.

## **Another example**



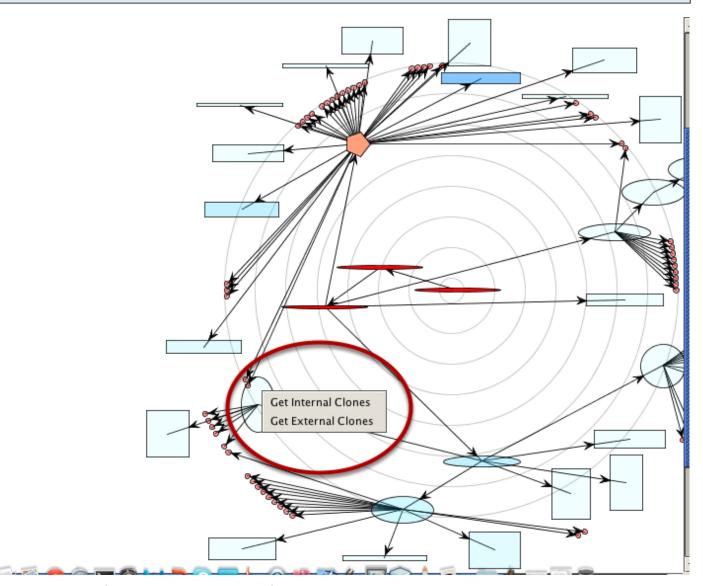
The above figure shows the hierarchical dependency graph for the *jhotdraw/draw* directory. In this case, node's (representing the directory) width and height represent the clone cohesion and coupling value. Node color value represent the external cloning relation level with the target directory.

#### Two mode of interaction



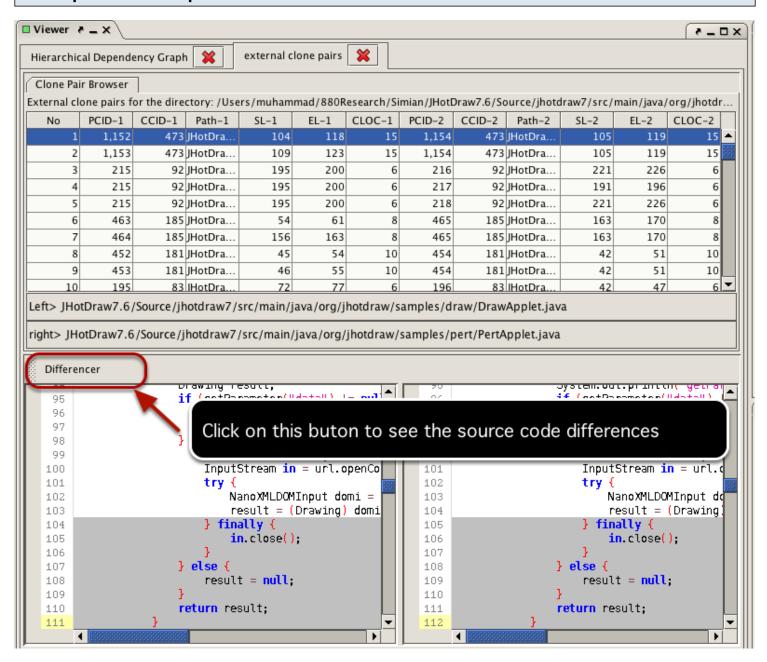
In the *picking* mode, you can select individual node and move them. In the *transforming* mode, you can move the entire graph.

## Obtaining external or internal clone pairs



Right click on a node( represnting a directory) opens a popup menu. You can then select the target operation. The clone pairs( internal/external) will be viewed in a new tab using the *clone pair browser* component (see the next figure).

## Example of a clone pair browser



## **Code Clone Metrics**

#### Introduction

For supporting in-depth clone analysis, VisCad can compute a set of metrics. We can divide the metrics into two broard categories. The first set of metrics (*clone system metric set*) relate clones with the organizational structure of the subject system and can be computed for different system boundaries, such as for the entire system, for subsystems/directories or for source files, as per the user's choice. Depending on the granularity of operation, we can again subdivide them into two groups, the *file metric set* and the *directory metric set*. The next set of metrics (*clone class metrics set*) deals with the clone classes.

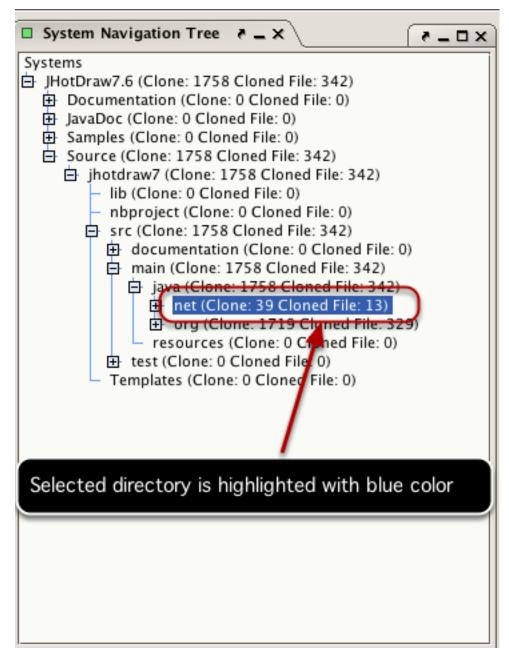
VisCad supports four operations for each metric set. These are:

- 1. Exporting: Results of metric computations can be exported in CSV (comma separated values) format.
- 2. Plotting : Although metrics are important for quantitative analysis, identifying important patterns from a large set of data is difficult. To avoid such difficulties, users can plot the metrics values with a bar chart which helps in identifying an anomaly within clone patterns easily.
- 3. Browsing Clone Code: Depending on the metrics values, user may be interested to explore the clone fragments located within a file or directory. VisCad also supports such operation.
- 4. Sorting: Values can be sorted to locate the maximum or the minimum value easily.

## **Obtaining Metrics For Files**

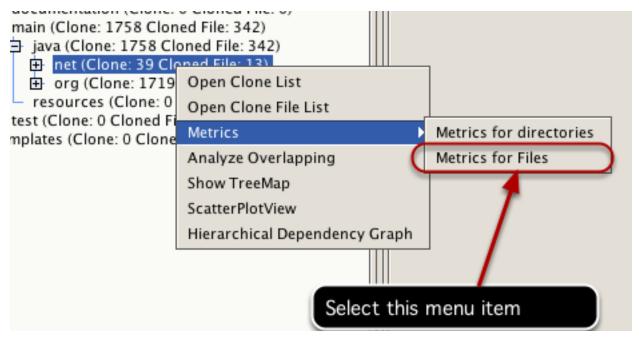
This section discusses the steps for obtaining various metrics values for the clone files located in a directory.

## Select a directory

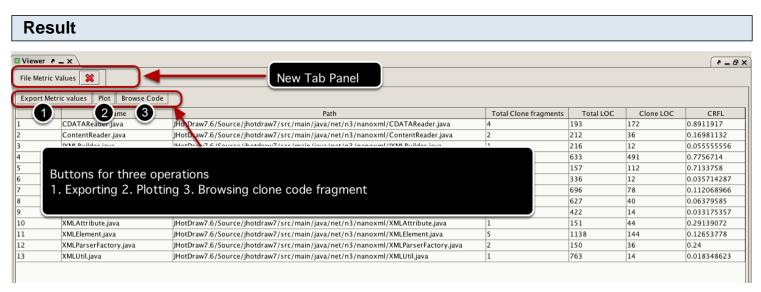


Click on a target directory to select it from the system navigation tree. Metrics will be computed for all clone files located within this directory

## Open the popup menu and select the operation

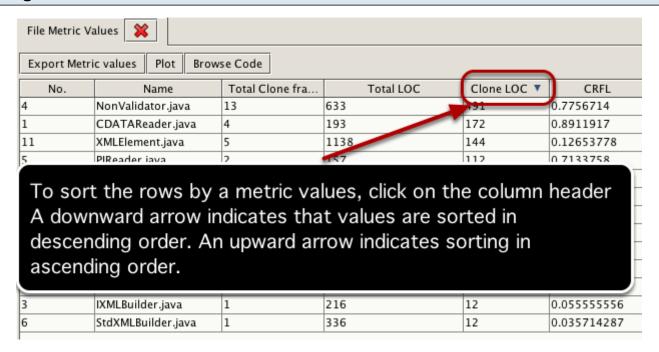


Right click on the selected directory to open the popup menu. Select Metrics->Metrics for files from the menu.

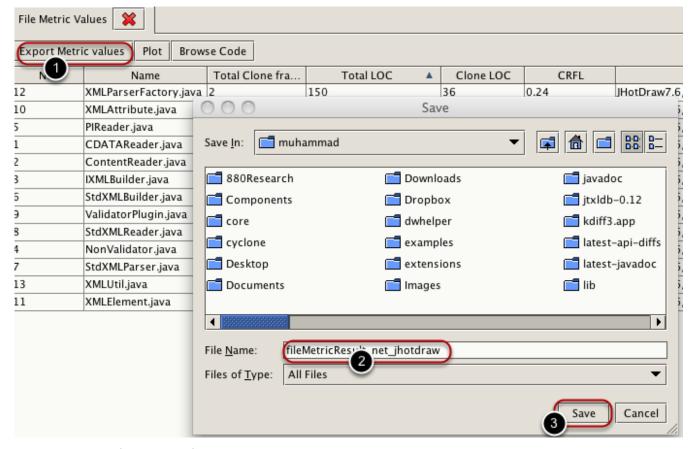


The computed metrics values for all clone files will be displayed as a new tab panel(titled *File metric Values*) in the viewer window.

## Sorting metric values



## **Exporting/Saving metric values**



To save the result, follow the following steps:

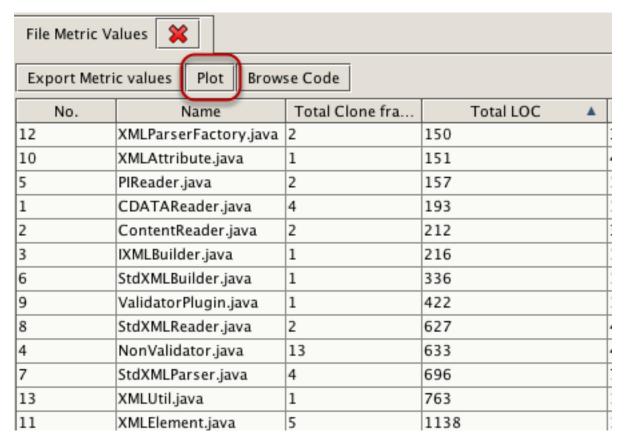
Step-1: To save the result, click on the Export Metric Values button. This will open a dialog box. Select

the folder where you want to save the result.

Step-2: Provide the name of the file you want to save .

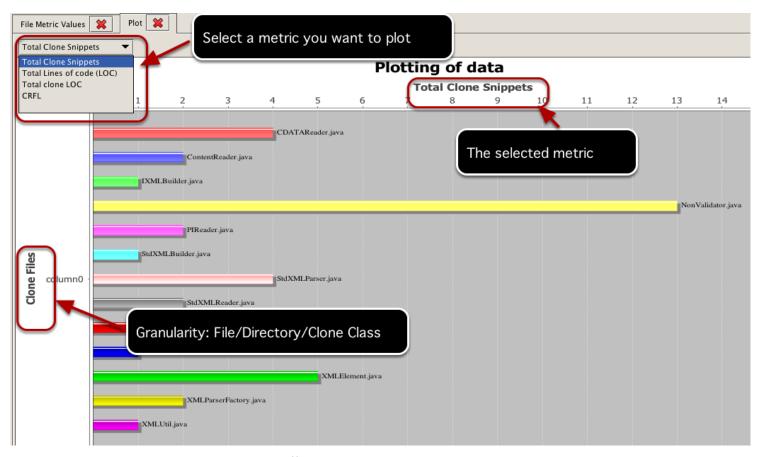
Step-3: Click on the Save button to complete the operation.

## **Plotting metric values**



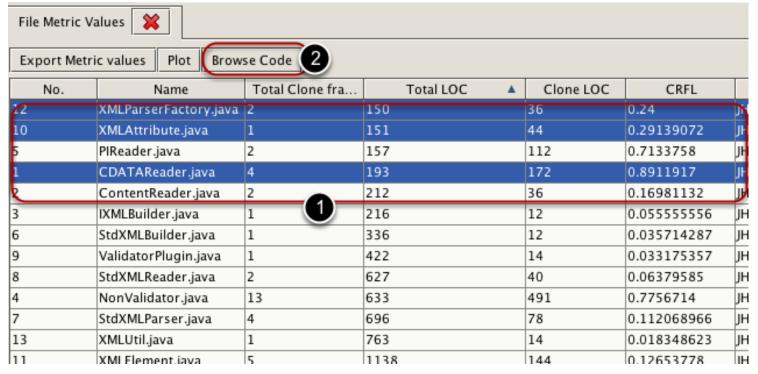
To plot the metric values, click the on the *Plot* button.

## **Plotting metric values (Continued)**



You can change the selection to plot different metric values. Putting the cursor over a bar provides the detail path information.

## **Browsing clone code fragments**

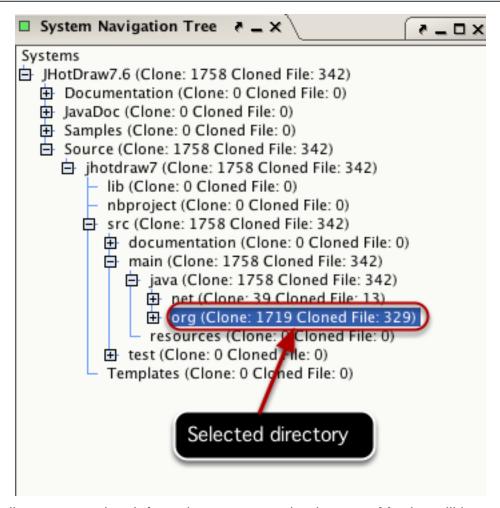


Step-1: Select one or more rows by holding the command key in MAC or control key in Windows. Step-2: Now click on the *Browse Code* Button to analyze the clone code fragments in the selected file(s) with the *source code browser*.

## **Obtaining Metrics For Directories**

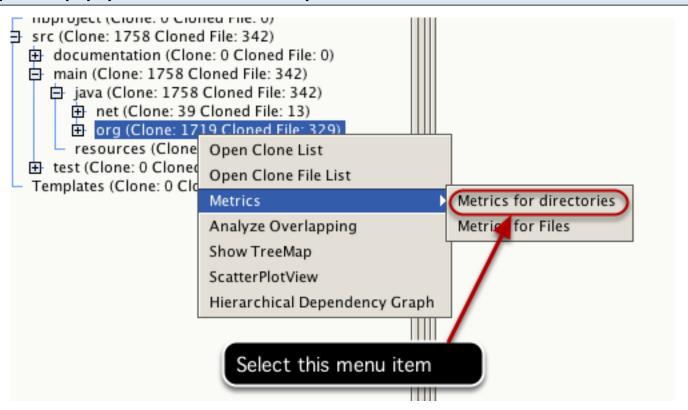
This section discusses the steps for obtaining various metrics values for all clone directories within a selected directory. A clone directory is a directory that contains at least one clone fragment.

## Select a directory



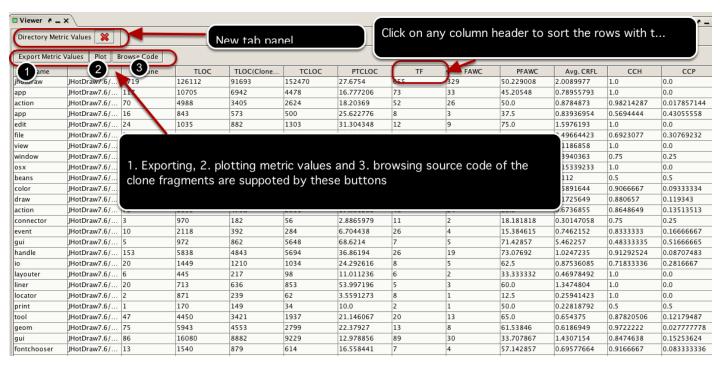
Click on a target directory to select it from the system navigation tree. Metrics will be computed for all clone directories within it.

## Open the popup menu and select the operation



Right click on the selected directory to open the popup menu. Select *Metrics->Metrics for directories* from the menu.

#### Result



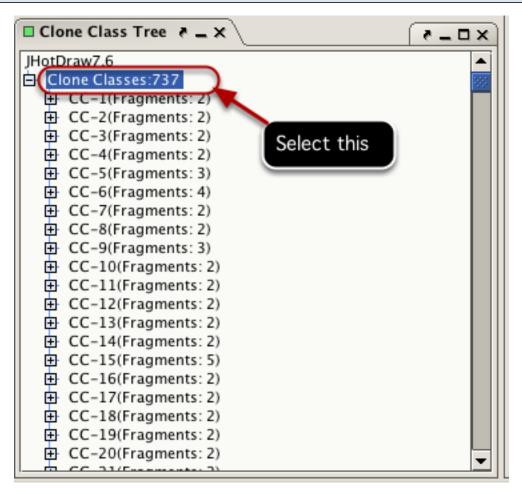
The computed metrics values for all clone directories will be displayed as a new tab panel(titled *Directory Metric Values*) in the viewer window.

Other operations
Various operations such as sorting, exporting and plotting computed metrics values, and browsing source code of the clone fragments are also supported. For details, see those operations in the discussion for files (Obtaining Metrics For Files).

## **Obtaining Metrics For Clone Classes**

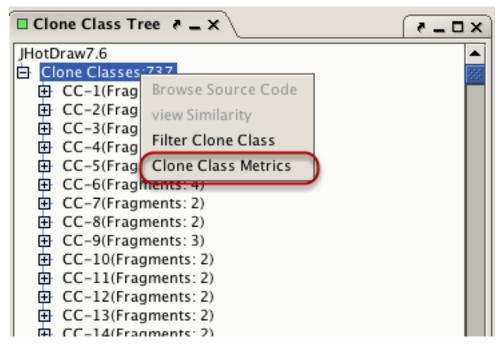
This section discusses the steps for obtaining various metrics values for the clone classes.

### Make the selection

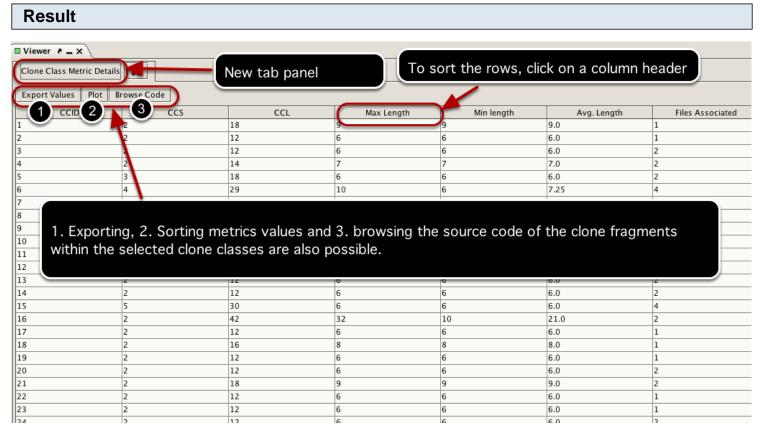


Click on the root node of the clone classes to select it from the clone class tree. Metrics will be computed for all clone classes within it.

## Open the popup menu and select the operation



Right click on the selected node to open the popup menu. Select *Clone Class metrics* menu item from the popup menu.



The computed metrics values for all clone classes will be displayed as a new tab panel(titled *Clone Class Metric Details*) in the viewer window.

Other operations
Various operations such as sorting, exporting and plotting computed metrics values, and browsing source code of the clone fragments are also supported. For details, see those operations in the discussion for files (Obtaining Metrics For Files).

# **Filtering**

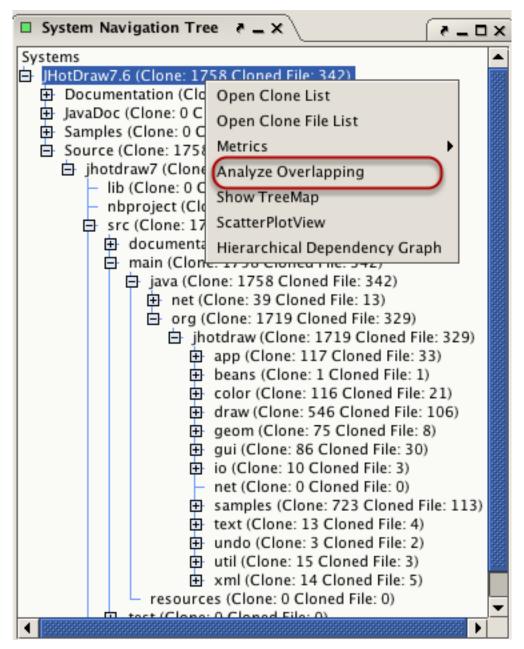
#### Introduction

The first and foremost challenge in clone analysis is the large volume of clone detection results. Not all clones are useful to the user and the objective of the analysis at hand governs the set of the useful clones. Here, the term 'useful clones' refers to those clone fragments that the maintenance engineers are looking for or are interested in. For example, when the objective is to analyze the inter-project clones, users may be more interested in the clone classes whose fragments are distributed across different projects and these clone fragments form the set of useful clones. In that case, we can filter out the clone pairs that are not distributed across different projects. VisCad supports a set of filtering operations to remove clones that are not useful/interesting to the users.

## **Overlapping Clone filterng**

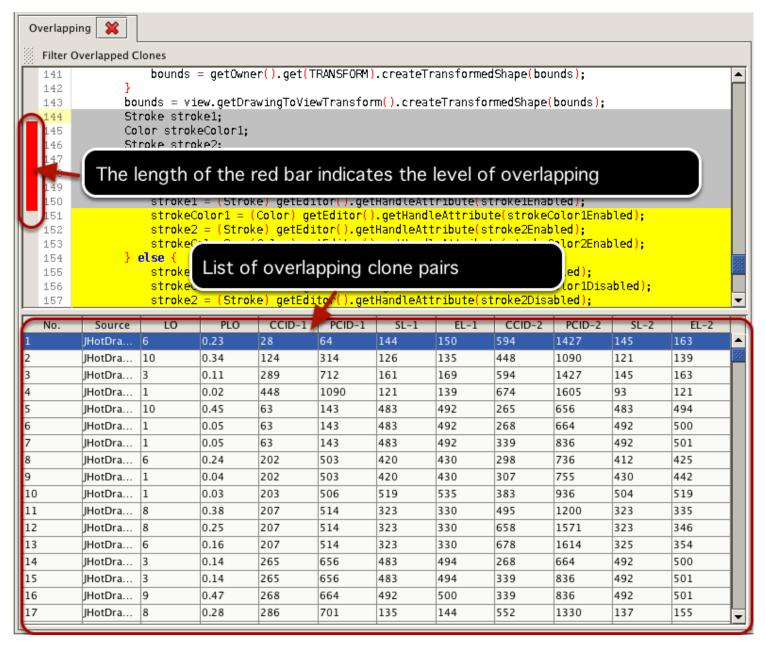
Clones may overlapped each other and removing those overlapping clones also reduces the size of the result set.

## Analyze overlapping



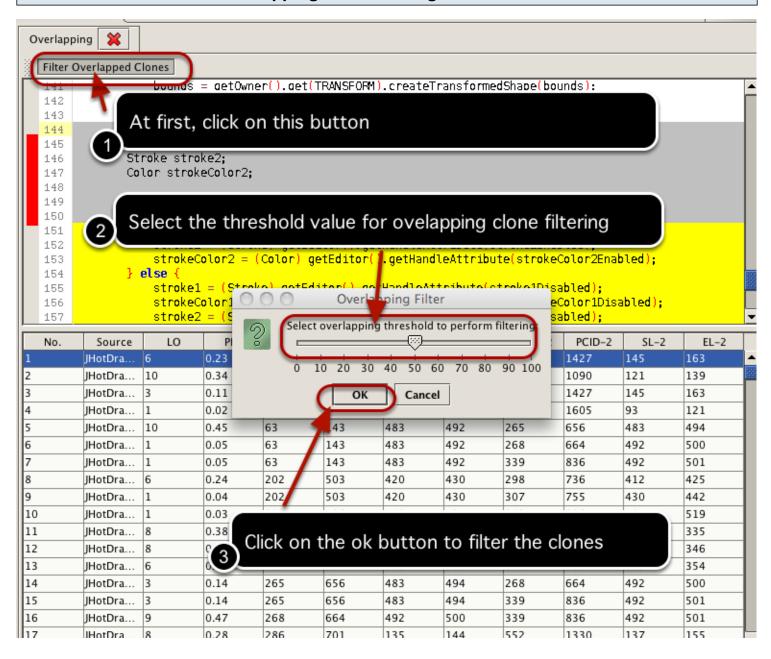
Click on a directory to selct it. Right click on it to open the popup menu and select *Analyze Overlapping* menu item.

## **Overlapping Clone pairs list**

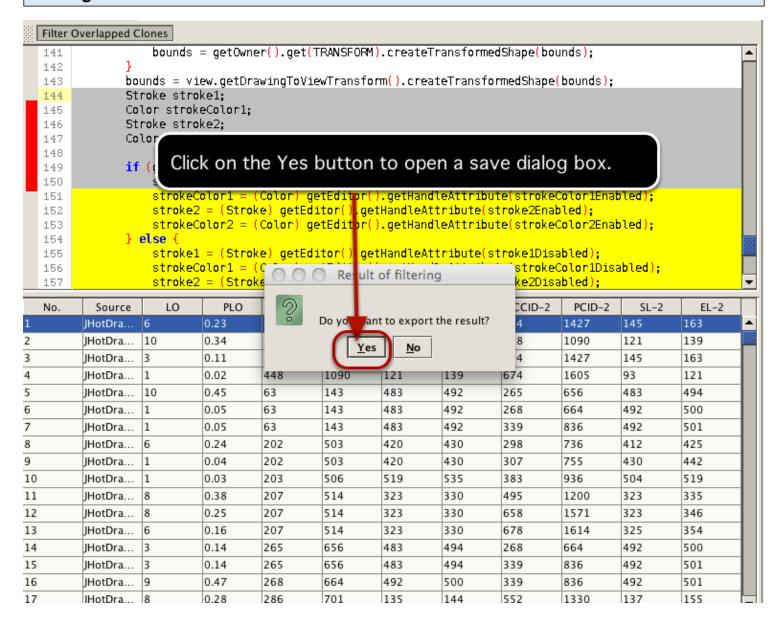


VisCad analyzes and shows the list of overlapping clone pairs.

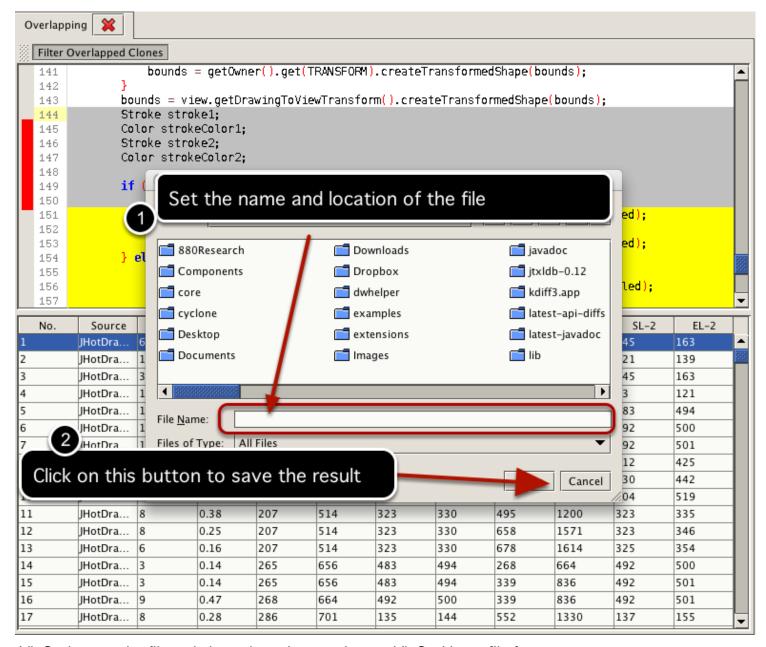
## Set threshold value for overlapping clone filtering



## Saving the result



## Saving the result(Continued)

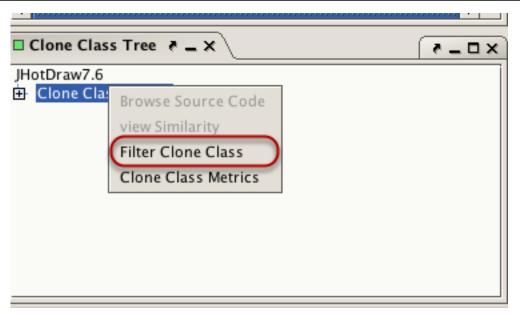


VisCad saves the filtered clone detection result as a VisCad input file format.

## **Textual filtering**

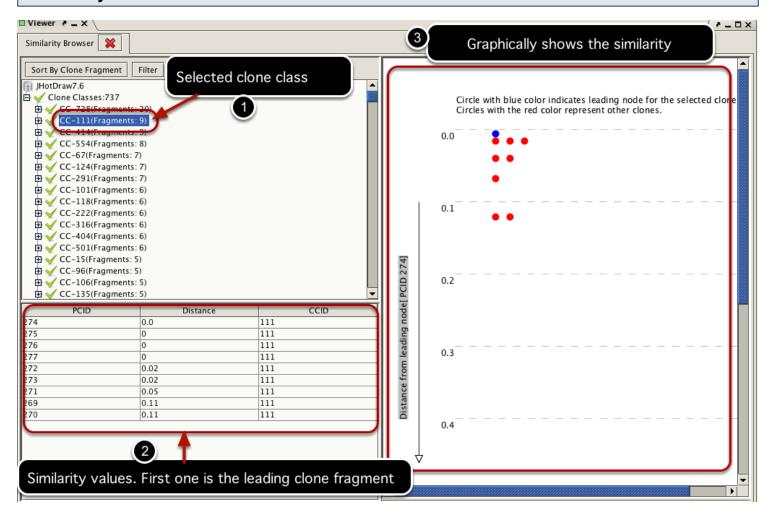
Textual filtering allows to remove clones that are only structurally similar without having any semantic similariy. For each clone class, VisCad determines the clone fragment that maximizes the sum of the textual similarity to all other fragments of that class. We call this fragment as the 'leading clone fragment' for that class. If the textual similarity between the 'leading clone fragment' and any other clone fragment in the clone class falls below a given threshold value, we remove the fragment from the analysis. We discard an entire clone class from the analysis when the textual similarities between the leading clone fragment and all other non-leading clone fragments of that clone class fall below the threshold value.

## Make the selection

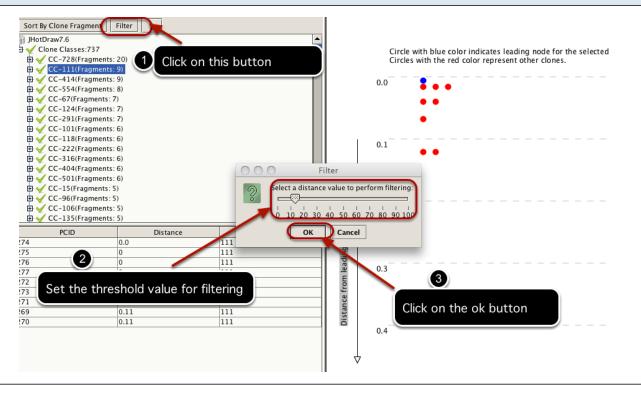


Select the *Clone Classes* node from the *Clone Class Tree*. Right click on it to bring the popup menu and select the *Filter Clone Class* option. This open the *similarity browser* view.

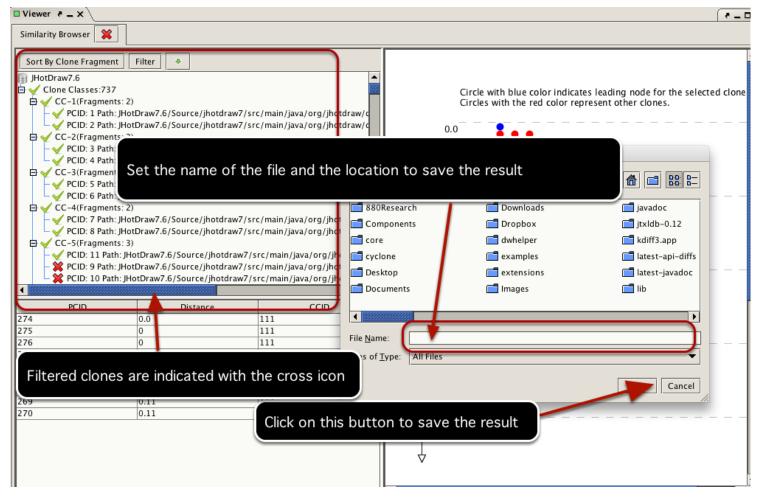
## **Similarity Browser View**



#### Filter the result



## Filter the result(Continued)



VisCad saves the filtered clone detection result in VisCad input file format.