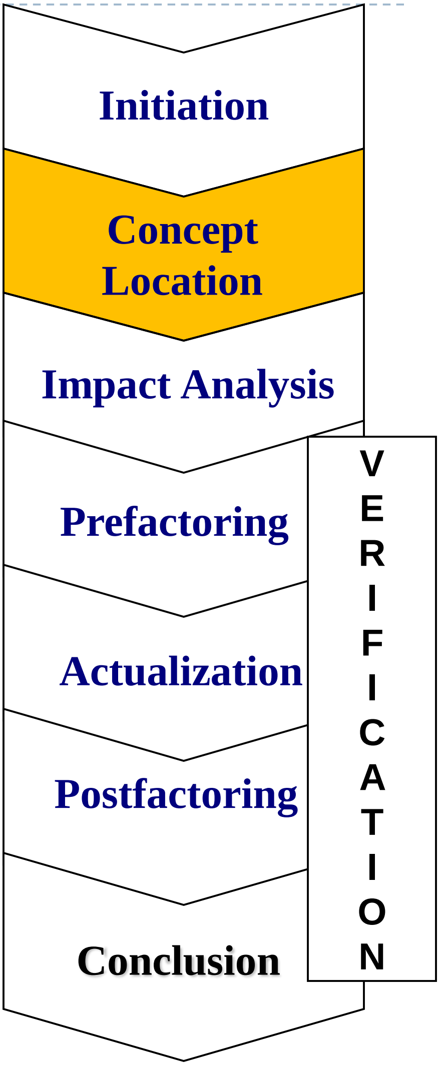
Concepts and Concept Location

Roberto A. Bittencourt
Based on Rajlich’s slides

Concepts and concept location



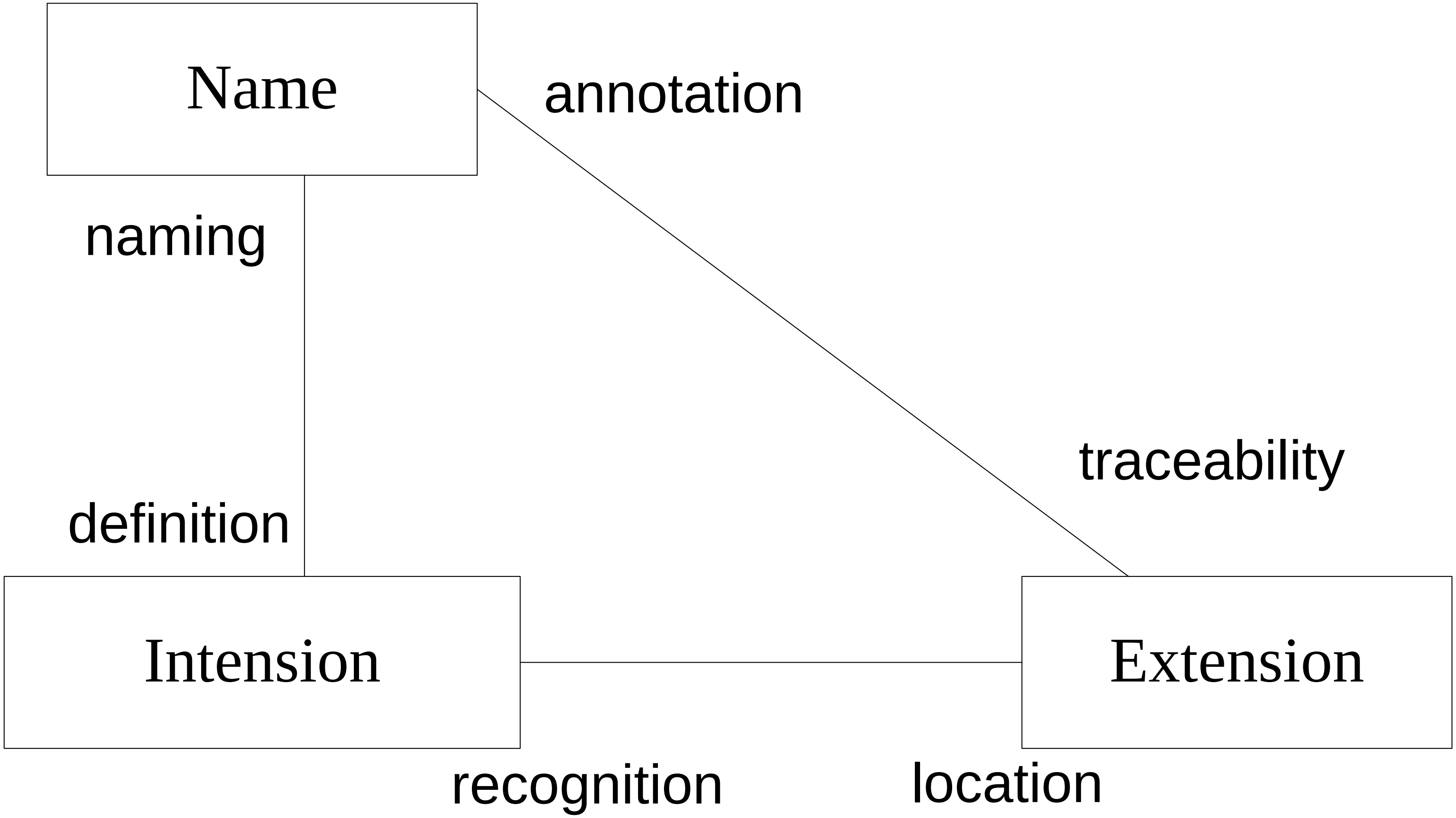
Role of concept location

- Concept location finds code snippet where a change is to be made
- Change requests are most often formulated in terms of domain concepts
  - Example: “Correct error that arises when trying to paste a text”
  - the programmer must find in the code the locations where concept “paste” is located
  - this is the start of the change

Partial comprehension of a code

- Large programs cannot be completely comprehended
  - programmers seek the minimum essential understanding for the particular software task
  - they use an as-needed strategy
  - they attempt to understand how certain specific concepts are reflected in the code
- Analogy: visiting a large city

Concept triangle



Name

annotation

naming

traceability

definition

Intension

Extension

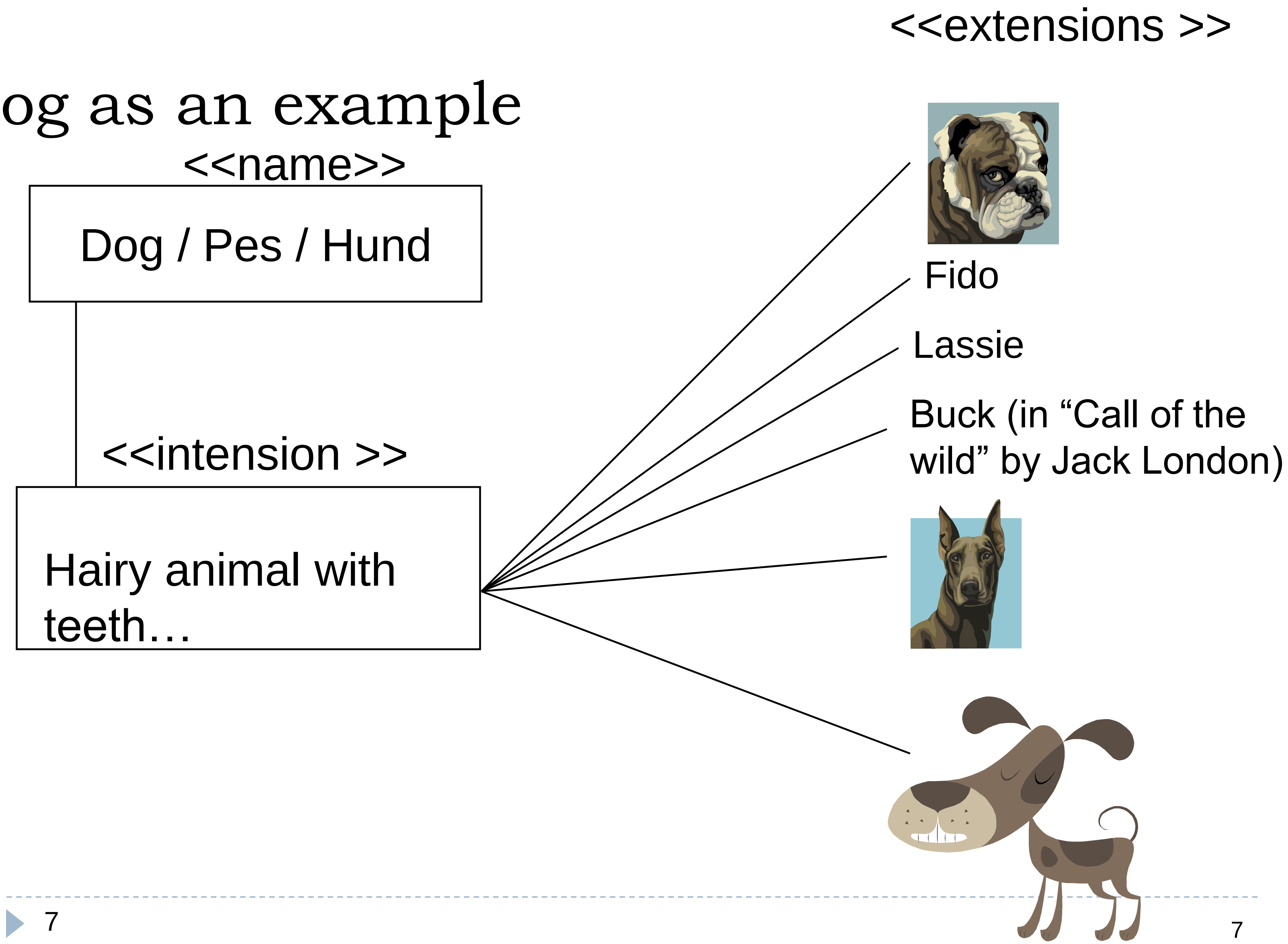
recognition

location

Spelling corner (Merriam-Webster)

- Intension \in-’ten(t)-shən\
  - synonym CONNOTATION
    - the suggesting of a meaning by a word apart from the thing it explicitly names or describes b: something suggested by a word or thing — W. R. Inge> an essential property or group of properties of a thing named by a term in logic
- Intention \in-’ten(t)-shən\
  - synonyms INTENT , PURPOSE , DESIGN , AIM , END , OBJECT , OBJECTIVE, GOAL mean what one intends to accomplish or attain.
    - INTENTION implies little more than what one has in mind to do or bring about <announced his intention to marry>. . .

Dog as an example



<<name>>

Dog / Pes / Hund

<<intension >>

Hairy animal with teeth…

<<extensions >>

Fido

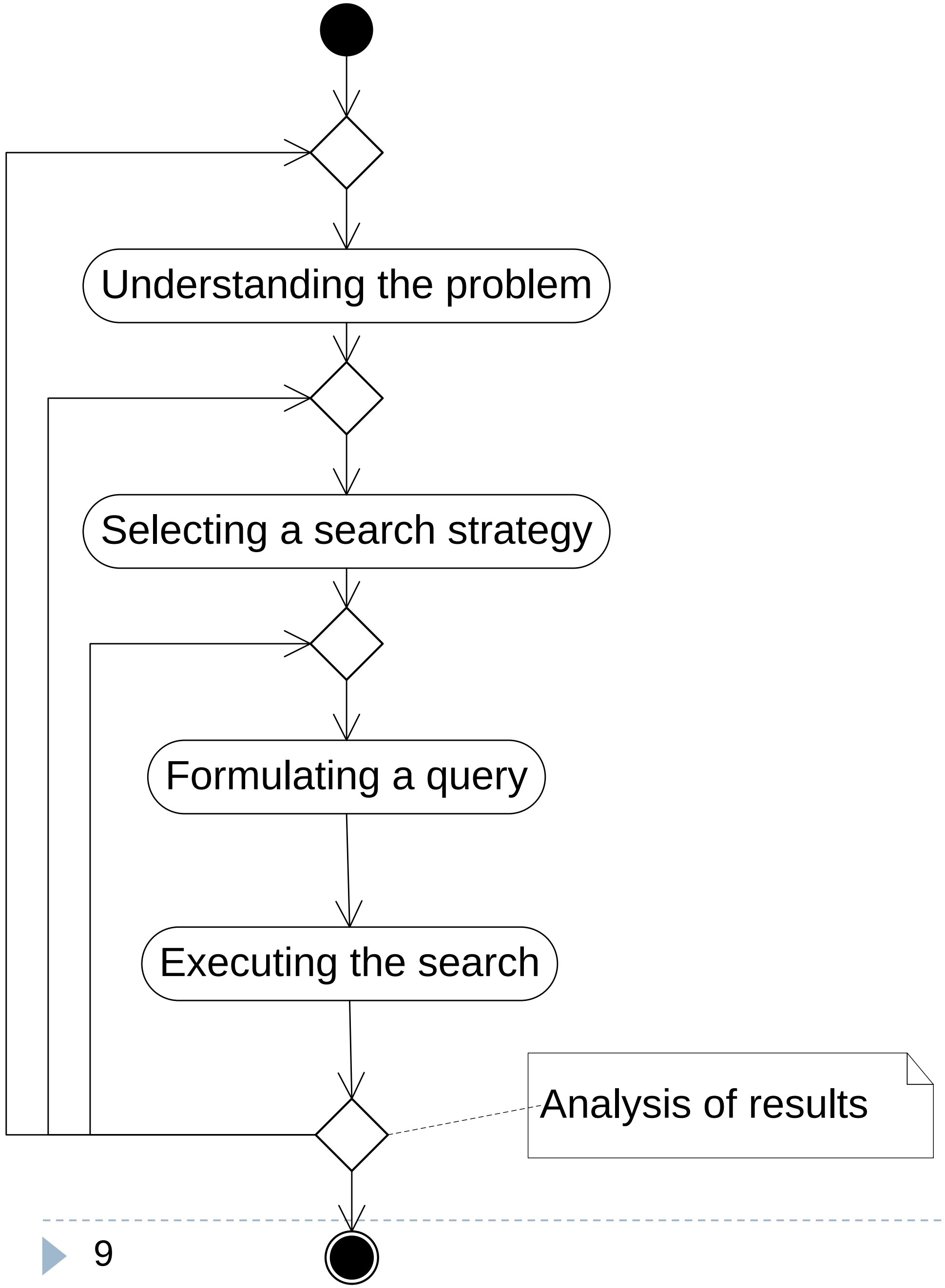
Lassie

Buck (in “Call of the wild” by Jack London)

Concept location

- Concept extensions are implemented as code fragments
  - variables, classes, methods, or other
- Programmers finds these code fragments
  - easy in small programs or in the programs that the programmer knows well
  - hard in large programs or programs that the programmer does not know
    - Watchmaker anecdote

Search in the unknown parts of system



Formulating a query

- Extract the set of concepts used in the change request
- Delete the concepts intended for the communication with the programmers
- Delete the concepts that are unlikely to be implemented in the code
  - concepts related to the things that are outside of the scope of the program
  - concepts that are to be implemented in the future.
- Rank the remaining concepts by the likelihood that they can be easily located

Example

- Point of Sale system
- Change request is "Implement a credit card payment"
- Identify the concepts
  - "Implement"
  - "Credit card"
  - "Payment"

Example

- Point of Sale system
- Change request is "Implement a credit card payment"
- Identify the concepts
  - "Implement" … communication with programmer
  - "Credit card"
  - "Payment"

Example

- Point of Sale system
- Change request is "Implement a credit card payment"
- Identify the concepts
  - "Implement" … communication with programmer
  - "Credit card" … to be implemented, not in the old code
  - "Payment"

Example

- Point of Sale system
- Change request is "Implement a credit card payment"
- Identify the concepts
  - "Implement" … communication with programmer
  - "Credit card" … to be implemented,
    not in the old code
  - **"Payment"** **!!!** Significant concept,
    find it in the code

Recognize concept

- Reading code
  - Comments and identifiers
  - Characteristic algorithm (plan)
- Small modification
  - Change the code slightly, execute
  - Throw away this modification!

Concept location methodologies

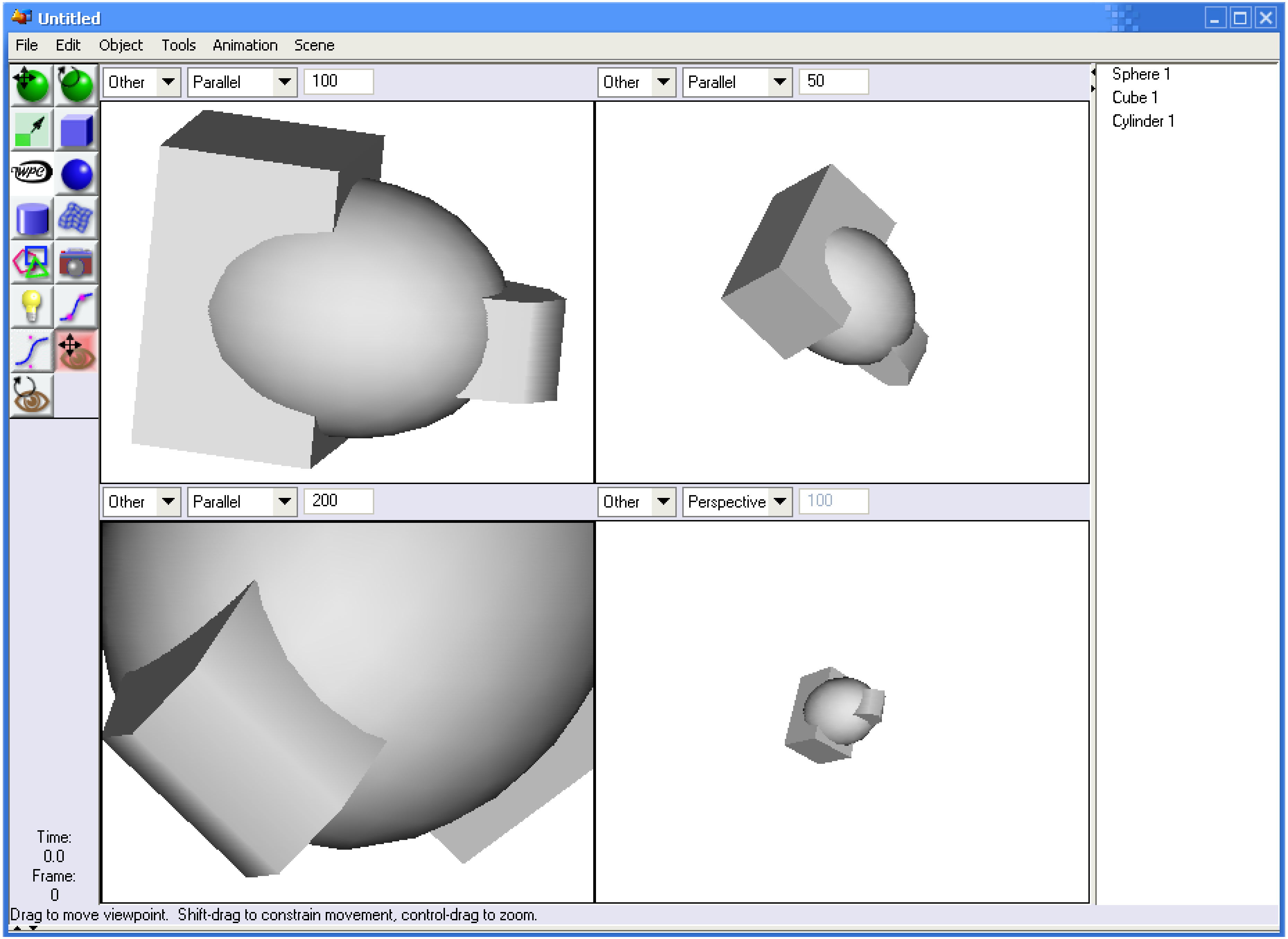
- Human knowledge
- Traceability tools
- Dynamic search (execution traces)
- Static search
  - dependency search
  - "grep" (pattern matching)
  - information retrieval techniques

GREP Search Technique

- GREP is an acronym for "global regular expression print".
  - GREP prints out the lines that contain a match for a regular expression.
  - Programmer iteratively formulates search query and then investigates the results.
  - If the results are too big to review, programmer either performs further search within these results or reformulates the search query.

Example: Art of Illusion

- 3D modeling studio, written in Java
- More than 600 classes, 100,838 LOC.
- Implement a zooming control
  - currently, the only way to zoom is to enter the zooming value into the specific text box
    - a value of the zoom has to be typed in by the user
    - the default value is 100%.
  - implement zooming control that uses arrow keys



Other Parallel 100

Other Parallel 50

Other Parallel 200

Other Perspective 100

Sphere 1
Cube 1
Cylinder 1

Time:
0.0
Frame:
0

Drag to move viewpoint. Shift-drag to constrain movement, control-drag to zoom.

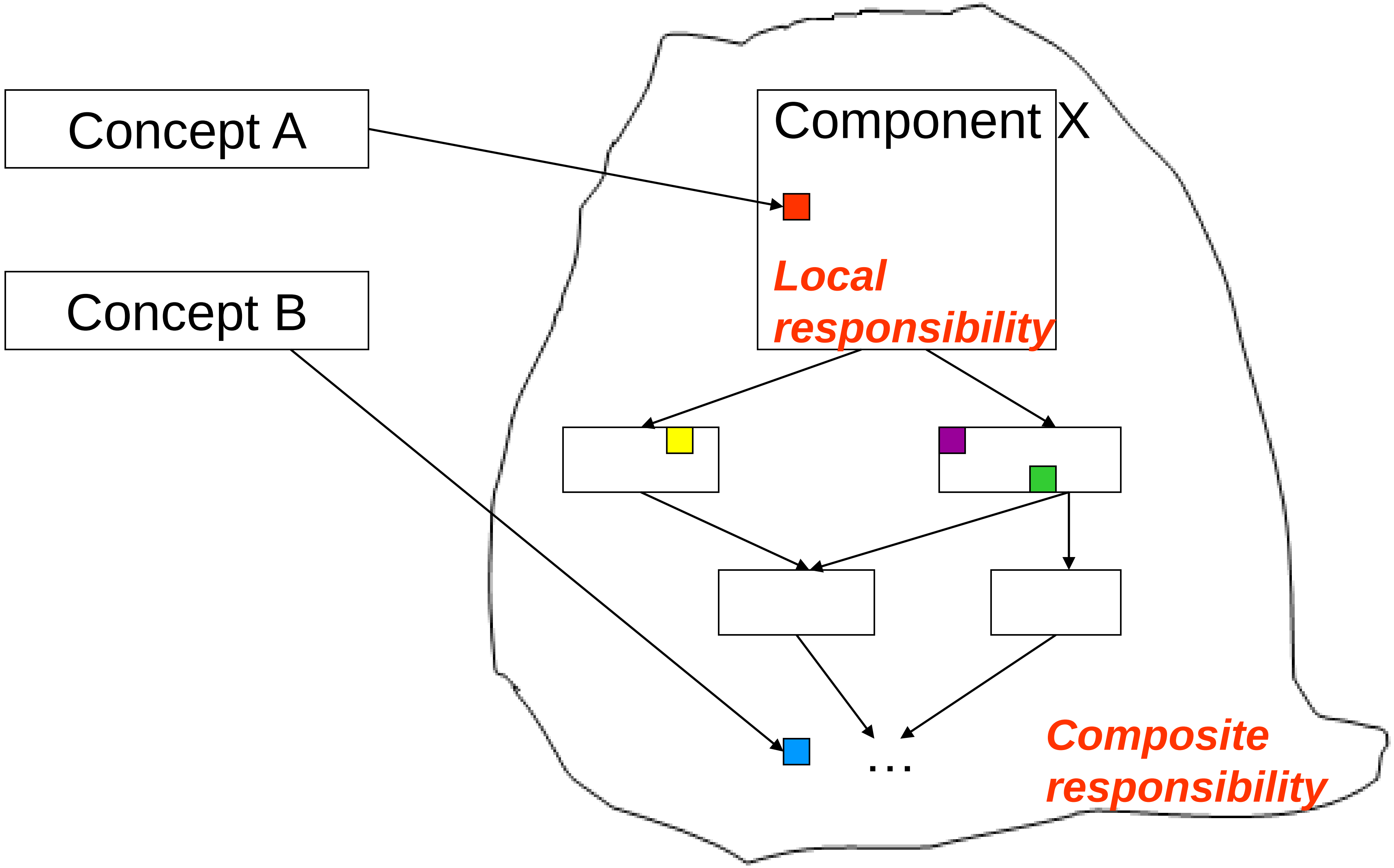
GREP example

- First search:“*zoom*”
  - The query produced irrelevant 6 lines
- Second search: “*scale*”
  - returned in 1,544 lines, too large for inspection.
- Third search:“*100*”
  - default scaling value is 100
  - search within the results of the previous search
  - returned 4 lines from the ViewerCanvas.java file.
- Inspection
  - one of the lines is the location

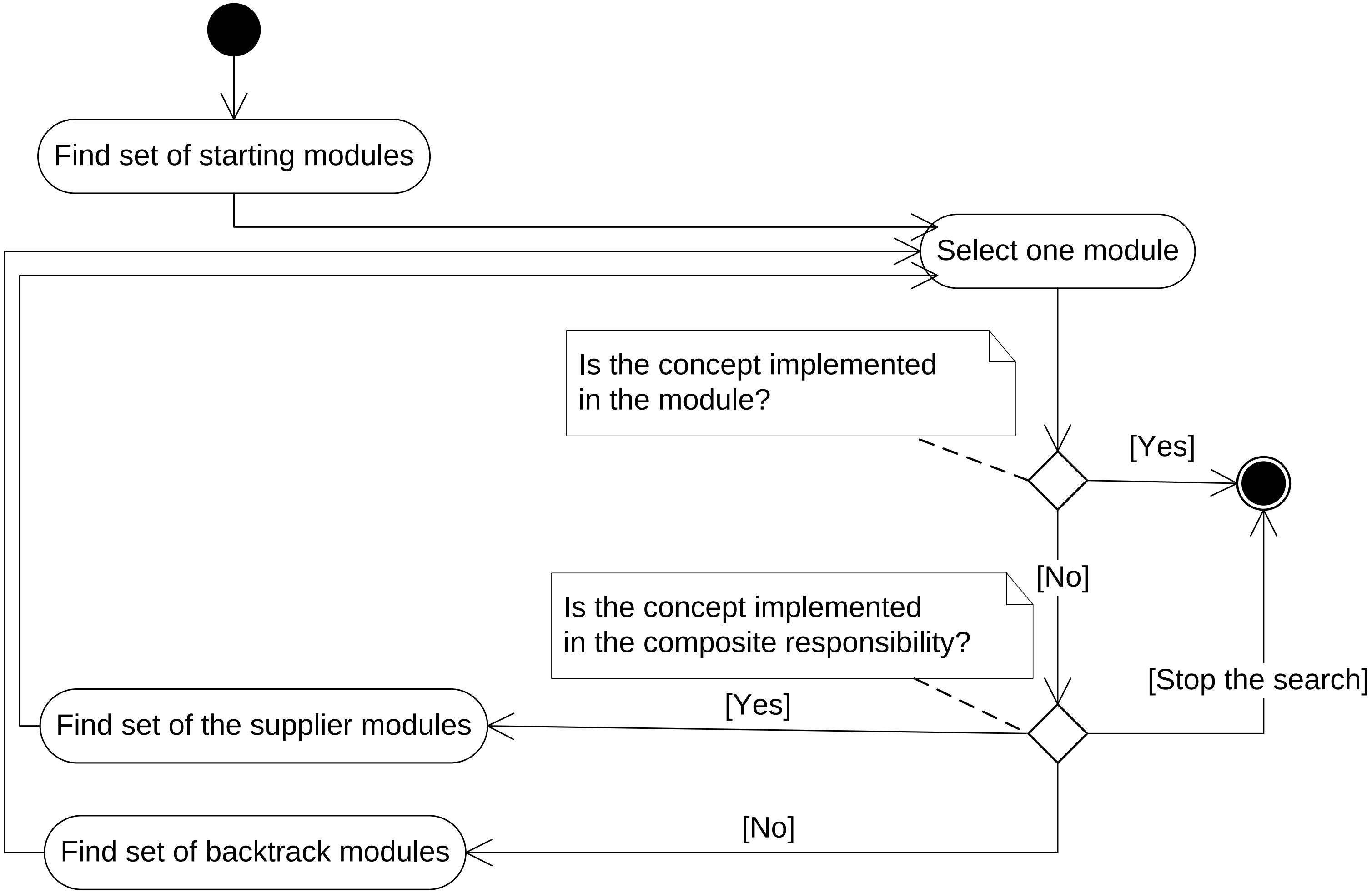
Dependency Search Technique

- Uses Class Dependency Graphs (CDG)
  - extracted form the existing code
- Local functionality
  - consists of concepts that are actually implemented in the module and are not delegated to others.
- Composite functionality
  - as the complete functionality of a module combined with all its supporting modules.
- Determined by reading code and documentation

Functionalities of component X



Concept location by dependency search



Find set of starting modules

Select one module

Is the concept implemented in the module?

[Yes]

[No]

Is the concept implemented in the composite responsibility?

[Stop the search]

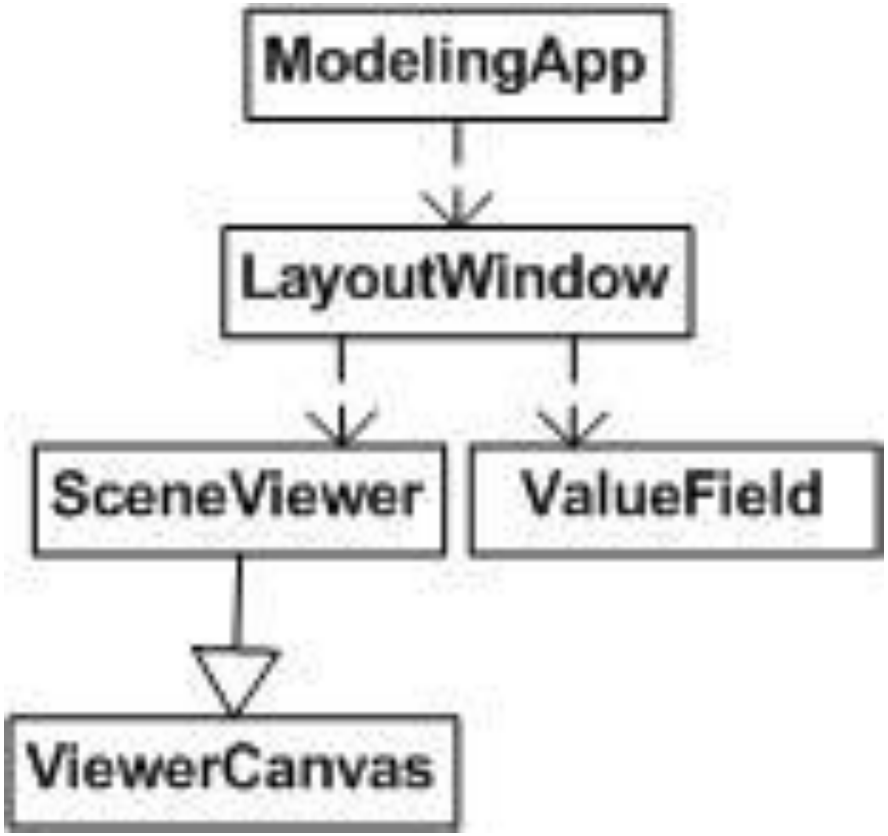
[Yes]

Find set of the supplier modules

[No]

Find set of backtrack modules

Progress of the search



Dependency search part 1

- Start at the ModelingApp class
  - concept not contained within its local responsibility
- The next step: inspect LayoutWindow
  - responsible for constructing the main AOI window
  - composite responsibility contains the concept, but the local responsibility does not.
- There were clues to search ValueField
  - It implements the text box.
  - concept is not present in the composite responsibility
  - backtrack to the LayoutWindow class

Dependency search part 2

- The SceneViewer class
  - several functions are responsible for responding to events from the user
  - function updateImage() was responsible for repainting the screen
  - we determined that the composite responsibility of this function contained the concept.
  - local responsibility of SceneViewer still did not contain the concept
- ViewerCanvas class
  - Contains the concept

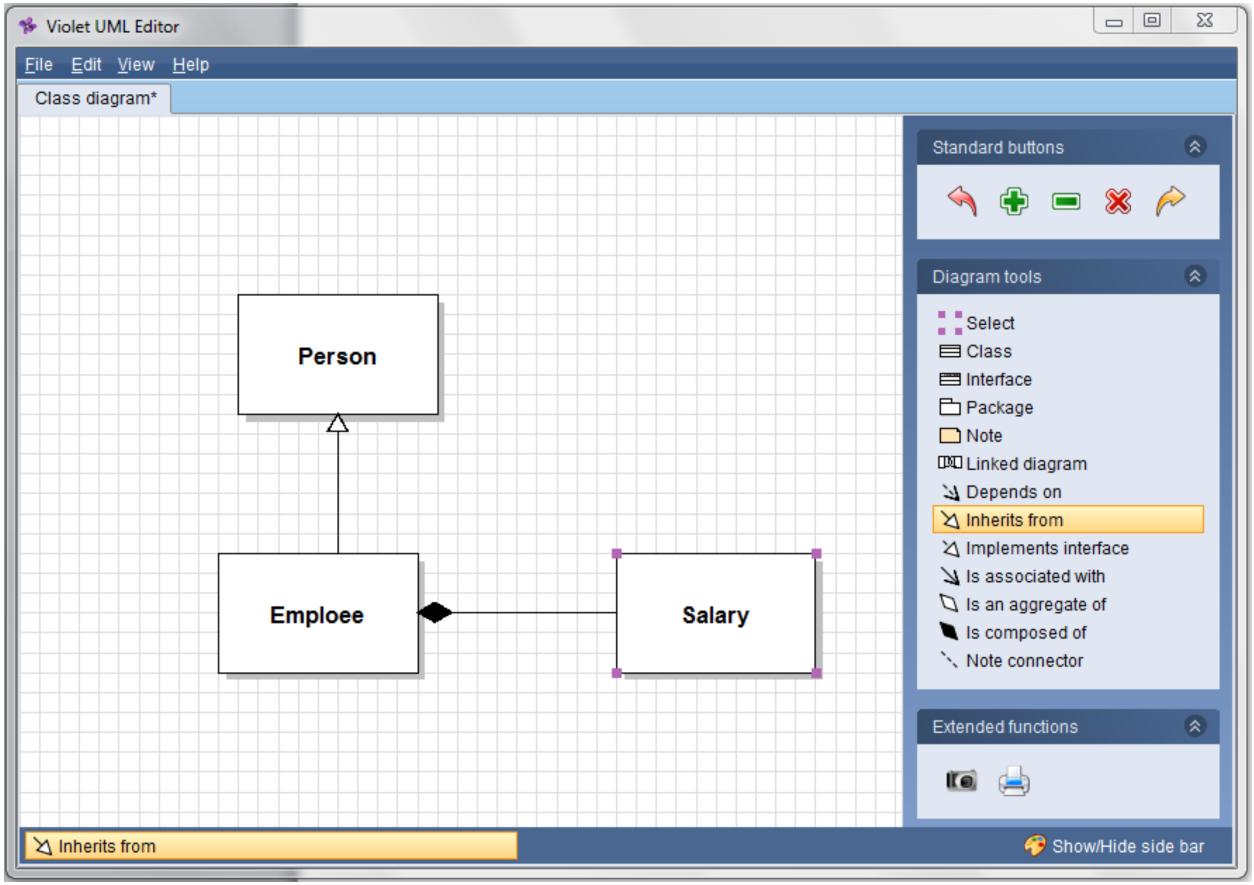
Comparison of the Techniques

- The grep-based
  - depend on the use of naming conventions
  - independent of class structure
    - GREP tools provide just the list of search results
  - suitable for explicit concepts only
- The static dependency search technique
  - utilizes the class structure
  - needs correct understanding of composite and local functionality
  - Suitable for both explicit and implicit concepts

Example Violet

- Violet
  - Open source UML editor
- Supports drawing UML Diagrams
  - Class diagram, Sequence diagram, State diagram, Object diagram, Use case diagram
- 60 classes and 10,000 lines of code
  - http://sourceforge.net/projects/violet/

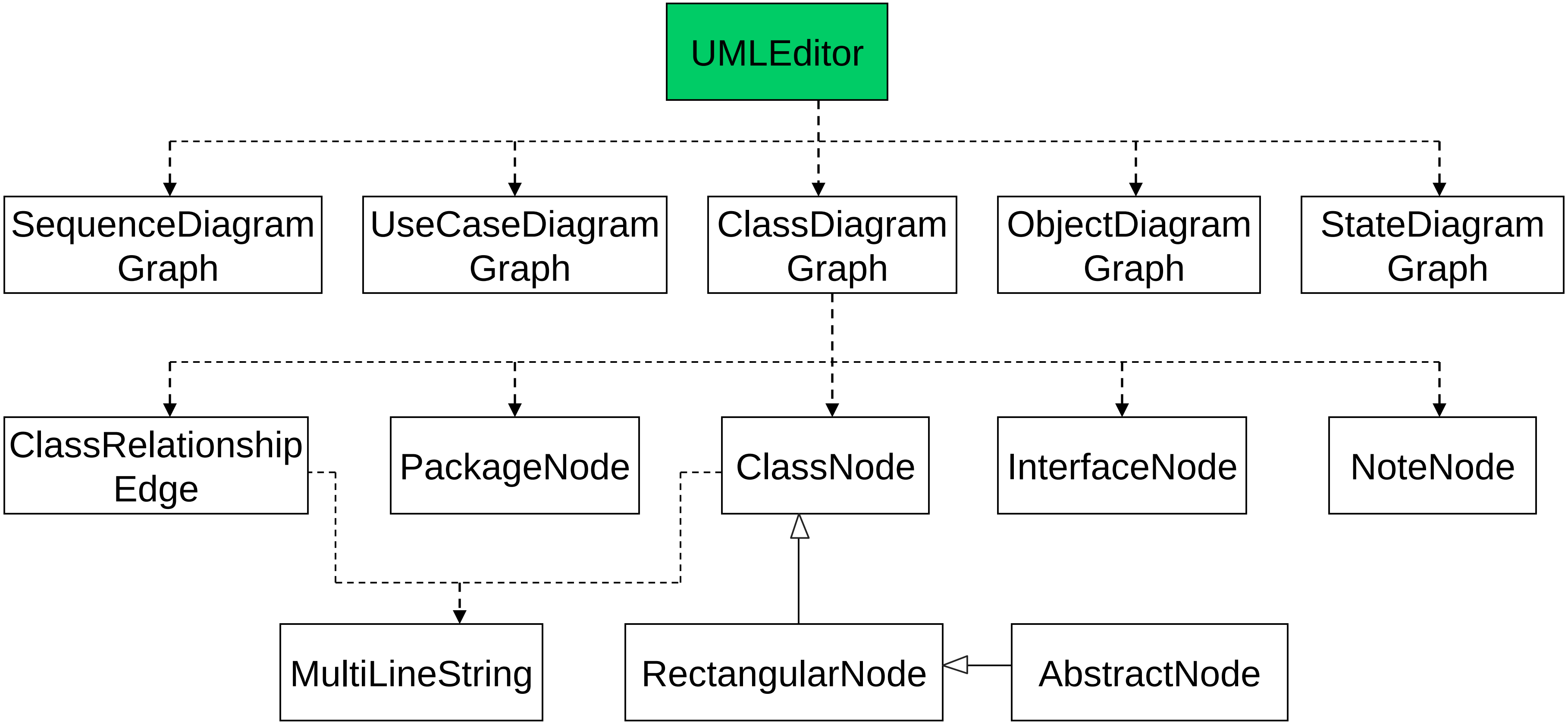
GUI of Violet



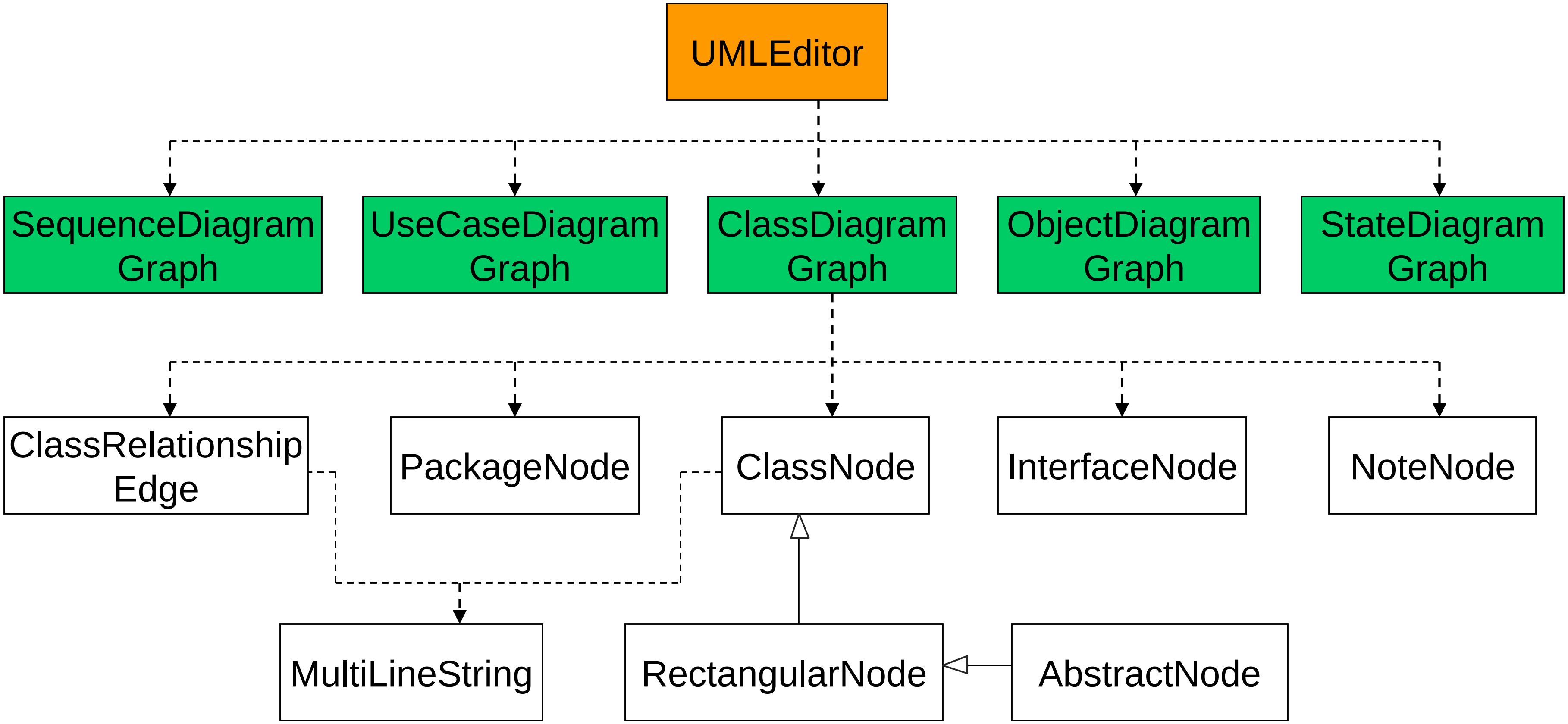
Change Request

- Record the author for each figure
- This change will make Violet more versatile
  - Support for cooperative work
  - The author created a figure
    - Author knows the semantics of the figure
- Name of concept: "author"
  - Implicit concept extension
    - The extension is not present in the current code
  - Belongs to the set of the figure properties

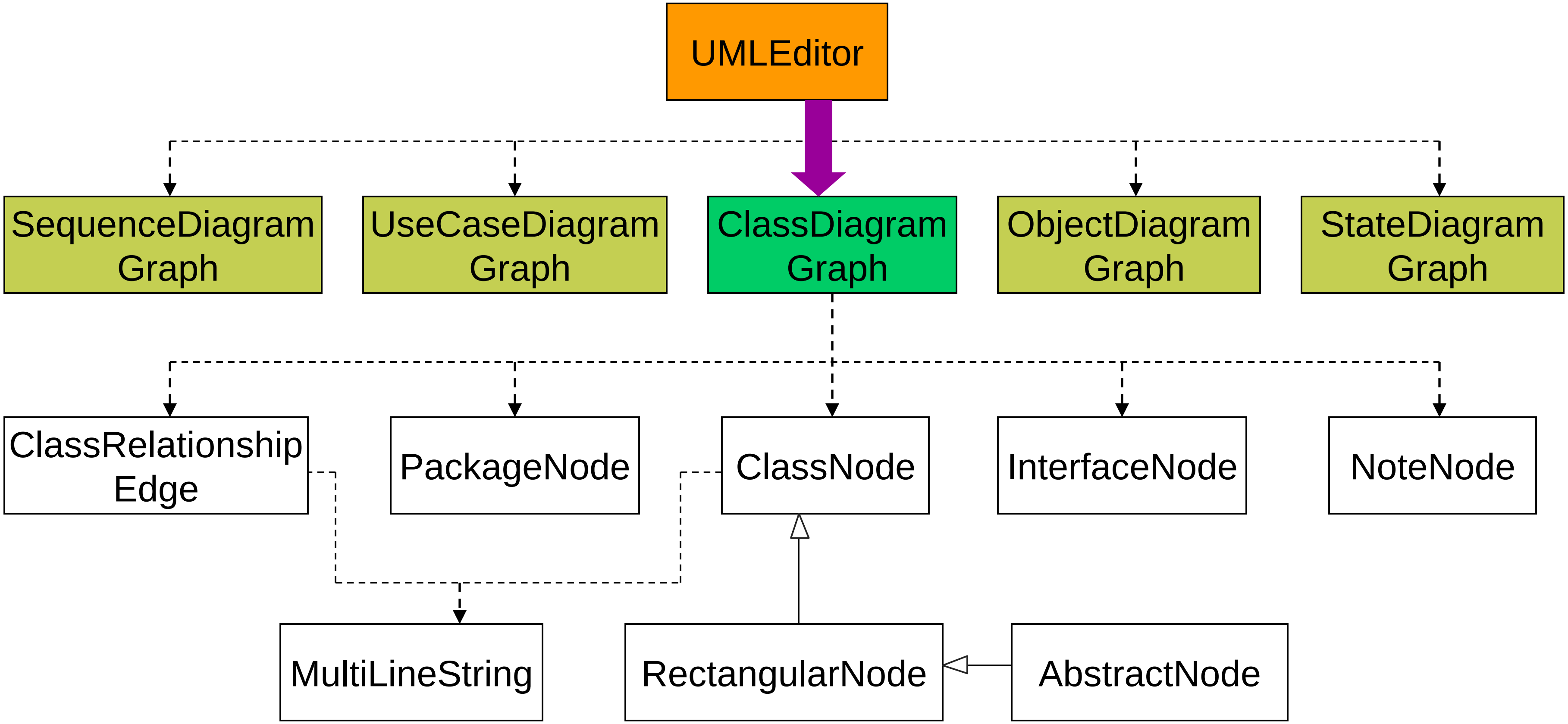
Locating figure properties: Start



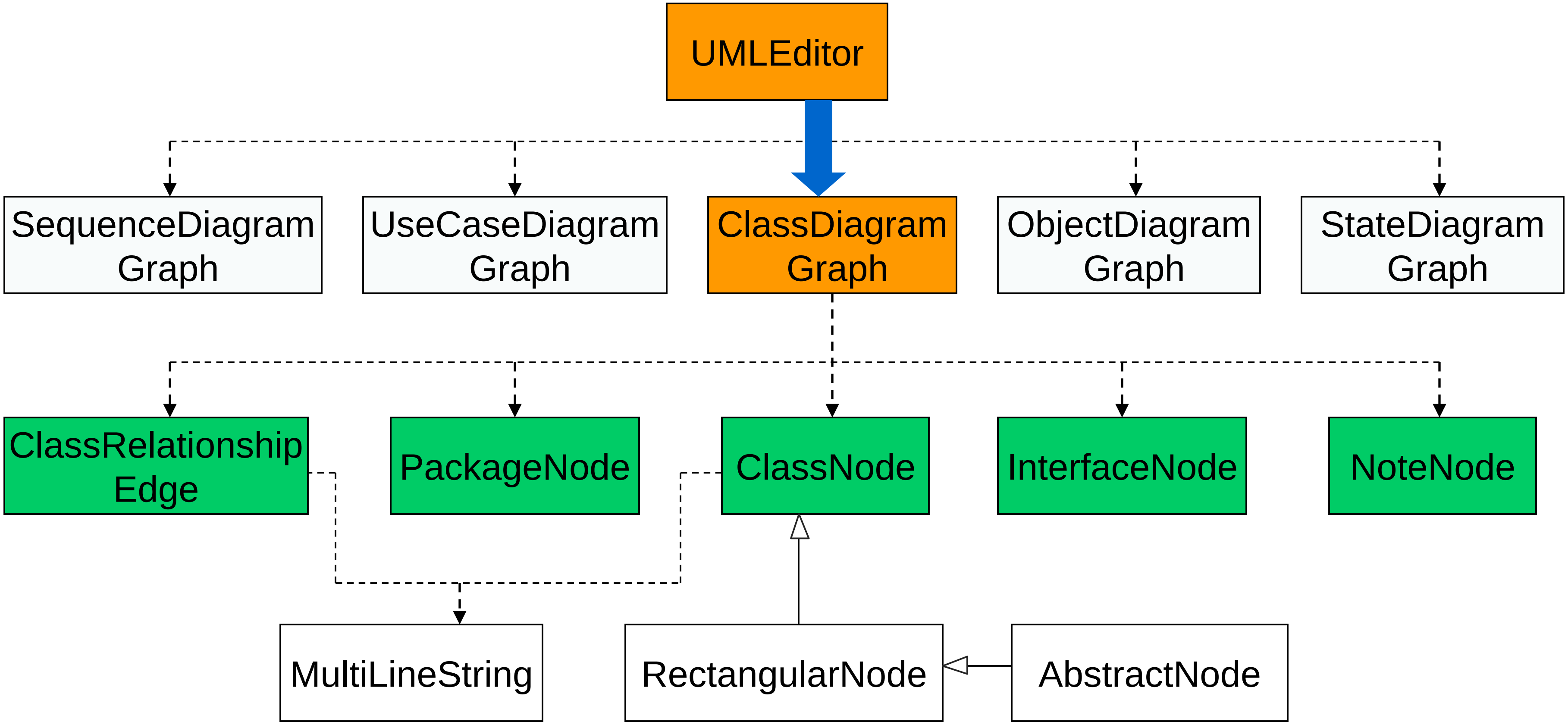
Classes to inspect



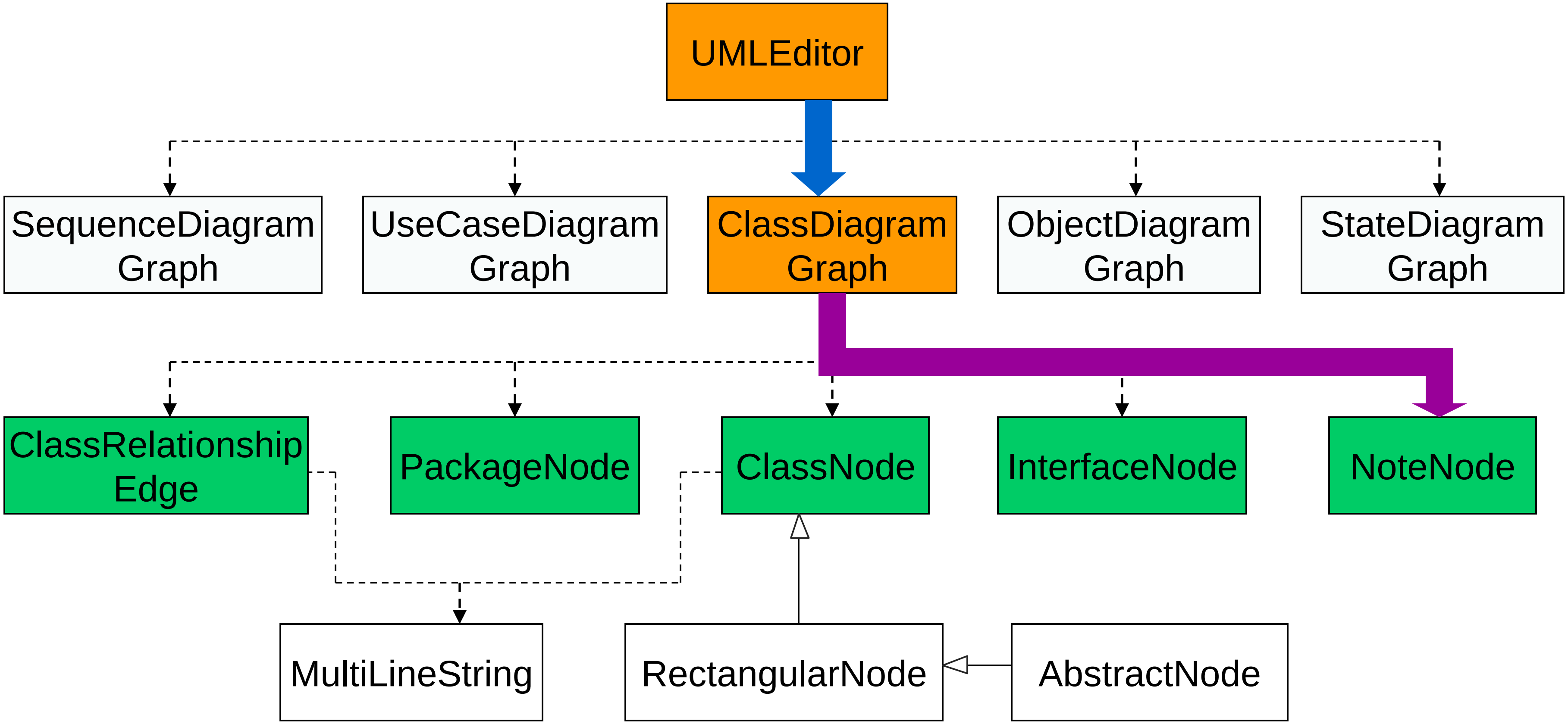
Most likely supplier



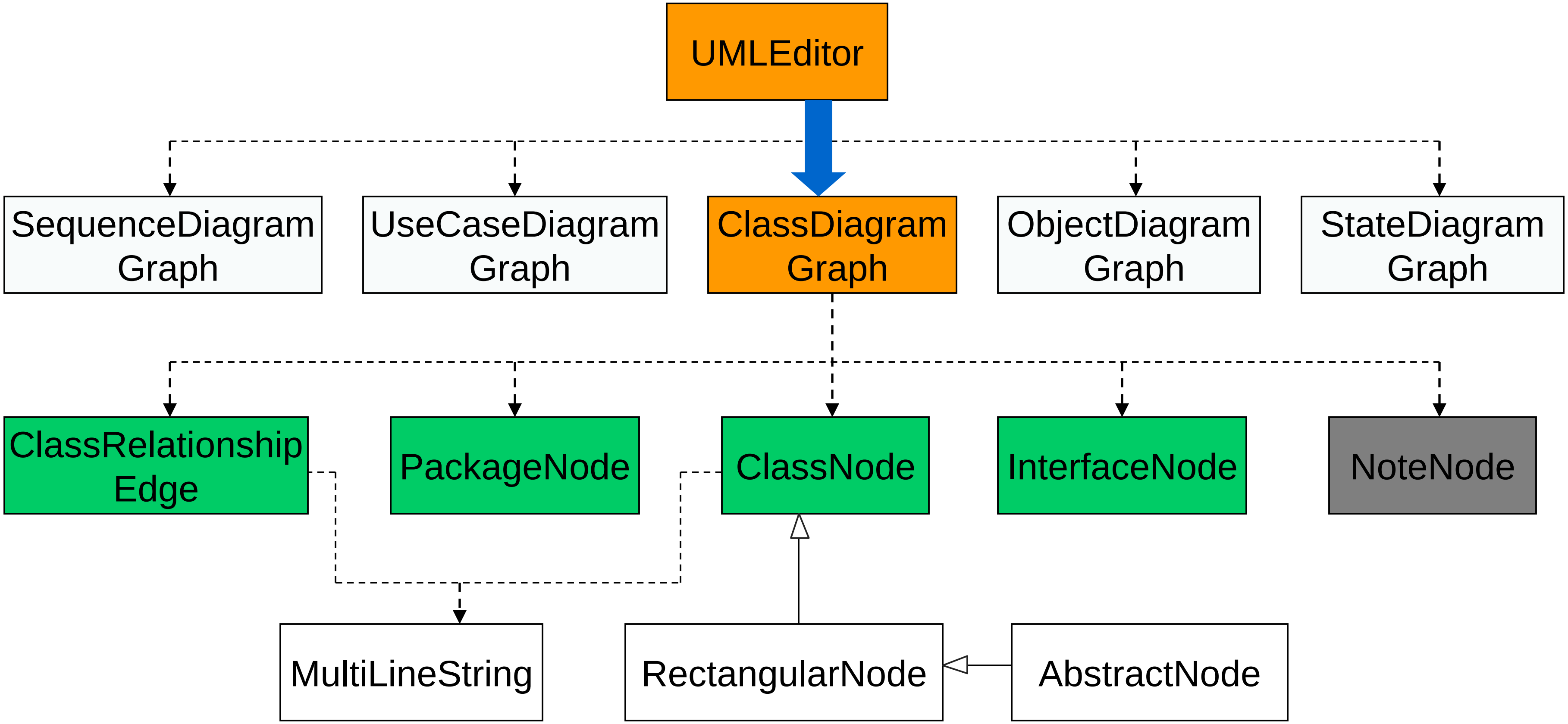
Next classes to inspect



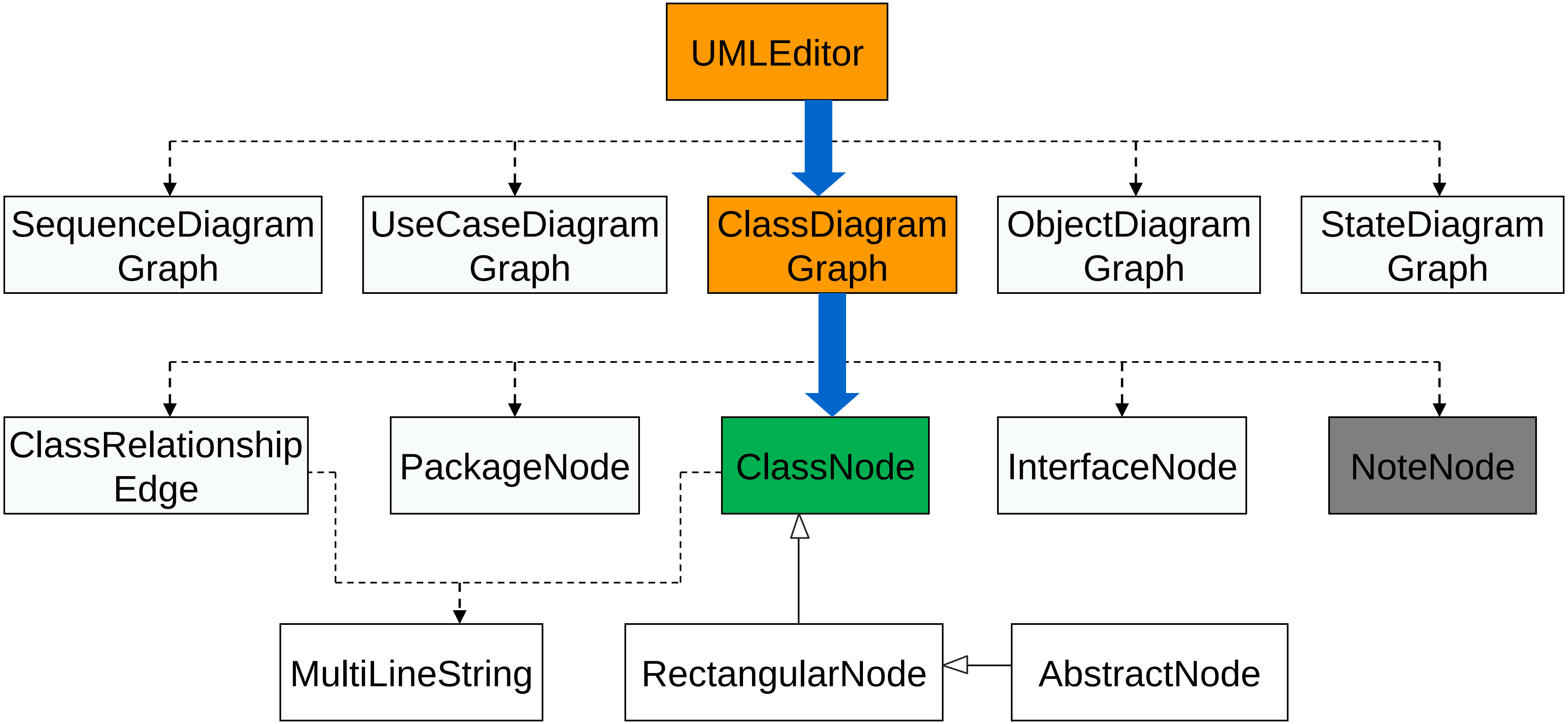
Wrong way


UMLEditor
SequenceDiagram Graph
UseCaseDiagram Graph
ClassDiagram Graph
ObjectDiagram Graph
StateDiagram Graph
ClassRelationship Edge
PackageNode
ClassNode
InterfaceNode
NoteNode
MultiLineString
RectangularNode
AbstractNode

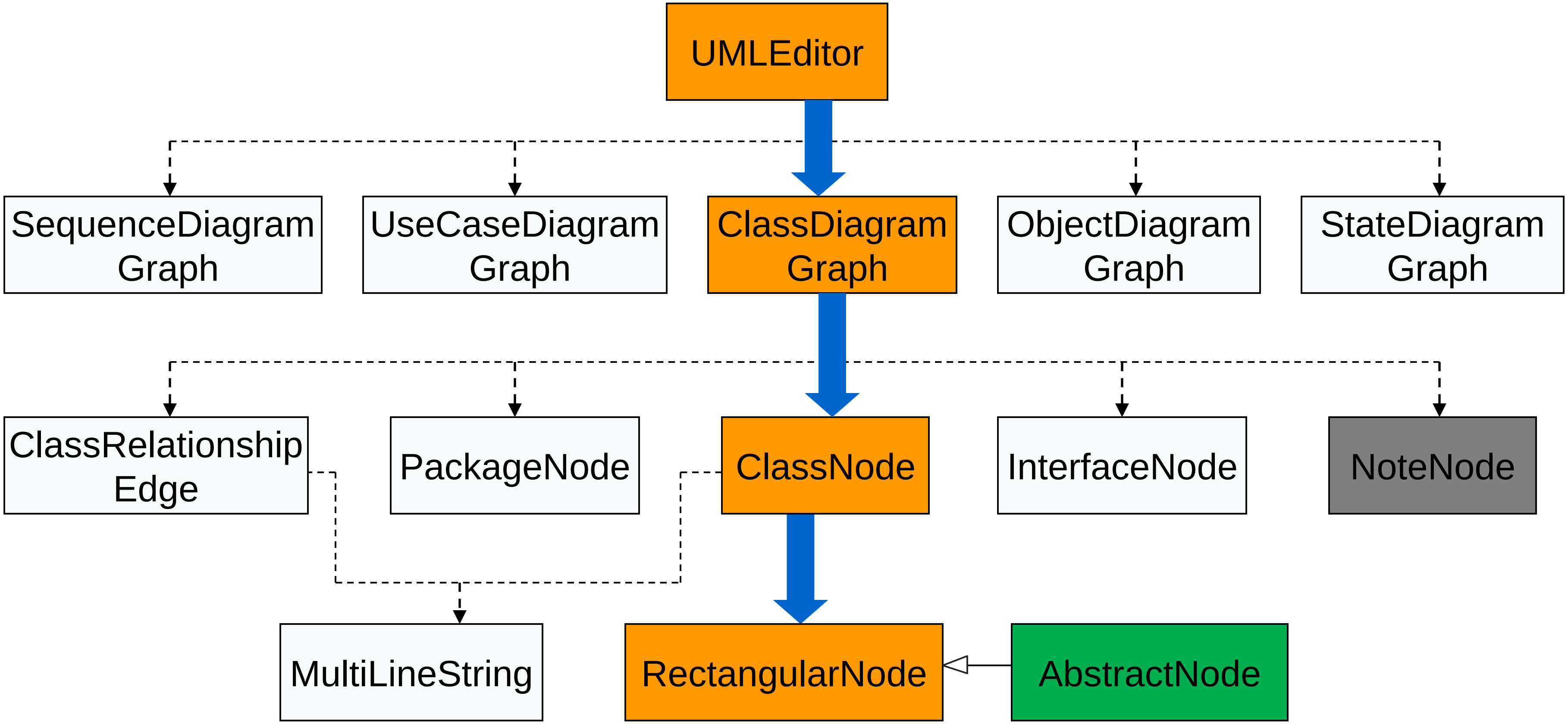
Backtrack


UMLEditor
SequenceDiagram Graph
UseCaseDiagram Graph
ClassDiagram Graph
ObjectDiagram Graph
StateDiagram Graph
ClassRelationship Edge
PackageNode
ClassNode
InterfaceNode
NoteNode
MultiLineString
RectangularNode
AbstractNode

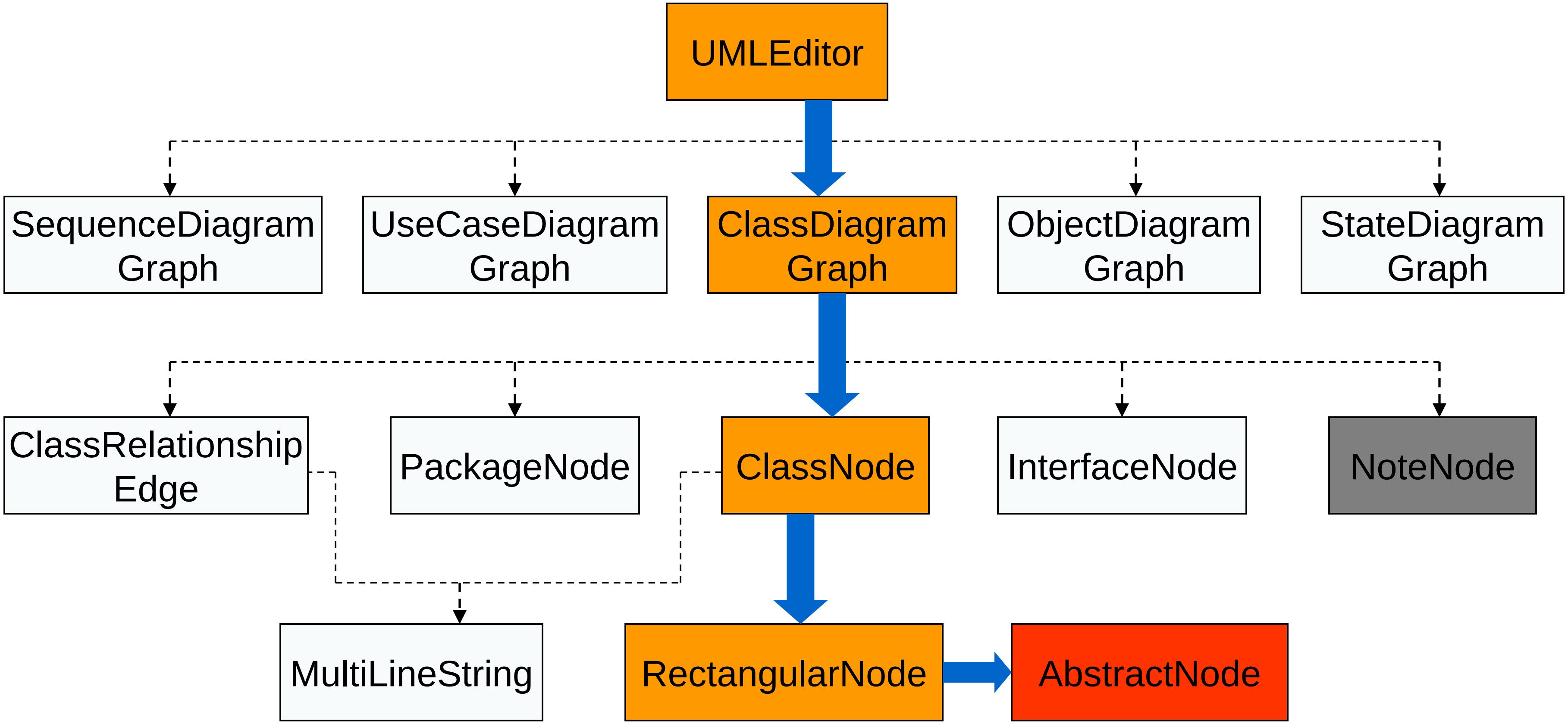
Concept location found



Possible extension of the search



Another location found



Example 2: Eclipse 3.3

- Eclipse – Integrated Development Environment for Java

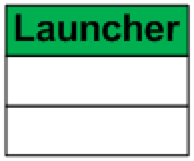
- 15,479 classes, 156,334 functions

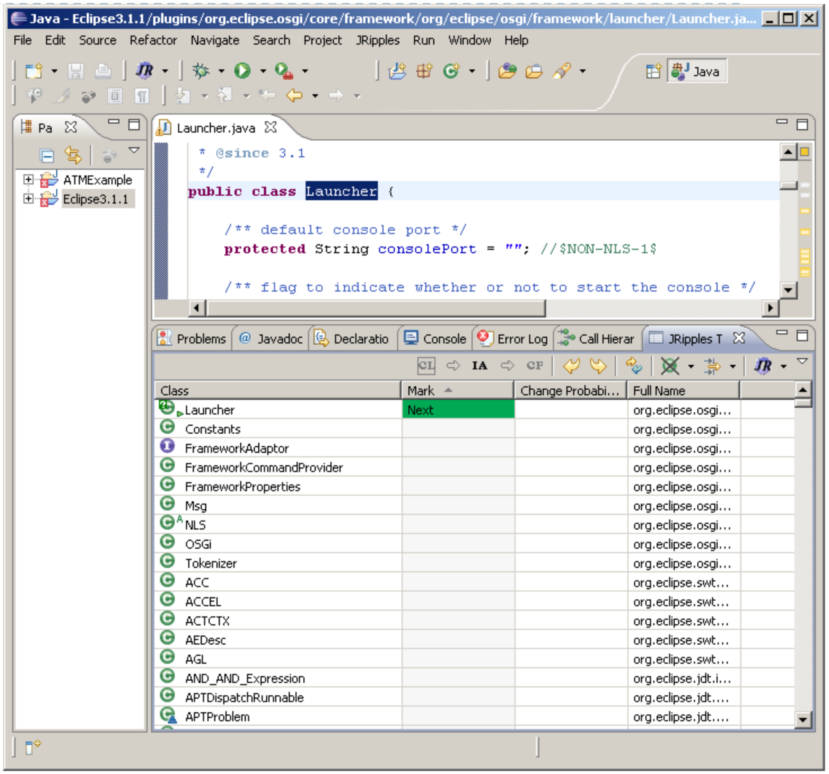
- Change request: BugID 172261*: *[Actions] When rename a file in one project's navigator, the other selected file's name is renamed*

- Concept location technique: CLDS

*https://bugs.eclipse.org/bugs/show_bug.cgi?id=172261

CLDS in Eclipse 3.3

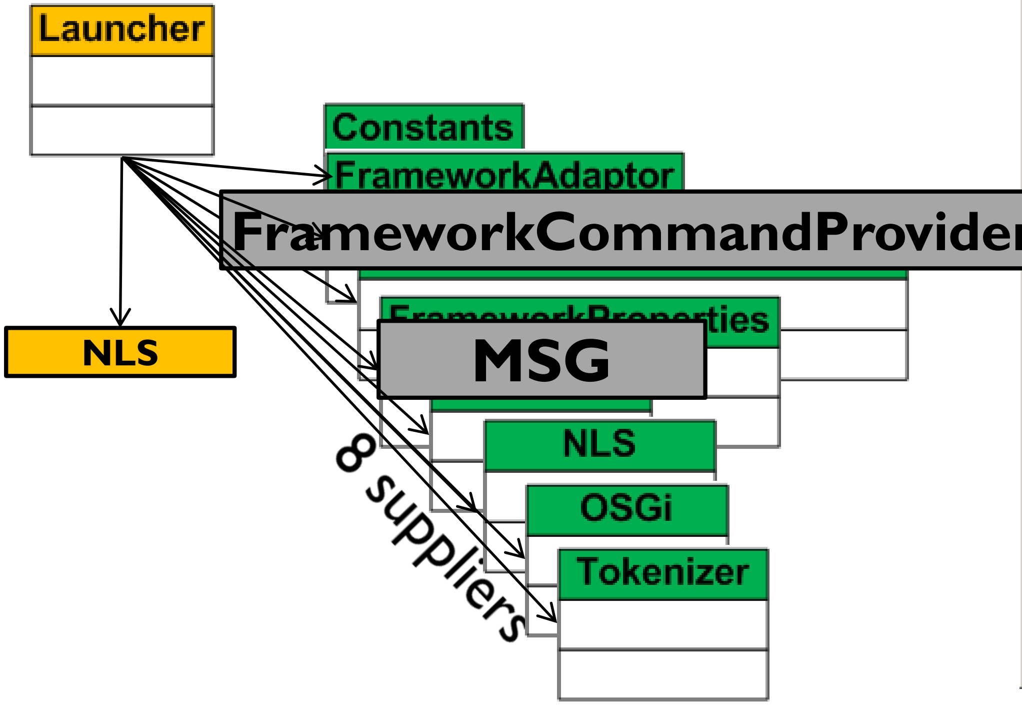


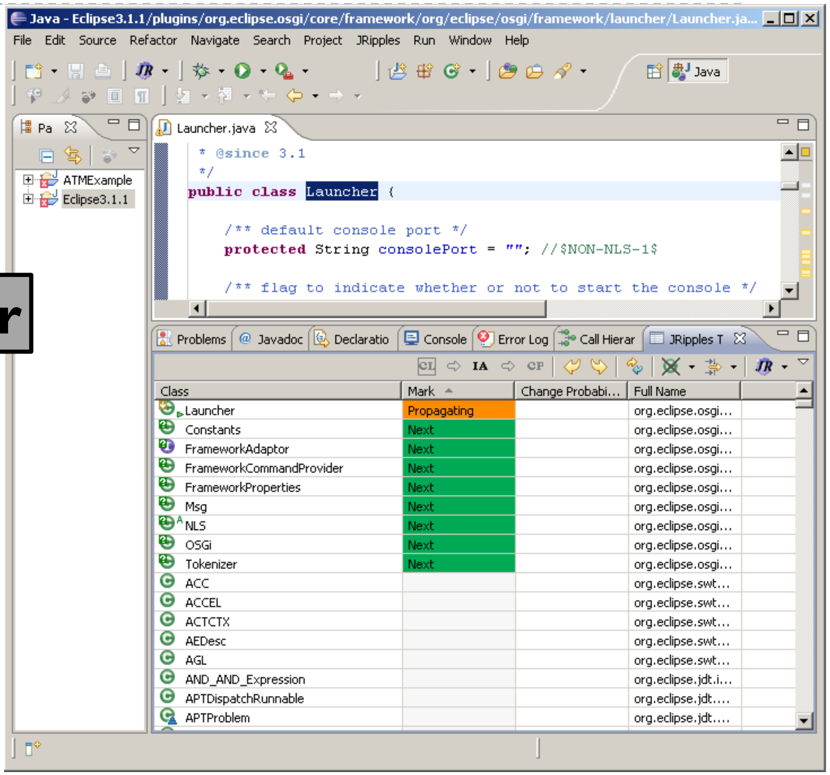

Java - Eclipse3.1.1/plugins/org.eclipse.osgi/core/framework/org/eclipse/osgi/framework/launcher/Launcher.ja...
File Edit Source Refactor Navigate Search Project JRipples Run Window Help
Java
Pa
ATMExample
Eclipse3.1.1
Launcher.java
Problems
Javadoc
Declaratio
Console
Error Log
Call Hierar
JRipples T
CL
IA
CP

```
 * @since 3.1
 */
public class Launcher {

    /** default console port */
    protected String consolePort = ""; //$NON-NLS-1$

    /** flag to indicate whether or not to start the console */
```

Class	Mark	Change Probabi...	Full Name
Launcher	Next		org.eclipse.osgi...
Constants			org.eclipse.osgi...
FrameworkAdaptor			org.eclipse.osgi...
FrameworkCommandProvider			org.eclipse.osgi...
FrameworkProperties			org.eclipse.osgi...
Msg			org.eclipse.osgi...
NLS			org.eclipse.osgi...
OSGi			org.eclipse.osgi...
Tokenizer			org.eclipse.osgi...
ACC			org.eclipse.swt...
ACCEL			org.eclipse.swt...
ACTCTX			org.eclipse.swt...
AEDesc			org.eclipse.swt...
AGL			org.eclipse.swt...
AND_AND_Expression			org.eclipse.jdt.i...
APTDispatchRunnable			org.eclipse.jdt....
APTProblem			org.eclipse.jdt....

CLDS in Eclipse 3.3

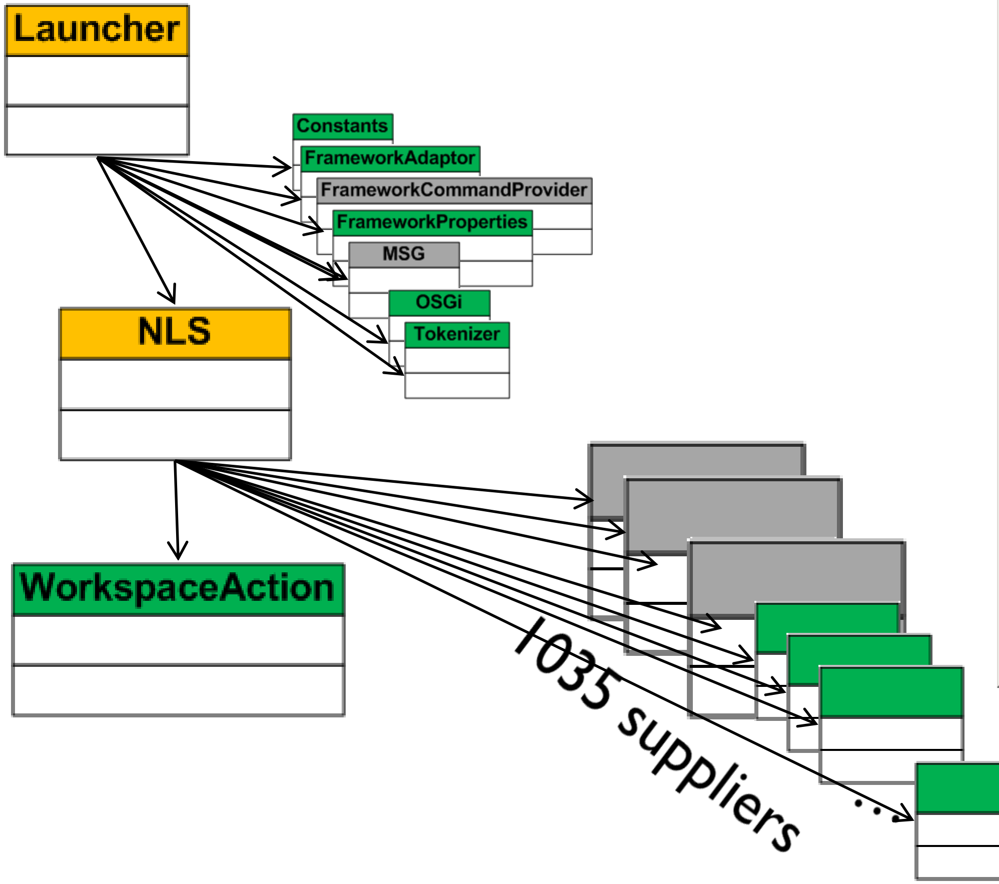


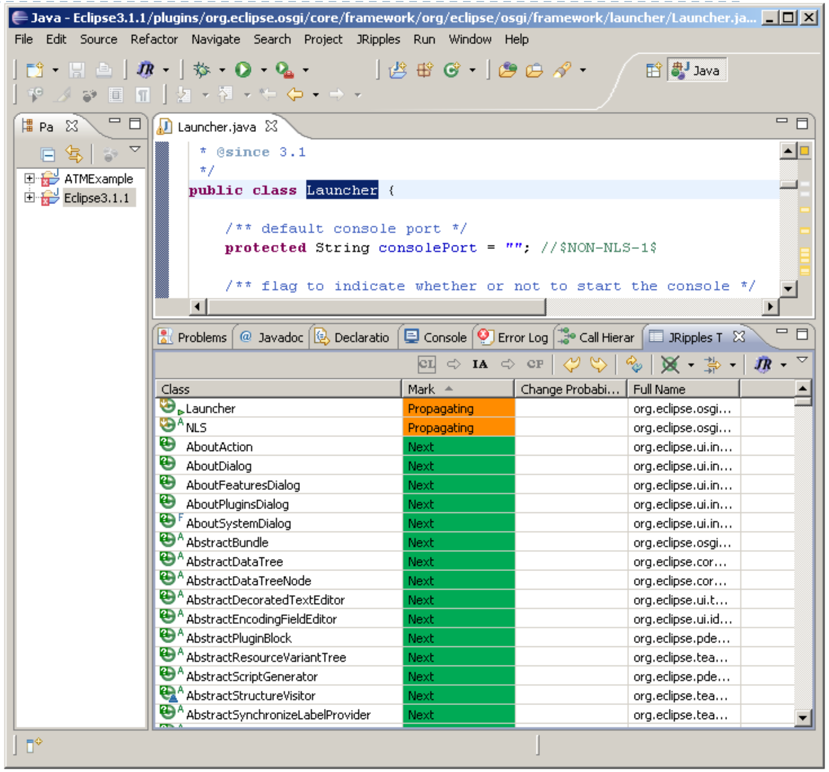

Java - Eclipse3.1.1/plugins/org.eclipse.osgi/core/framework/org/eclipse/osgi/framework/launcher/Launcher.ja...
File Edit Source Refactor Navigate Search Project JRipples Run Window Help
Java
Pa
ATMExample
Eclipse3.1.1
Launcher.java
Problems
Javadoc
Declaratio
Console
Error Log
Call Hierar
JRipples T
CL
IA
CP

```
 * @since 3.1
 */
public class Launcher {

    /** default console port */
    protected String consolePort = ""; //$NON-NLS-1$

    /** flag to indicate whether or not to start the console */
```

Class	Mark	Change Probabi...	Full Name
Launcher	Propagating		org.eclipse.osgi...
Constants	Next		org.eclipse.osgi...
FrameworkAdaptor	Next		org.eclipse.osgi...
FrameworkCommandProvider	Next		org.eclipse.osgi...
FrameworkProperties	Next		org.eclipse.osgi...
Msg	Next		org.eclipse.osgi...
NLS	Next		org.eclipse.osgi...
OSGi	Next		org.eclipse.osgi...
Tokenizer	Next		org.eclipse.osgi...
ACC			org.eclipse.swt...
ACCEL			org.eclipse.swt...
ACTCTX			org.eclipse.swt...
AEDesc			org.eclipse.swt...
AGL			org.eclipse.swt...
AND_AND_Expression			org.eclipse.jdt.i...
APTDispatchRunnable			org.eclipse.jdt....
APTProblem			org.eclipse.jdt....

CLDS in Eclipse 3.3

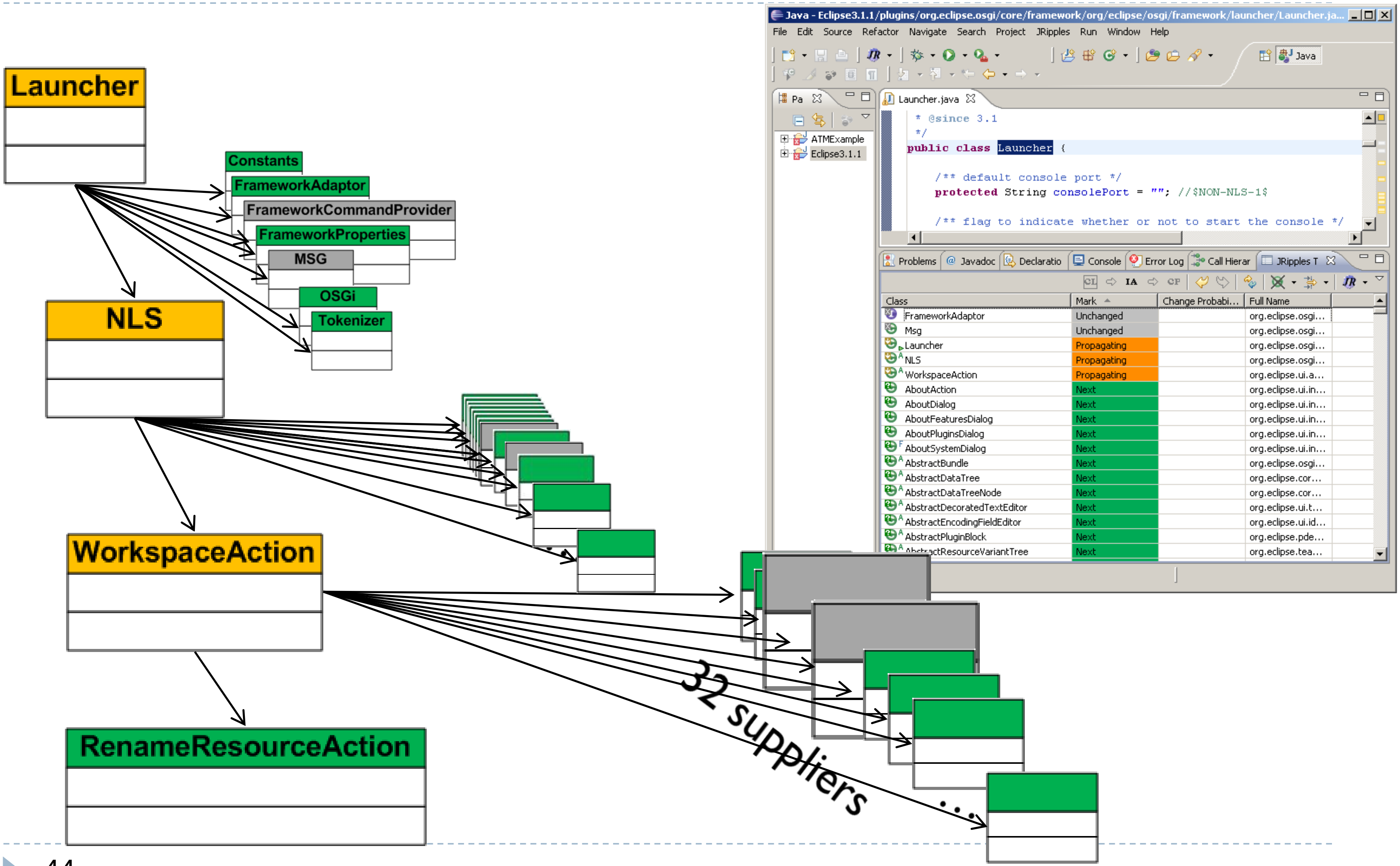



Java - Eclipse3.1.1/plugins/org.eclipse.osgi/core/framework/org/eclipse/osgi/framework/launcher/Launcher.ja...
File Edit Source Refactor Navigate Search Project JRipples Run Window Help
Java
Pa
ATMExample
Eclipse3.1.1
Launcher.java
Problems
Javadoc
Declaratio
Console
Error Log
Call Hierar
JRipples T
CL
IA
CP

```
 * @since 3.1
 */
public class Launcher {

    /** default console port */
    protected String consolePort = ""; //$NON-NLS-1$

    /** flag to indicate whether or not to start the console */
```

Class	Mark	Change Probabi...	Full Name
Launcher	Propagating		org.eclipse.osgi...
NLS	Propagating		org.eclipse.osgi...
AboutAction	Next		org.eclipse.ui.in...
AboutDialog	Next		org.eclipse.ui.in...
AboutFeaturesDialog	Next		org.eclipse.ui.in...
AboutPluginsDialog	Next		org.eclipse.ui.in...
AboutSystemDialog	Next		org.eclipse.ui.in...
AbstractBundle	Next		org.eclipse.osgi...
AbstractDataTree	Next		org.eclipse.cor...
AbstractDataTreeNode	Next		org.eclipse.cor...
AbstractDecoratedTextEditor	Next		org.eclipse.ui.t...
AbstractEncodingFieldEditor	Next		org.eclipse.ui.id...
AbstractPluginBlock	Next		org.eclipse.pde...
AbstractResourceVariantTree	Next		org.eclipse.tea...
AbstractScriptGenerator	Next		org.eclipse.pde...
AbstractStructureVisitor	Next		org.eclipse.tea...
AbstractSynchronizeLabelProvider	Next		org.eclipse.tea...

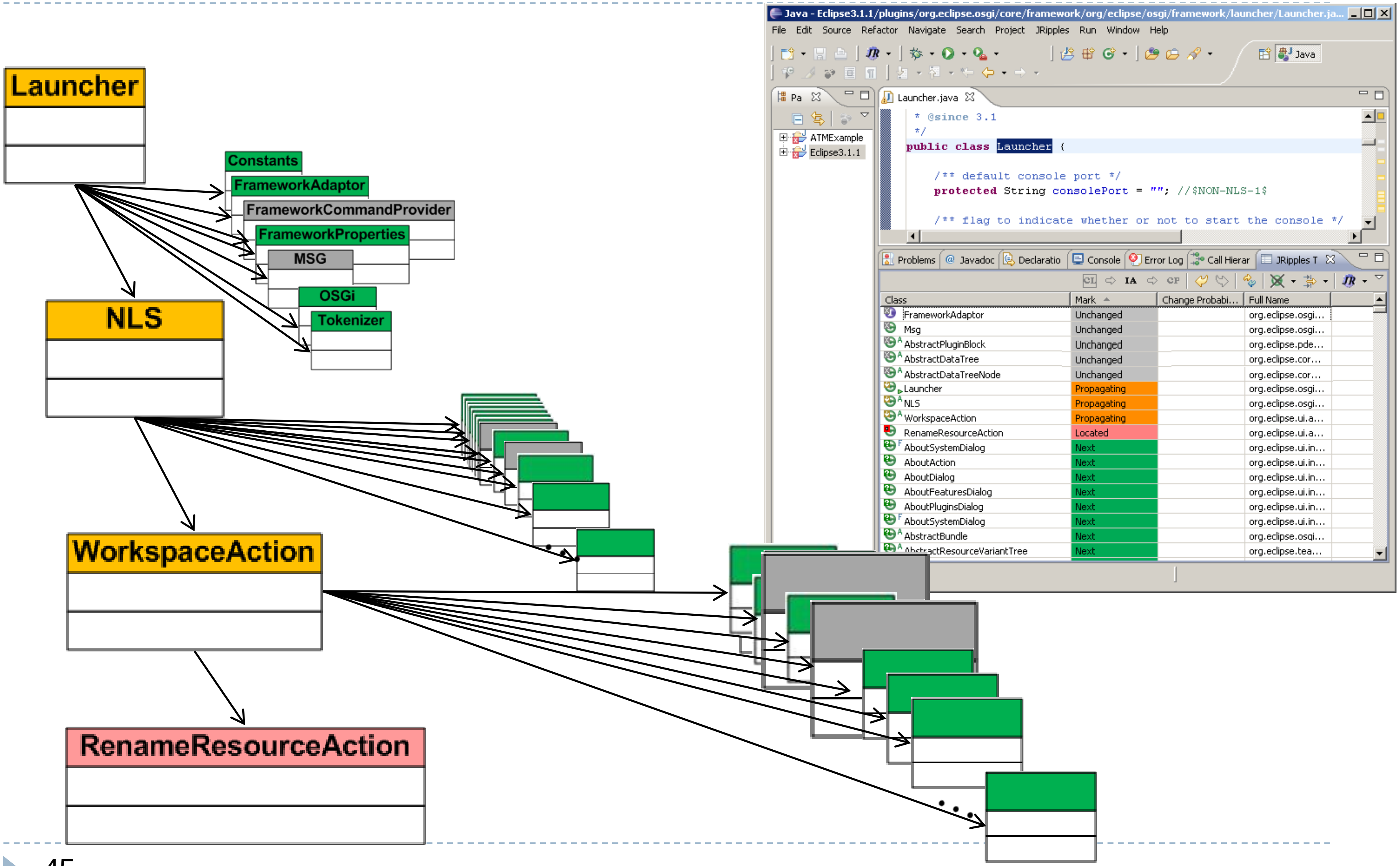
CLDS in Eclipse 3.3



```
 * @since 3.1
 */
public class Launcher {

    /** default console port */
    protected String consolePort = ""; //$NON-NLS-1$

    /** flag to indicate whether or not to start the console */
```

Class	Mark	Change Probabi...	Full Name
FrameworkAdaptor	Unchanged		org.eclipse.osgi...
Msg	Unchanged		org.eclipse.osgi...
Launcher	Propagating		org.eclipse.osgi...
NLS	Propagating		org.eclipse.osgi...
WorkspaceAction	Propagating		org.eclipse.ui.a...
AboutAction	Next		org.eclipse.ui.in...
AboutDialog	Next		org.eclipse.ui.in...
AboutFeaturesDialog	Next		org.eclipse.ui.in...
AboutPluginsDialog	Next		org.eclipse.ui.in...
AboutSystemDialog	Next		org.eclipse.ui.in...
AbstractBundle	Next		org.eclipse.osgi...
AbstractDataTree	Next		org.eclipse.cor...
AbstractDataTreeNode	Next		org.eclipse.cor...
AbstractDecoratedTextEditor	Next		org.eclipse.ui.t...
AbstractEncodingFieldEditor	Next		org.eclipse.ui.id...
AbstractPluginBlock	Next		org.eclipse.pde...
AbstractResourceVariantTree	Next		org.eclipse.tea...

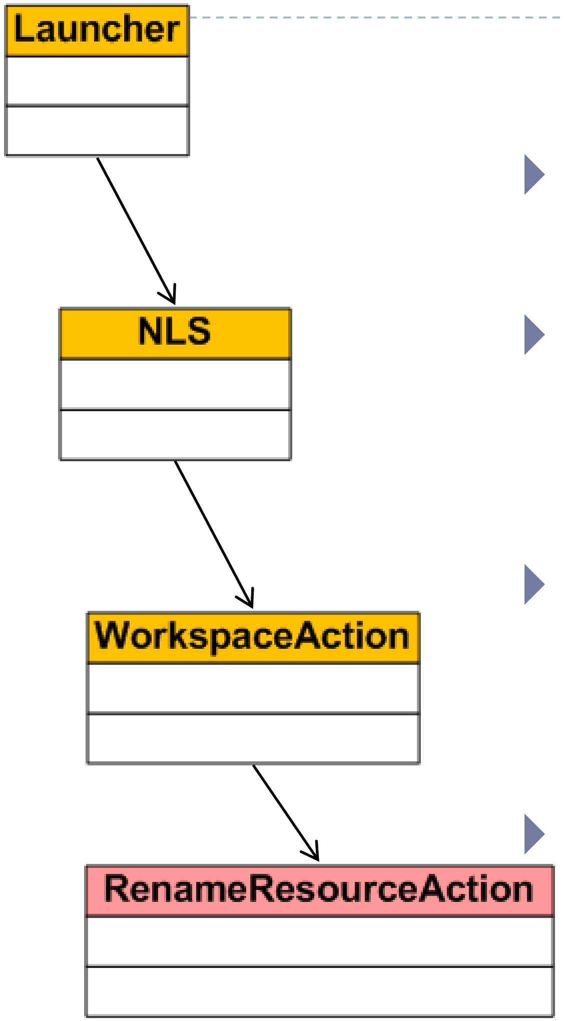
CLDS in Eclipse 3.3



```
 * @since 3.1
 */
public class Launcher {

    /** default console port */
    protected String consolePort = ""; //$NON-NLS-1$

    /** flag to indicate whether or not to start the console */
```

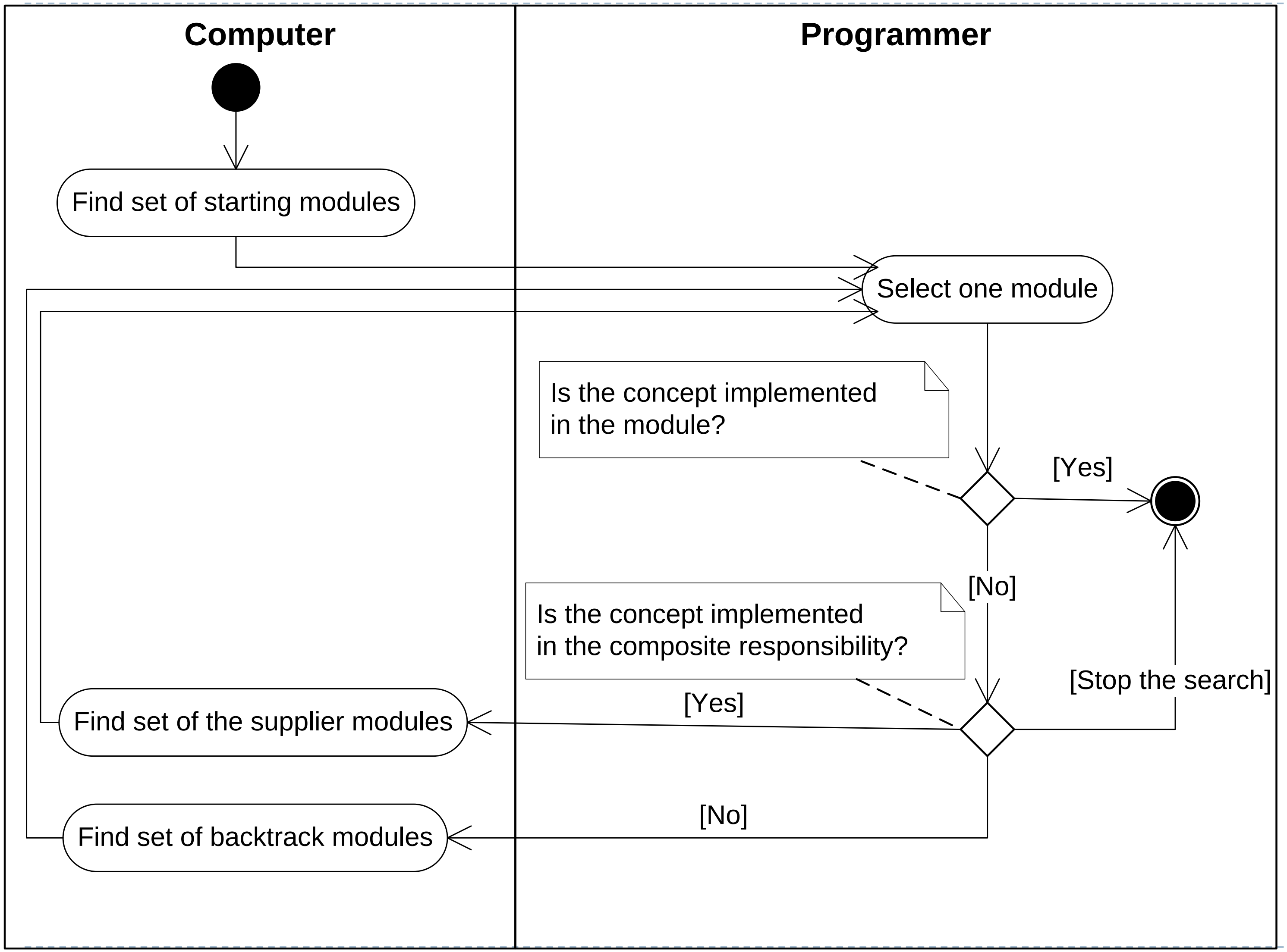
Class	Mark	Change Probabi...	Full Name
FrameworkAdaptor	Unchanged		org.eclipse.osgi...
Msg	Unchanged		org.eclipse.osgi...
AbstractPluginBlock	Unchanged		org.eclipse.pde...
AbstractDataTree	Unchanged		org.eclipse.cor...
AbstractDataTreeNode	Unchanged		org.eclipse.cor...
Launcher	Propagating		org.eclipse.osgi...
NLS	Propagating		org.eclipse.osgi...
WorkspaceAction	Propagating		org.eclipse.ui.a...
RenameResourceAction	Located		org.eclipse.ui.a...
AboutSystemDialog	Next		org.eclipse.ui.in...
AboutAction	Next		org.eclipse.ui.in...
AboutDialog	Next		org.eclipse.ui.in...
AboutFeaturesDialog	Next		org.eclipse.ui.in...
AboutPluginsDialog	Next		org.eclipse.ui.in...
AboutSystemDialog	Next		org.eclipse.ui.in...
AbstractBundle	Next		org.eclipse.osgi...
AbstractResourceVariantTree	Next		org.eclipse.tea...

CLDS - results



- Main class: Launcher
- Minimum possible number of classes visited = 4
- Number of classes proposed to be visited =1076 (1+8+1035+32)
- Programmers can easily get lost when JRipples requires them to visit a large number of suppliers for a class

Interactive tool for concept location

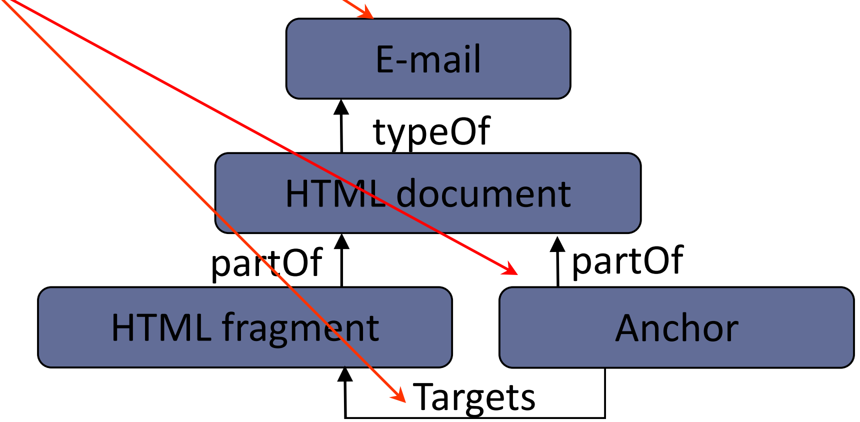

Computer
Programmer
Find set of starting modules
Select one module
Is the concept implemented
in the module?
[Yes]
[No]
Is the concept implemented
in the composite responsibility?
[Stop the search]
[Yes]
Find set of the supplier modules
[No]
Find set of backtrack modules

Concept location by grep

- Classical technique for concept location
  - based on pattern matching
- Programmer formulates a query
  - based on concept name(s)
- Grep searches the files
  - finds corresponding lines of code (“hits”)
  - programmer investigates the hits
- If a search fails, new query is tried
  - programmer learns from failed search

Example: “Anchors” bug in Mozilla*

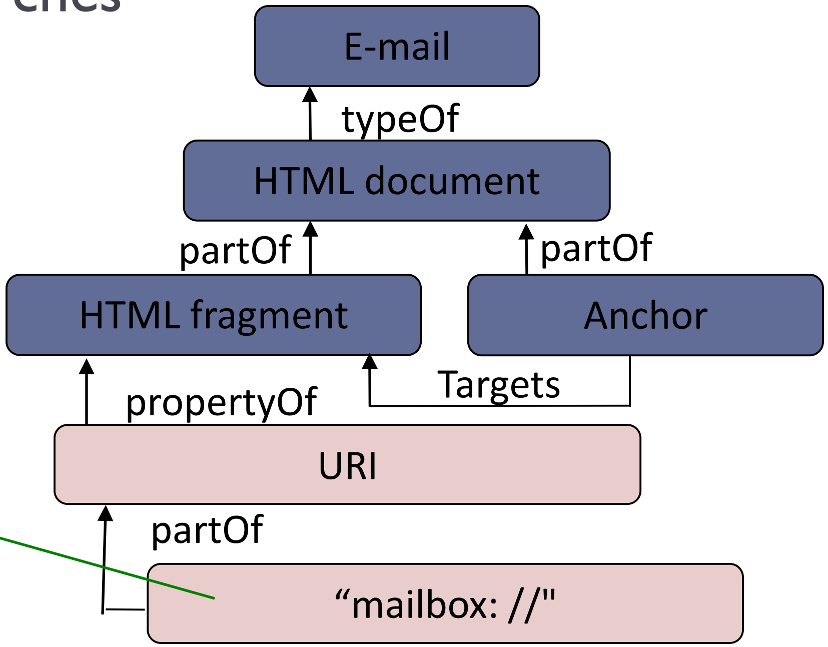
- Change request
  - *Anchors in e-mails are broken (Clicking Anchor doesn't go to target in e-mail)*
- Initial knowledge



*M. Petrenko, V. Rajlich, R. Vanciu, “Partial Domain Comprehension in Software Evolution and Maintenance”, ICPC 2008, 13-22

Knowledge after learning

- Result of repeated search
  - Programmer learns from the searches



Winning grep query:
mailbox: //