CLICS: A Tool for the Investigation of Duplication in Software

Cory Kapser and Michael W. Godfrey
Software Architecture Group (SWAG)
School of Computer Science, University of Waterloo
{cjkapser, migod}@uwaterloo.ca

1 Introduction

Code duplication, or code cloning, is generally believed to be common in large industrial systems. Code cloning is the copying of code in order to duplicate behaviors or algorithms in the software. It can be an indication various problems in a system such as poor design or increasing difficulty in maintenance. Problems associated with cloning include increased code size, copied bugs and increased maintenance effort due to multiple copies of code. Management of code cloning and the various problems associated with it is important for the successful evolution of software systems.

One problem when analyzing clones is that detection methods return a large set of suspected clones, but provide little or no additional information about them to aid the user in their interpretation. This makes clone management cumbersome at best and intractable in general.

The tool we will demonstrate is the Clone Interpretation and Navigation System (CLICS). This tool tries to aid in the process of clone filtering and classification by using automatic classification of clones and tying them to system architecture. This work has been developed through case studies on large software systems.

2 Goals

The core challenge to the maintenance and management of cloning in software systems is comprehending the actual types of clones and the dependencies they create in the software system. However, clone detection tools can return very large result sets and viewing every possible clone is infeasible. To address this problem we aim to provide:

- facilities to evaluate overall cloning activity.
- mechanisms to guide users toward clones that will be most effectively used in their task.
- methods for filtering and refining the analysis of the clones.

3 Features

The tool provides features that allow varying levels of abstraction in the analysis. High level information is provided for initial system evaluation. Navigation mechanisms are provided to guide user to the most interesting and relevant clones within the system. The navigation features provide multiple levels of abstraction when viewing information and results. The tool also provides automatic filtering using filters that have been developed through several case studies. The user is also able to manually filter clones, as well as remove and add subsystems from and to the analysis.

The high level information provided can be used as a way to for an initial evaluation of the system and estimate the cost required to manage or investigate the clones in the system. Information is provided in the form of metrics encompassing several aspects of the system.

There are several modes of navigation in CLICS. These modes include navigatable system architecture graphs, clone classification trees, and system architecture trees. Each of these modes uses a hierarchy of containment in order to provide multiple levels of abstraction in the analysis.

Currently only limited query support is implemented in CLICS. CLICS supports querying clones based on location, based on clones relations to code segments, and clones of a particular size. All query results are displayed in a categorized clone navigation tree.

4. More Information

For more information on the tool, and the work related to it, please see [1].

References