A Software Package for Generating Code Coverage Reports With Gcov

Zhenmai Hu
Dr. Michael D. Adams, Supervisor
Dr. Wu-Sheng Lu, Committee Member
Department of Electrical and Computer Engineering
University of Victoria
1. Introduction
   ❖ The Importance of Software Testing
2. Background
   ❖ Software Testing
     • Software Testing Methods
     • Structural Coverage Analysis
     • Code Coverage Criteria
   ❖ GCC and Gcov
   ❖ Name Mangling
3. The Gcov Report Generator (GRG) Software
   ❖ Software Introduction
   ❖ The GRG Library
   ❖ The Application Program Coverage
   ❖ Usage Example
   ❖ Software Installation
4. Conclusion and Future Work
Introduction

❖ The Importance of Software Testing

1. The European Space Agency’s Ariane 5 Flight 501 failed forty seconds after takeoff on June 4, 1996.
2. The NASA Mars Climate Orbiter approached Mars at the wrong angle when entering the upper atmosphere and disintegrated in 1998.
Software testing considers many aspects of software behavior. For example, testing considers if the software being tested:
• responds correctly to all required inputs;
• has acceptable performance consumption towards time and memory;
• is sufficiently usable for applications;
• works appropriately in all intended environments;
• meets the general needs of users; and
• achieves the design and development requirements without side effects.
Background - Software Testing

- Structural Coverage Analysis

**Structural coverage analysis** is frequently utilized to evaluate testing thoroughness by determining the exercised code structure in testing procedures. It can typically be classified into two categories: **control flow criteria** (the flow of operations and paths performed during software execution) and **data flow criteria** (flow of data through variable assignments and references).

A **control-flow graph (CFG)** is a directed graph depicting all execution paths in code.

Each **node** in a CFG is a straight-line code sequence with one entry point and one exit point, and represents a **basic block** in computer programming.

Each **edge** that connects nodes in a CFG, corresponds to a **branch** in code.

A **condition** is a boolean expression that does not contain boolean operators like AND, OR and NOT. A **decision** is a boolean expression composed of conditions and zero or more boolean operators.
Background - Software Testing

❖ Code Coverage Criteria

Code coverage testing measures the percentage of the program source code exercised based when running specific test suites.

Code coverage criteria is the requirement that a test suite needs to fulfill during software testing

- **Function Coverage** = \( \frac{\text{number of functions executed}}{\text{total number of functions}} \times 100\% \)

- **Statement Coverage** = \( \frac{\text{number of statements executed}}{\text{total number of statements}} \times 100\% \)

- **Branch Coverage** = \( \frac{\text{number of branches executed}}{\text{total number of branches}} \times 100\% \)

Full coverage: any particular coverage type with 100% coverage

---

```plaintext
if(a < 0) {
  x = 1;
}
if(b < 0) {
  x = 2;
}
x++;
```

<table>
<thead>
<tr>
<th>a = -1</th>
<th>b = -1</th>
</tr>
</thead>
<tbody>
<tr>
<td>a = 1</td>
<td>b = 1</td>
</tr>
</tbody>
</table>

---

University of Victoria
Background - GCC and Gcov

 GCC

The GNU Compiler Collection (GCC), includes the front ends and libraries to support several programming languages such as C, C++, Objective-C, and etc. It needs to add [-fprofile-arcs] and [-ftest-coverage] or [--coverage] to support Gcov.

 Gcov

Gcov is a code coverage tool that comes as a utility from GCC and only works on code compiled with GCC. It uses two types of files, gcno and gcda, to generate coverage information that can be used to analyze various types of code coverage.

The node files with extension gcno produced during compilation, containing the data needed to reproduce the graph of basic blocks and the corresponding source line number to each block.

The count files with extension gcda produced at program termination, containing the execution count for every basic block and branch and some summary information.

Let’s take hello_world.cpp as an example:

```cpp
#include <iostream>

int main() {
    std::cout << "Hello, World!\n";
    int status;
    if (std::cout) {
        status = 0;
    } else {
        status = 1;
    }
    return status;
}
```
Background - GCC and Gcov

1. Compile the file with --coverage using the GCC C++ compiler. The gcno file will be generated in this step.
   
   ```
   g++ --coverage hello_world.cpp -o hello_world
   ```
   
2. Run the program generated and print the string “Hello, World!” The gcda file will be generated in this step.
   
   ```
   ./hello_world
   ```
   
3. Use gcov command to generate summary information and the coverage report hello_world.gcda.
   
   ```
   gcov hello_world
   ```

Summary information of hello_world generated with gcov

```
File 'hello_world.cpp'
Lines executed: 83.33% of 6
Creating 'hello_world.cpp.gcov'
File '/usr/include/c++/10/iostream'
No executable lines
Removing 'iostream.gcov'
```
Background - GCC and Gcov

Gcov provides numerous command-line options for obtaining additional information to the coverage report.

For example: `gcov hello_world -a -b -j`

- `-a` or `--all-blocks` : adds the execution count for every basic block in the coverage report
- `-b` or `--branch-probabilities` : shows the branch frequencies in percent in the coverage report
- `-j` or `--json-format` : generates the coverage data in JSON format.
Background – Name Mangling

- Name Mangling

In C++, multiple functions can have the same name in the source code due to overloading or visibility within different scopes. To differentiate them when linking and compiling, the compiler using a specific set of rules assigns each function in the source code a unique name is a process known as name mangling.

The unique name produced by name mangling is called the mangled name. The original function name in the source code is called the demangled name.

For example, in the Itanium application binary interface (ABI) standard for C++, the functions can be mapped as:

```cpp
int add (int) _Z3addi
int add (double) _Z3addd
```

Specifically, the compiler can implement each constructor using multiple functions, and the same holds for destructors. These associated functions of a constructor or destructor typically have the same function name in the source code but different mangled name. For example, a virtual base destructor with the demangled name `base::~base()` can have two different mangled names: `_ZN4baseD0Ev` and `_ZN4baseD2Ev`. 
Software

Software Introduction

The Gcov Report Generator (GRG) software consists of:

• A C++ header-only library to generate nicely-formatted customized code-coverage report in PDF format.
• Program called coverage to run the library through command line.

The code coverage report generated in PDF format is formatted as:

• The summary section includes six summary tables for function, statement, and branch coverage on per-file and per-function basis (eg: Per-File Function Coverage, Per-Func Function Coverage).
• The detail section includes the function source code and detailed coverage information (eg: execution counts of lines and branches).

The important customized features provided by the GRG are as follow:

• filter coverage results using file and function patterns (i.e., regular expressions);
• calculate the branch coverage with or without exceptional branches;
• format the code coverage report by keeping only the selected section contents;
• select the coverage information included by setting the function, statement, and branch coverage threshold on a per-file and per-function basis; and
• aggregate coverage results for functions with the same names in source code but different mangled names.
The GRG Library

Application programming interface (API) of the GRG library: One primary class Report_maker (constructor + function call operator) + Two secondary classes Global_options and Report_options as input parameters.
The Application Program Coverage

Coverage is controlled by numerous command line options. There are two types of these options: Global Options and Per-Report Options. Global options apply to all reports generated by the program, while per-report options apply only to a single report.

Global Options

- `-c` or `--aggregate-constructors-destructors`
  When this option is specified, coverage aggregates all of the C++ compiler-generated functions from one particular C++ constructor as if they were the same function when determining code coverage. This option also has a similar effect on destructors.

- `-t` or `--aggregate-templates`
  When this option is specified, coverage aggregates all instantiations of the same template function as if they were the same function when determining code coverage.

- `-h` or `--help`
  When this option is specified, coverage prints help information about using coverage to the standard output and exits without doing any further processing.

Global Options — Function Selection

- `-x REGEX` or `--file-pattern REGEX`
  When this option is specified, coverage sets the current value of the file pattern to the extended regular expression specified by REGEX. The value of REGEX defaults to string "/(.*)/" (i.e., select all files).

- `-y REGEX` or `--function-pattern REGEX`
  When this option is specified, coverage sets the current value of the function pattern to the extended regular expression specified by REGEX. The value of REGEX defaults to string "/(.*)/" (i.e., select all functions).

- `-p EXPRESSION` or `--pattern EXPRESSION`
  When this option is specified, coverage adds a new file-function pattern to the pattern list that consists of the file-function patterns. The EXPRESSION is a string chosen from “keep” or “delete”, through which we can determine whether to keep or discard the current file-function pattern when generating the report. The value of EXPRESSION defaults to keep, which means the file-function pattern that matches the EXPRESSION will be kept.

- `-r` or `--reset`
  When this option is specified, coverage resets all of the per-report options to their default values.
The Application Program Coverage

**Per-Report Options**

- **-n DIRNAME or --temporary-directory-name DIRNAME**
  When this option is specified, coverage sets the DIRNAME as the name of the temporary directory used to store the LaTeX source file. The directory specified must already exist. For example, "/tmp/foobar" or "apple" can be used. If this option is not specified, the system directory for temporary files will be used (i.e., the directory returned by the C++ standard library function std::filesystem::temp_directory_path).

- **-k or --keep-temporary-directory**
  When this option is specified, coverage keeps the temporary directory that contains the LaTeX source file. If this option is not specified, the directory will be deleted.

- **-e or --ignore-exceptional-branches**
  When this option is specified, coverage treats exceptional branches as if they do not exist when determining branch coverage.

- **-o PATHNAME or --output PATHNAME**
  When this option is specified, coverage sets PATHNAME as the output pathname to which to write the code coverage report. If not specified, the report is printed to standard output.

- **-m PROGRAM or --pdflatex-program PROGRAM**
  This option sets PROGRAM as the pathname of the program to use to convert LaTeX source to a PDF document. Any program that has a CLI compatible with pdflatex can be used. The pdflatex, lualatex, or pdflatex_frontend are possible values for PROGRAM. The value of PROGRAM defaults to pdflatex.

**Per-Report Options – Report Formatting**

- **-d or --detail**
  When this option is specified, coverage includes the detail section in the report, which contains detailed coverage information for each selected function.

- **--function-coverage-summary-per-file**
  When this option is specified, coverage includes the per-file function/branch coverage summary table in the code coverage report.

- **--statement-coverage-summary-per-function**
  When this option is specified, coverage includes the per-function statement/branch coverage summary table in the code coverage report.

- **--branch-coverage-summary-per-function**
  When this option is specified, coverage includes the per-function statement/branch coverage summary table in the code coverage report.

- **--function-coverage-summary**
  This option works for selecting both the --function-coverage-summary-per-file and the --function-coverage-summary-per-function options (same for statement and branch).
The Application Program Coverage

Per-Report Options — Function Selection

- **-F PERCENT or --func-cov-per-file-threshold PERCENT**
  When this option is specified, coverage selects all functions in source files with function coverage less than the real number PERCENT, of which the value starts from 0 and is allowed to be greater than 100. The value of the PERCENT defaults to 101. The lines in the selected function are marked in the detail section of the report.

- **-f PERCENT or --func-cov-per-func-threshold PERCENT**
  When this option is specified, coverage selects all functions with function coverage less than the real number PERCENT, of which the value starts from 0 and is allowed to be greater than 100. The value of the PERCENT defaults to 101. The lines in the selected function are marked in the detail section of the report.

- **-S PERCENT or --state-cov-per-file-threshold PERCENT**
  When this option is specified, coverage selects functions in source files with statement coverage less than the real number PERCENT, of which the value starts from 0 and is allowed to be greater than 100. The value of the PERCENT defaults to 101. The lines in the selected function are marked in the detail section of the report.

- **-s PERCENT or --state-cov-per-func-threshold PERCENT**
  When this option is specified, coverage selects all functions with statement coverage less than the real number PERCENT, of which the value starts from 0 and is allowed to be greater than 100. The value of the PERCENT defaults to 101. The lines in the selected function are marked in the detail section of the report.

- **-B PERCENT or --branch-cov-per-file-threshold PERCENT**
  When this option is specified, coverage selects all functions in source files with branch coverage less than the real number PERCENT, of which the value starts from 0 and is allowed to be greater than 100. The value of the PERCENT defaults to 101. The lines in the selected function are marked in the detail section of the report.

- **-b PERCENT or --branch-cov-per-func-threshold PERCENT**
  When this option is specified, coverage selects all functions with branch coverage less than the real number PERCENT, of which the value starts from 0 and is allowed to be greater than 100. The value of the PERCENT defaults to 101. The lines in the selected function are marked in the detail section of the report.
A source file `template.cpp`:

- a template function called `bubbleSort` to sort the input array with element type specified in ascending order;
- two overloading functions with the same name `sum` used to add two input numbers but differs in the return type and input data type;
- a function called `main` that initializes some variables and calls the other functions.

```cpp
#include <iostream>
#include <type_traits> // std::extent

// A template function
template <class T>
void bubbleSort(T a[], int n) {
    for (int i = 0; i < n - 1; i++) {
        for (int j = n - 1; i < j; j--) {
            if (a[i] < a[j]) {
                std::swap(a[i], a[j]);
            }
        }
    }
}

// Two overloading functions
int sum(int a, int b) {
    return a + b;
}
double sum(double a, double b) {
    return a + b;
}

int main() {
    // Calls template function
    int a1[5] = {3, 5, 1, 2, 4};
    double a2[5] = {2.3, 5.6, 1.1, 2.2, 10.0};
    bubbleSort<int>(a1, std::extent<int[5]>::value);
    bubbleSort<double>(a2, std::extent<int[5]>::value);

    // Calls overload functions and outputs the results
    int a = 1;
    int b = 2;
    double c = 1.2;
    double d = 2.4;
    std::cout << sum(a, b) << '\n';
    std::cout << sum(c, d) << '\n';

    return 0;
}
```
Software

Usage Example

Source code for full_report.pdf that uses the GRG library API with default options

```cpp
#include <iostream>
#include <fstream>
#include "grg/grg.hpp"

int main() {
    // Initialize the global options
    grg::Report_maker::Global_options global_options;
    global_options.aggregate_templates = true;
    File_function_regex cur_pattern = (".*template.cpp", ".*?");
    global_options.filter_patterns.push_back(cur_pattern);
    std::vector<std::string> gcda_files_vec;
    std::string input_file("template.gcda");
    gcda_files_vec.push_back(input_file);

    // Call the constructor of Report_maker class
    grg::Report_maker maker(global_options, gcda_files_vec);

    // Initialize the report options
    grg::Report_maker::Report_options report_options;
    report_options.function_coverage_summary = true;
    report_options.statement_coverage_summary = true;
    report_options.branch_coverage_summary = true;
    report_options.detail = true;

    // Call the operator() to generate the coverage report stream
    // and write it to a PDF file named full_report.pdf
    std::ofstream myfile("full_report.pdf", std::ios::out);
    int status = maker(report_options, myfile);
    myfile.close();

    return status;
}
```

Source code for the file specified_report.pdf that uses the GRG library API with particular options

```cpp
#include <iostream>
#include <fstream>
#include "grg/grg.hpp"

int main() {
    // Initialize the global options
    grg::Report_maker::Global_options global_options;
    global_options.aggregate_templates = true;
    File_function_regex cur_pattern = (".*template.cpp", ".*?");
    global_options.filter_patterns.push_back(cur_pattern);
    std::vector<std::string> gcda_files_vec;
    std::string input_file("template.gcda");
    gcda_files_vec.push_back(input_file);

    // Call the constructor of Report_maker class
    grg::Report_maker maker(global_options, gcda_files_vec);

    // Initialize the report options
    grg::Report_maker::Report_options report_options;
    report_options.function_coverage_summary = true;
    report_options.statement_coverage_summary = true;
    report_options.branch_coverage_summary = true;
    report_options.detail = true;

    double threshold = 100;
    report_options.state_cov_per_func_threshold = threshold;
    report_options.branch_cov_per_func_threshold = threshold;

    // Call the operator() to generate the coverage report stream
    // and write it to a PDF file named specified_report.pdf
    std::ofstream myfile("specified_report.pdf", std::ios::out);
    int status = maker(report_options, myfile);
    myfile.close();

    return status;
}
```
Usage Example

Run coverage to generate the file full_report.pdf containing all the files and functions with default option settings:

```
./coverage --function-coverage-summary
--statement-coverage-summary
--branch-coverage-summary
--detail
--output full_report.pdf template.gcda
```

Run coverage to generate the file specified_report.pdf with specified file-function pattern and coverage thresholds:

```
./coverage --file-pattern .*template.cpp --pattern keep
--function-coverage-summary
--statement-coverage-summary
--branch-coverage-summary
--detail
--state-cov-per-func-threshold 100
--branch-cov-per-func-threshold 100
--output specified_report.pdf template.gcda
```
Software

Usage Example

Partial contents of the file full_report.pdf generated by GRG software.

1 Summary

1.1 Function Coverage

1.1.1 Pre-File Function Coverage

<table>
<thead>
<tr>
<th>Coverage</th>
<th>Pathname</th>
</tr>
</thead>
<tbody>
<tr>
<td>5/5 (100.0%)</td>
<td>/home/sint/Coverage/tests/template.cpp</td>
</tr>
<tr>
<td>4/4 (100.0%)</td>
<td>/usr/include/c++/5/10/bits/move.h</td>
</tr>
</tbody>
</table>

1.1.2 Pre-Function Function Coverage

<table>
<thead>
<tr>
<th>Coverage</th>
<th>Pathname</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/1 (100.0%)</td>
<td>/home/sint/Coverage/tests/template.cpp</td>
</tr>
<tr>
<td>1/1 (100.0%)</td>
<td>/home/sint/Coverage/tests/template.cpp</td>
</tr>
<tr>
<td>1/1 (100.0%)</td>
<td>/home/sint/Coverage/tests/template.cpp</td>
</tr>
<tr>
<td>1/1 (100.0%)</td>
<td>/home/sint/Coverage/tests/template.cpp</td>
</tr>
<tr>
<td>12/12 (100.0%)</td>
<td>/usr/include/c++/5/10/bits/move.h</td>
</tr>
</tbody>
</table>

1.2 Statement Coverage

1.2.1 Pre-File Statement Coverage

<table>
<thead>
<tr>
<th>Coverage</th>
<th>Pathname</th>
</tr>
</thead>
<tbody>
<tr>
<td>6/6 (100.0%)</td>
<td>/home/sint/Coverage/tests/template.cpp</td>
</tr>
<tr>
<td>0/0 (100.0%)</td>
<td>/usr/include/c++/5/10/bits/move.h</td>
</tr>
</tbody>
</table>

1.2.2 Pre-Function Statement Coverage

<table>
<thead>
<tr>
<th>Coverage</th>
<th>Pathname</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/1 (100.0%)</td>
<td>/home/sint/Coverage/tests/template.cpp</td>
</tr>
<tr>
<td>1/1 (100.0%)</td>
<td>/home/sint/Coverage/tests/template.cpp</td>
</tr>
</tbody>
</table>

2 Details

2.1 /home/sint/Coverage/tests/template.cpp

void bubbleSort(double, double) |

2.2 /usr/include/c++/5/10/bits/move.h

std::remove_reference<bubbleSort<double>::type>::std::move<double>(double)
Software

Usage Example

The file specified_report.pdf generated the by GRG software.
Software

Software Installation
The following versions of required software packages have been verified to work with the GRG software:

- GCC 10.3.0;
- Boost 1.71.0; and
- pdfTeX 1.40.22 (TeX Live 2021).

The GRG software uses CMake for compiling, building, testing, and installation:

1. Generate the native build files by running the command:
   \[\text{cmake -H$\text{SOURCE\_DIR} -B$\text{BUILD\_DIR}}\]
2. Build the GRG software and test files by running the command:
   \[\text{cmake --build $\text{BUILD\_DIR}}\]
3. Test the GRG software built by running the command follows:
   \[\text{cmake --build $\text{BUILD\_DIR} --target test}\]
4. Install the GRG software by running the command:
   \[\text{cmake --build $\text{BUILD\_DIR} --target install}\]

The text like the following will be printed to the standard output upon successful completion of the tests, and the code coverage reports will be presented in the directory named CoverageTestResults under the $\text{SOURCE\_DIR}$ directory.

\[100\% \text{ tests passed , 0 tests failed out of 4}\]
Conclusion and Future Work

• The importance of software testing and the criteria of structural coverage analysis have been studied.

• The GRG software developed based on Gcov containing a library and a program called coverage.

• The GRG software can generate customized code coverage report in PDF format for function, statement, and branch coverage on per-file and per-function basis.

• The GRG software provides rich functionalities, including filtering functions using regular expressions, calculating the branch coverage with or without exceptional branches, selecting the report contents by setting coverage threshold, and aggregating associated functions with different mangled names.

• In the future, the GRG software could allow users to output coverage statistics into a plain text file.

• In addition, the GRG software may need to add mangled function names when presenting the report.
Thank you!

Any Questions?