



- (51) **International Patent Classification:**
G06F 11/36 (2006.01)
- (21) **International Application Number:**
PCT/US2013/025585
- (22) **International Filing Date:**
11 February 2013 (11.02.2013)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
201210033077.X 14 February 2012 (14.02.2012) CN
13/763,497 8 February 2013 (08.02.2013) US
- (71) **Applicant:** ALIBABA GROUP HOLDING LIMITED
[—/US]; Fourth Floor, One Capital Place, P.O. Box 847,
George Town, Grand Cayman (KY).
- (72) **Inventors:** ZHOU, Qi; c/o Alibaba Group Legal Depart-
ment, 10/F, Building A, The West Lake International Plaza
of S&T, No. 391 Wen'er Road, Hangzhou (CN). CAI,
Hua; c/o Alibaba Group Legal Department, 10/F, Building

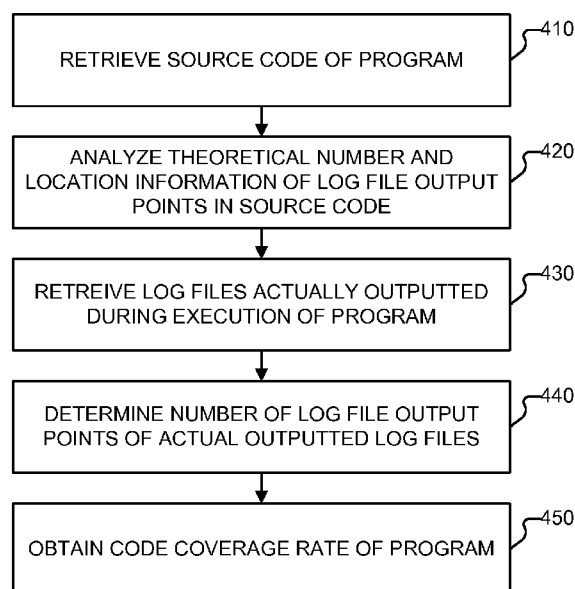
A, The West Lake International Plaza of S&T, No. 391
Wen'er Road, Hangzhou (CN). SUN, Tingtao; c/o Alibaba
Group Legal Department, 10/F, Building A, The West
Lake International Plaza of S&T, No. 391 Wen'er Road,
Hangzhou (CN).

(74) **Agent:** FU, Diana, Y.; Van Pelt, Yi & James LLP, 10050
N. Foothill Blvd., Suite 200, Cupertino, CA 95014 (US).

(81) **Designated States** (*unless otherwise indicated, for every
kind of national protection available*): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,
BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP,
KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD,
ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI,
NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU,
RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ,
TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA,
ZM, ZW.

[Continued on next page]

(54) **Title:** CODE COVERAGE RATE DETERMINATION METHOD AND SYSTEM



400

FIG. 4

(57) **Abstract:** Embodiments of the present application relate to a code coverage rate determination method, a code coverage rate determination system, and a computer program product for determining code coverage rate. A code coverage rate determination method is provided. The method includes retrieving source code of a program, determining theoretical number of log file output points included in the source code and location information of the log file output points, retrieving log files actually outputted during the execution of the program, determining an actual number of log file output points of the actual outputted log files based on the location information of the corresponding log file output points recorded in each log file, and determining a code coverage rate of the program based on the theoretical number of log file output points and the actual number of the log file output points of the actual outputted log files.



(84) **Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK,

SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

CODE COVERAGE RATE DETERMINATION METHOD AND SYSTEM**CROSS REFERENCE TO OTHER APPLICATIONS**

[0001] This application claims priority to People's Republic of China Patent Application No. 201210033077.X entitled CODE COVERAGE RATE DETERMINATION METHOD AND SYSTEM AND CODE COVERAGE DETECTION METHOD AND SYSTEM, filed February 14, 2012 which is incorporated herein by reference for all purposes.

FIELD OF INVENTION

[0002] The present application relates to a code coverage rate determination method and system.

BACKGROUND OF THE INVENTION

[0003] In a computer program testing process, code coverage rate is an indicator used to assess program code coverage. The code coverage rate measures the extent a source code is used during testing. Statistics relating to the code coverage rate enables one to understanding how much of the code in a program has been executed and helps determine where the code coverage is insufficient to thereby discover potential problems with the program.

[0004] One code coverage rate determination method includes tagging runtime situations of code to be tested, take a final count of the quantity of run code executed based on the number of tags encountered, and determine the code coverage rate by comparing the quantity of run code based on the quantity of tags encountered against the total quantity of code. For example, the method comprises: determining a total quantity of code in a program to be tested, a base address of the program to be tested, and a relative virtual address of each line of code of the program; setting up break points in each line of code based on the base address and the relative virtual address; determining whether a break point occurs in the execution of the program to be tested; and tagging the line of code in which the break point has occurred. Finally, the cumulative total of tagged lines of code is used to obtain the total quantity of executed code. The total quantity of executed code divided by the total quantity of code of the program to be tested is the code coverage rate.

[0005] Another method comprises: before a testing process begins, revising a source code and configuration of the program to be tested by adding test information output settings to the

source code for which a code coverage rate is to be determined; during the testing process, internally recording tags output by the source code that has been tested based on the test information output settings added into the source code; and upon conclusion of the testing process, retrieving the source code that has been executed based on the tags, and comparing a quantity of source code executed based on the number of tags output with the total quantity of source code to obtain the code coverage rate.

[0006] The methods described above all add tags to the source code in advance to determine the code coverage rate. When the volume of source code is very large, the above methods generate a large amount of processing, which affects the testing time. Also, because the above methods require processing of the source code, the above code coverage rate determination methods only apply to the testing stage of a program and cannot be implemented during the actual execution of the program.

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] Various embodiments of the invention are disclosed in the following detailed description and the accompanying drawings.

[0008] FIG. 1 is a schematic diagram illustrating an embodiment of a code coverage detection system;

[0009] FIG. 2A is a schematic diagram illustrating an embodiment of a number of log file output points determination system;

[0010] FIG. 2B is a schematic diagram illustrating another embodiment of a number of log file output points determination system;

[0011] FIG. 2C is a schematic diagram illustrating yet another embodiment of a number of log file output points determination system;

[0012] FIG. 3 is a schematic diagram illustrating another embodiment of a code coverage rate determination system;

[0013] FIG. 4 is a flow chart illustrating an embodiment of a code coverage rate determination method;

[0014] FIG. 5A is a schematic diagram illustrating an embodiment of a number of log file output points determination method;

[0015] FIG. 5B is a schematic diagram illustrating another embodiment of a number of log file output points determination method;

[0016] FIG. 5C is a schematic diagram illustrating yet another embodiment of a number of log file output points determination method;

[0017] FIG. 6 is a flow chart illustrating another embodiment of the code coverage rate determination method;

[0018] FIG. 7 is a flow chart illustrating an embodiment of a code coverage detection method; and

[0019] FIG. 8 is a flowchart illustrating another embodiment of a code coverage detection method.

DETAILED DESCRIPTION

[0020] The invention can be implemented in numerous ways, including as a process; an apparatus; a system; a composition of matter; a computer program product embodied on a computer readable storage medium; and/or a processor, such as a processor configured to execute instructions stored on and/or provided by a memory coupled to the processor. In this specification, these implementations, or any other form that the invention may take, may be referred to as techniques. In general, the order of the steps of disclosed processes may be altered within the scope of the invention. Unless stated otherwise, a component such as a processor or a memory described as being configured to perform a task may be implemented as a general component that is temporarily configured to perform the task at a given time or a specific component that is manufactured to perform the task. As used herein, the term 'processor' refers to one or more devices, circuits, and/or processing cores configured to process data, such as computer program instructions.

[0021] A detailed description of one or more embodiments of the invention is provided below along with accompanying figures that illustrate the principles of the invention. The invention is described in connection with such embodiments, but the invention is not limited to any embodiment. The scope of the invention is limited only by the claims and the invention encompasses numerous alternatives, modifications and equivalents. Numerous specific details are

set forth in the following description in order to provide a thorough understanding of the invention. These details are provided for the purpose of example and the invention may be practiced according to the claims without some or all of these specific details. For the purpose of clarity, technical material that is known in the technical fields related to the invention has not been described in detail so that the invention is not unnecessarily obscured.

[0022] FIG. 1 is a schematic diagram illustrating an embodiment of a code coverage detection system. The system 100 includes a log file output point information retrieval unit 110, a log file output point determination unit 120, an actual log file outputted retrieval unit 130, a log file output point quantization unit 140, a code coverage rate determination unit 150, and a code execution unit 160.

[0023] The log file output point information retrieval unit 110 is configured to retrieve source code of a program.

[0024] The log file output point determination unit 120 is configured to determine theoretical number and location information of log file output points included in the source code. Details of the determination are described below in connection with FIG. 4.

[0025] The actual log file outputted retrieval unit 130 is configured to retrieve log files actually outputted during the execution of the program.

[0026] The log file output point quantization unit 140 is configured to determine a number of log file output points of the actual outputted log files based on the location information of the corresponding log file output points recorded in each log file. In one example, the log file output point quantization unit 140 optionally includes a reading element or a receiving element. The reading element is configured to read the log files outputted by the program from a predetermined storage location. The storage location can be predetermined by the program. The receiving element is configured to receive log files transmitted by the program from a predetermined interface. The predetermined interface can be predetermined by the program.

[0027] Programmers often use log files to record information about programs' operation. As used herein, a log file output point refers to a point in the program where a log is outputted. The log file output point can indicate a time, an error level (e.g., fatal, error, warning, info, debug, etc.), and log information related to a corresponding source code's position (e.g., file, line number, etc.). For example, a log that is output at a log file output point can include information relating to the

program's operations such as information regarding the operating environment, and can have a log level such as a warning.

[0028] As a result, a log program in an online system can automatically open and store for a period of time the output of the program, and can obtain log file information in various runtime environments easier. In some embodiments, the period of time can be set to several days. The code coverage rate determination unit 150 is configured to obtain the code coverage rate of the program based on the theoretical number of log file output points included in the source code and the number of log file output points of the actual outputted log files.

[0029] The code execution unit 160 is configured to compile and execute the source code.

[0030] The number of log file output points of the actual outputted log files may be determined using, without limitation, the following systems:

[0031] FIG. 2A is a schematic diagram illustrating an embodiment of a log file output points determination system configured to determine the number of log file output points. The system 200 is an implementation of the log file output point quantization unit 140. The system 200 includes a log file analysis unit 210, a merging unit 220, a merged log file counting unit 230, and an input/output unit 235.

[0032] The log file analysis unit 210 is configured to analyze content of the actual outputted log files to obtain the location information of the corresponding log file output points.

[0033] The merging unit 220 is configured to merge log files including the same location information into one merged log file.

[0034] The merged log file counting unit 230 is configured to count the total number of the log file output points of the actual outputted log files based on the total number of merged log files.

[0035] The input/output unit 235 is configured to input and output log files.

[0036] FIG. 2B is a schematic diagram illustrating another embodiment of log file output points determination system configured to determine the number of log file output points. The system 240 is an implementation of the log file output point quantization unit 140. The system

240 includes: an actual outputted log file counting unit 250, a subtraction unit 260, and an output unit 265.

[0037] The actual outputted log file counting unit 250 is configured to count the actual number of log files outputted for each log file output point based on the location information of the corresponding each log file output point. The reason that more than one log file can be output is that the log file can form another log file after a certain size has been reached. To obtain the output with respect to a specific log file output point, all the log files outputted within a certain time can be gathered.

[0038] The subtraction unit 260 is configured to in the event that the actual number of log files outputted for each log file output point is greater than 1, subtract a value from the actual number of log files outputted to obtain the number of repetitions. The value can be 1.

[0039] The output unit 265 is configured to output a final number of log file output points of the actual outputted log files based on the number of repetitions of log files outputted.

[0040] To obtain the final number, the number of repetitions of log files is subtracted from the actual number of log files outputted to obtain a final number of log files, and the final number of log files is assigned to the final number of log file output points.

[0041] FIG. 2C is a schematic diagram illustrating yet another embodiment of a number of log file output points determination system. The system 270 is an implementation of the log file output point quantization unit 140. The system 270 comprises: an establishment unit 272, a selection unit 274, a retrieval unit 276, a location determination unit 278, an addition unit 280 and a quantity determination unit 282.

[0042] The establishment unit 272 is configured to establish a quantity record list. The initial value of the quantity record list can be 0.

[0043] The selection unit 274 is configured to select an actual outputted log file.

[0044] The retrieval unit 276 is configured to retrieve the location information of the corresponding log file output point of the selected actual outputted log file.

[0045] The location determination unit 278 is configured to determine whether the location information is included in the quantity record list.

[0046] In the event that the location information is not included in the quantity record list, The addition unit 280 is configured to add the location information to the quantity record list and increment a quantity in the quantity record list by 1. The quantity determination unit 282 is configured to perform a determining operation.

[0047] The determining operation comprises determining whether any actual outputted log files remain, in the event that the actual outputted log files remain, the selection unit 274 selects another actual outputted log file, and in the event that the actual outputted log files do not remain, the quantity included in the quantity record list is assigned to be the quantity of log file output points of the actual outputted log files.

[0048] In the event that the location information is included in the quantity record list, the quantity determination unit 282 is configured to perform the above determining operation.

[0049] FIG. 3 is a schematic diagram illustrating another embodiment of a code coverage rate determination system. The system 300 includes a log file output point determination unit 310, a comparison unit 320, an uncovered code determination unit 330, and an input/output unit 340.

[0050] The log file output point determination unit 310 is configured to determine log file output points of actual outputted log files in a source code based on location information in the source code of the corresponding log file output points recorded in the actual outputted log files.

[0051] The comparison unit 320 is configured to determine log file output points of un-outputted log files based on the log file output points of the actual outputted log files and theoretical log file output points. In some embodiments, the comparison unit 320 is configured to compare the actual outputted log files in different runtime environments to determine the log file output points of log files which are outputted in some runtime environments but not outputted in other runtime environments.

[0052] For example, output from two different runtime environments can be stored in a file. The file can indicate various output points and corresponding log level for the two different runtime environments. For each output point, the file can indicate whether the two different runtime environments output a log. In some embodiments, the file can indicate whether the log was output or not for each output point for each environment.

[0053] The uncovered code determination unit 330 is configured to determine code that is not covered or uncovered code based on the log file output points of the un-outputted log files.

[0054] The input/output unit 340 is configured to input and output log files.

[0055] FIG. 4 is a flow chart illustrating an embodiment of a code coverage rate determination method. The method 400 can be implemented on, for example, the system 100. The code coverage rate determination method 400 includes:

[0056] In step 410, a system is configured to retrieve source code of a program.

[0057] In step 420, the system is configured to analyze theoretical number and location information of log file output points included in the source code.

[0058] Generally, in order to more easily discover and resolve actual problems during subsequent program maintenance and inspection processes, the log file output points are set up in the source code to record runtime situations of the corresponding lines of code. One log file output point usually corresponds to a number of lines of code. In the event that the lines of code are executed normally, a corresponding log file will be outputted. The log file output points are generally marked using a predetermined format. The predetermined format of the log file output points allows a scanner to analyze the source code of the program to obtain the log file output points, the location information of the log file output points in the source code, and corresponding relationships to other log file content (such as, e.g., log write records, log levels, etc.). The information obtained via the analysis facilitates the obtaining of the theoretical number of the log file output points. The theoretical number of the log file output points relates to the total number of the log file output points of outputted log files assuming that the source code of the program is executed normally.

[0059] A conventional scanning process includes scanning the entire source code with respect to the log program characteristics to determine log information included in the source code as a whole. A method such as ctag can also be used in place of a full text scan of the source code to determine log information.

[0060] In step 430, the system is configured to retrieve the log files actually outputted during the execution of the program.

[0061] In step 440, the system is configured to determine an actual number of log file output points of the actually outputted log files based on the location information of the corresponding log file output points recorded in each log file.

[0062] In the event that code configured to output a log file location is executed, a corresponding log file is outputted. By counting the log files outputted during the execution of the program, and analyzing the location information of the corresponding log file output points recorded therein, the system can count the number of log file output points in the actual outputted log files. Accordingly, it is possible to count the number of lines of code that has been executed.

[0063] The log files actually outputted during the execution of the program are generally stored in a predetermined path, and the log files can be read directly from the storage location. The path of the storage location of the outputted log files can be changed during the execution of the program or the actually outputted log files can be retrieved from a predetermined interface. In some embodiments, in order to avoid affecting the normal execution of the program, the frequency with which the program outputs log files to the code coverage rate determination system can be set. For example, the frequency with which the program outputs log files can be set to output on a half-hourly basis, on an hourly basis, or for a one-time output after the program run is complete.

[0064] In step 450, the system is configured to obtain a code coverage rate for the program based on the theoretical number of log file output points included in the source code and the number of log file output points of the actual outputted log files.

[0065] The ratio of the theoretical number of log file output points included in the source code and the number of log file output points of the actual outputted log files can correspond to the code coverage rate. Because the code coverage rate does not involve the actual number of code, the code coverage rate obtained may not correspond to the absolute numerical equivalent of the actual code coverage. As another aspect, the determining of the code coverage rate does not merely correspond to the obtaining of an absolute numerical value. Instead, one of the benefits of the code coverage rate is to aid in identifying problems that could potentially arise in the source code.

[0066] The result obtained using the code coverage rate determination method 400 does not necessarily correspond to an accurate numerical value of code coverage. The result

expressed by the log files outputted therein is the certain number of lines of run code that correspond to the log file output point. Thus, the overall runtime situation of each portion of the code included in the source code is reflected accurately. For example, if no log file is outputted for a log file output point that is expected to be outputted during an actual run, this scenario indicates that a problem has occurred in the source code in this location. Even if a portion of a certain number of lines of code corresponding to the log file output point were run, the lines do not need to be considered when determining the code coverage rate, because whether the lines are considered has no effect on the identification of problems in the source code.

[0067] Determination of the log file output points of the actual outputted log files can occur in a variety of ways. In the event that each log file output point only outputs one log file, the total number of actual outputted log files can correspond to the number of log file output points of the actual outputted log files. However, in actual operation, the same log file output point can be output to a plurality of log files. In order to ensure accuracy of the total number of log file output points, the repeated log files can be removed, and after removal, the number of log files is assigned to the total number of log file output points. In another example, the removal of the repeated log files can be implemented using at least the following methods:

[0068] FIG. 5A is a schematic diagram illustrating an embodiment of a number of log file output points determination method. The method 500 includes:

[0069] In step 510, the system is configured to analyze content of the actual outputted log files to obtain the location information of the corresponding log file output points recorded in the source code.

[0070] In step 520, the system is configured to merge log files including the same location information into one merged log file.

[0071] In step 530, the system is configured to determine the total number of the log file output points of the actually outputted log files by counting the total number of merged log files. The total number of merged log files corresponds to the number of log file output points of the actual outputted log files.

[0072] FIG. 5B is a schematic diagram illustrating another embodiment of a number of log file output points determination method. The method 540 includes:

[0073] In step 550, the system is configured to count the actual number of log files outputted for each log file output point based on location information of the corresponding each log file output point recorded in the log files.

[0074] In step 555, in the event that the number of log files is greater than 1, the system is configured to subtract a value (e.g., 1) from the actual number of log files outputted for each log file output point to obtain the number of repetitions outputted for each log file output point.

[0075] In step 560, the system is configured to output a final number of log file output points of the actual outputted log files based on the number of repetitions of log files outputted. The total number of actual outputted log files minus the number of repetitions corresponds to the final number of log file output points of the actual outputted log files.

[0076] For example, if the total number of actually outputted log files is 200, and the actual number of log files outputted at a certain location is 10, then 10 minus 1 corresponds to the number of repetitions of 9 where each location relates to a specific log output point. Furthermore, each log file can include its location in the code so that repetitious log files can be determined with respect to a location. The number $200 - 9 = 191$ is used to correspond to the number of log file output points of the actual outputted log files. In case other locations also have repetitions, the same process is repeated until all repetitions have been subtracted to obtain the final number, i.e., the number of log file output points of the actual outputted log files.

[0077] FIG. 5C is a schematic diagram illustrating yet another embodiment of a number of log file output points determination method. The method 570 comprises:

[0078] In step 572, the system is configured to establish a quantity record list. The initial value of the quantity in the quantity record list can be 0.

[0079] In step 574, the system is configured to select an actual outputted log file.

[0080] In step 576, the system is configured to retrieve location information of the corresponding log file output point of the selected actual outputted log file.

[0081] In step 578, the system is configured to determine whether the location information is included in the quantity record list.

[0082] In step 580, in the event that the location information is not included in the quantity record list, the system is configured to add the location information to the quantity record list and increase the quantity by a value.

[0083] In step 582, in the event that the location information is included in the quantity record list, the system is configured to determine whether any actual outputted log files remain.

[0084] In step 584, in the event that actual outputted log files remain, the system is configured to proceed to step 574 to select another actual outputted log file.

[0085] In step 584, in the event that an actual outputted log file does not remain, the system is configured to assign the quantity included in the quantity record list to be the number of log file output points of the actual outputted log files.

[0086] One of ordinary skill in the art understands that other methods can also be used, as long as the repeated portions therein are subtracted.

[0087] FIG.6 is a flow chart illustrating another embodiment of the code coverage rate determination method. The method 600 can be implemented on, for example, the system 100. Because one program can include a plurality of functional units, during the writing of the source code, the code corresponding to each functional unit can be placed in a different folder to facilitate searching and differentiation. Therefore, the method 600 comprises:

[0088] In step 610, a device is configured to retrieve source code of a program.

[0089] In step 620, the device is configured to analyze theoretical number of log file output points of the source code included in different locations (e.g., file folders).

[0090] In step 630, the device is configured to retrieve log files actually outputted during the execution of the program.

[0091] In step 640, the device is configured to determine the number of log file output points of the log files actually outputted in each folder during the execution of the program.

[0092] In step 650, the device is configured to obtain code coverage rates for each folder based on the theoretical number of log file output points included in each folder and the number of log file output points of the actual outputted log files.

[0093] By analyzing the code coverage rates for each folder, determining where problems lie in a program's source code can be performed easier. For example, if the code coverage rate of a particular folder is determined to be 100%, and the code coverage rate of another folder is determined to be 30%, a determination can be made that problems may be present in the source code related to the latter folder. Thus, subsequent maintenance and revision of the source code can be facilitated.

[0094] In some embodiments, when the method 600 is used to determine the code coverage rate of the source code as a whole, the quantities determined in steps 410-450 for each different location (e.g. file folders) can be summed directly so that the method 600 can be implemented using the same method 400.

[0095] Additionally, the log files outputted during the execution of the program can include the following information: time, level (e.g., Fatal/Error/Warning/Info/Debug, etc.) in which file and at which line of the source code the log file output point is outputted, specific log contents, etc. By analyzing information included in the log files, a log file output situation during a particular execution of the program can be accurately obtained, and the relationship thereof to specific code in the source code can be accurately determined. For example, a log file outputted during a program run is as shown below:

```
[2011-09-21 11:44:57.009137] (time)
[WARNING]
[18316]
[build/debug64/component/sub_path/file.cpp:1181]
__TARGET_ID__:sys/component
PROFILING:Tubo start the fuxi work item:1316576697
```

[0096] As shown in the above example, the time of this log file is "2011-09-21 11:44:57.009137", the level is "WARNING", and the location in the source code of the corresponding log file output point is "build/debug64/component/sub_path/file.cpp:1181". Therefore, the method 600 can also implement code coverage detection based on the information included in the log files to determine the extent of code coverage of a particular program in a runtime environment.

[0097] FIG. 7 is a flow chart illustrating an embodiment of a code coverage detection method. The method 700 can be implemented on, for example, system 300. The method 700 comprises:

[0098] In step 710, a device is configured to determine log file output points of actual outputted log files in a source code of a program based on location information in the source code of the corresponding log file output points recorded in the actual outputted log files.

[0099] In step 720, the device is configured to determine log file output points of un-outputted log files based on the log file output points of the actual outputted log files and theoretical log file output points.

[0100] In step 730, the device is configured to determine uncovered code based on the log file output points of the un-outputted log files.

[0101] Because the location information in the source code of the corresponding log file output points is recorded in the log files, the device can determine locations of the log file output points that correspond to the run code of the source code by comparing the log file output points against the theoretical log file output points. Thus, the log file output points of the un-outputted log files can be determined, and thereby the code which is not covered can also be determined.

[0102] In one example, a program can be tested in different runtime environments to simulate the program's actual runtime environment as much as possible and to test the program from a plurality of angles. For example, unit testing, integration testing, production environments, etc. correspond to different runtime environments. In different runtime environments, code coverage rates may not be entirely the same. In other words, the actual outputted log files may not be the same in the different runtime environments. FIG. 8 is a flowchart illustrating another embodiment of a code coverage detection method. The method 800 can be implemented on, for example, the system 300. The method 800 can be used to compare code coverage in different runtime environments, and comprises:

[0103] In step 810, a device is configured to retrieve log files actually outputted by a program in different runtime environments.

[0104] In step 820, the device is configured to determine log file output points of the log files which are output in one or more runtime environments but not output in other runtime environments based on location information in the source code of the corresponding log file output points recorded in the log files actually outputted in each runtime environment.

[00105] In step 830, the device is configured to determine code which is not covered in the corresponding runtime environments based on the log file output points of un-outputted log files.

[00106] In the event that source code location of a certain log file output point is not recorded in log files of all runtime environments, the source code location of the certain log file output point not being recorded in the log files indicates the log files were not outputted for the log file output point corresponding to the source code location in certain environments. Thus, a determination can be made that uncovered code exists in the environment corresponding to the log file that is not recorded. For example, three different runtime environments of a program include a unit testing environment, an integration testing environment, and a production environment. By comparing the records included in the log files of the three runtime environments, in the event that corresponding log files are discovered for the unit testing environment and the integration testing environment and no log files is discovered for the production environment, this scenario indicates that some of the code covered in the unit testing environment and the integration testing environment is not covered in the production environment. Therefore, by comparing the information recorded in the log files, uncovered code can be found quickly, thereby facilitating subsequent revision and maintenance of the program.

[00107] The units described above can be implemented as software components executing on one or more general purpose processors, as hardware such as programmable logic devices and/or Application Specific Integrated Circuits designed to perform certain functions or a combination thereof. In some embodiments, the units can be embodied by a form of software products which can be stored in a nonvolatile storage medium (such as optical disk, flash storage device, mobile hard disk, etc.), including a number of instructions for making a computer device (such as personal computers, servers, network equipment, etc.) implement the methods described in the embodiments of the present invention. The units may be implemented on a single device or distributed across multiple devices. The functions of the units may be merged into one another or further split into multiple sub-units.

[00108] Although the foregoing embodiments have been described in some detail for purposes of clarity of understanding, the invention is not limited to the details provided. There are many alternative ways of implementing the invention. The disclosed embodiments are illustrative and not restrictive.

[00109] WHAT IS CLAIMED IS:

CLAIMS

1. A code coverage rate determination method, comprising:
 - retrieving source code of a program;
 - determining a theoretical number of log file output points included in the source code and
 - 5 location information of the log file output points included in the source code;
 - retrieving log files actually outputted during the execution of the program;
 - determining, using one or more processors, an actual number of log file output points of the actual outputted log files based on the location information of the corresponding log file output points recorded in each log file; and
 - 10 determining, using the one or more processors, a code coverage rate of the program based on the theoretical number of log file output points included in the source code and the actual number of the log file output points of the actual outputted log files.
2. The code coverage rate determination method as described in claim 1, wherein the determining of the number of the log file output points comprises:
 - 15 analyzing content of the actual outputted log files to obtain the location information of the corresponding log file output points;
 - merging log files containing the same location information into one merged log file; and
 - counting the total number of the log file output points of the actual outputted log files based on the total number of merged log files.
- 20 3. The code coverage rate determination method as described in claim 1, wherein the determining of the number of the log file output points comprises:
 - counting an actual number of log files outputted for each log file output point based on the location information of the corresponding each log file output point;
 - in the event that the actual number of log files outputted for each log file output point is
 - 25 greater than 1, subtracting a value from the actual number of log files outputted for each log file output point to obtain the number of repetitions outputted by each log file output point; and
 - outputting a final number of log file output points of the actual outputted log files based on the number of repetitions of log files outputted.
4. The code coverage rate determination method as described in claim 1, wherein the
- 30 determining of the number of log file output points comprises:
 - S1: establishing a quantity record list;
 - S2: selecting an actual outputted log file;

S3: retrieving the location information of the corresponding log file output points of the selected actual outputted log file;

S4: determining whether the location information is included in the quantity record list;

S5: in the event that the location information is not included in the quantity record list:

5 adding the location information to the quantity record list;

increasing a quantity of the quantity record list; and

proceeding to step S7;

S6: in the event that the location information is included in the quantity record list, proceeding to step S7; and

10 S7: determining whether an actual outputted log file remains;

S8: in the event that the actual outputted log file remains, proceeding to step S2;

S9: in the event that the actual outputted log file does not remain, assigning the quantity included in the quantity record list to be the number of log file output points of the actual outputted log files.

15 5. The code coverage rate determination method as described in claim 1, wherein the retrieving of the log files actually outputted during the execution of the program comprises reading the actually outputted log files from a predetermined storage location or retrieving the actually outputted log files transmitted by the program from a predetermined interface.

6. The code coverage rate determination method as described in claim 1, wherein the
20 determining of the theoretical number of the log file output points included in the source code comprises:

analyzing the theoretical number of log file output points of the source code included in different folders, wherein the number of log file output points of log files actually outputted during the execution of the program corresponds to the number of log file output points of actual
25 outputted log files included in each folder.

7. A code coverage rate determination system, comprising:

at least one processor configured to:

retrieve source code of a program;

determine theoretical number of log file output points included in the source code

30 and location information of the log file output points included in the source code;

retrieve log files actually outputted during the execution of the program;

determine an actual number of log file output points of actual outputted log files based on the location information of the corresponding log file output points recorded in each log file; and

determine a code coverage rate of the program based on the theoretical number of log file output points included in the source code and the actual number of the log file output points of the actual outputted log files; and

a memory coupled to the at least one processor and configured to provide the at least one processor with instructions.

8. The code coverage rate determination system as described in claim 7, wherein the retrieve of the log files actually outputted during the execution of the program comprises read the actually outputted log files from a predetermined storage location or retrieve the actually outputted log files transmitted by the program from a predetermined interface.

9. A code coverage detection method, comprising:

determining, using one or more processors, log file output points of actual outputted log files included in source code based on location information in the source code of the corresponding log file output points recorded in the actual outputted log files;

determining, using the one or more processors, log file output points of un-outputted log files based on the log file output points of the actual outputted log files and theoretical log file output points; and

determining, using the one or more processors, uncovered code based on the log file output points of the un-outputted log files.

10. The code coverage detection method, comprising:

retrieving log files actually outputted by a program in different runtime environments;

determining, using one or more processors, log file output points which are outputted in one runtime environment and not in another runtime environment based on location information in source code of corresponding log file output points recorded in the log files actually outputted in each runtime environment; and

determining, using the one or more processors, uncovered code when the program is executed in the corresponding runtime environments, the determination being based on the log file output points of the un-outputted log files.

11. A code coverage detection system, comprising:

at least one processor configured to:

retrieve log files actually outputted by a program in different runtime environments;
determine log file output points which are outputted in one runtime environment and
not in another runtime environment based on location information in source code of
corresponding log file output points recorded in the log files actually outputted in each
runtime environment; and

determine uncovered code when the program is executed in the corresponding
runtime environments, the determination being based on the log file output points of the un-
outputted log files; and

a memory coupled to the at least one processor and configured to provide the at least one
processor with instructions.

12. A computer program product for determining a code coverage rate, the computer program
product being embodied in a non-transitory computer readable storage medium and comprising
computer instructions for:

retrieving source code of a program;

determining theoretical number of log file output points included in the source code and
location information of the log file output points included in the source code;

retrieving log files actually outputted during the execution of the program;

determining an actual number of log file output points of the actual outputted log files based
on the location information of the corresponding log file output points recorded in each log file; and

determining a code coverage rate of the program based on the theoretical number of log file
output points included in the source code and the actual number of the log file output points of the
actual outputted log files.

13. The computer program product as described in claim 12, wherein the determining of the
number of the log file output points comprises:

analyzing content of the actual outputted log files to obtain the location information of
the corresponding log file output points;

merging log files containing the same location information into one merged log file; and

counting the total number of the log file output points of the actual outputted log files based
on the total number of merged log files.

14. The computer program product as described in claim 12, wherein the determining of the
number of the log file output points comprises:

counting an actual number of log files outputted for each log file output point based on the
location information of the corresponding each log file output point;

in the event that the actual number of log files outputted for each log file output point is greater than 1, subtracting a value from the actual number of log files outputted for each log file output point to obtain the number of repetitions outputted by each log file output point; and

outputting a final number of log file output points of the actual outputted log files based on the number of repetitions of log files outputted.

15. The computer program product as described in claim 12, wherein the determining of the number of log file output points comprises:

S1: establishing a quantity record list;

S2: selecting an actual outputted log file:

S3: retrieving the location information of the corresponding log file output points of the selected actual outputted log file;

S4: determining whether the location information is included in the quantity record list;

S5: in the event that the location information is not included in the quantity record list:

adding the location information to the quantity record list;

increasing a quantity of the quantity record list; and

proceeding to step S7;

S6: in the event that the location information is included in the quantity record list, proceeding to step S7; and

S7: determining whether an actual outputted log file remains;

S8: in the event that the actual outputted log file remains, proceeding to step S2;

S9: in the event that the actual outputted log file does not remain, assigning the quantity included in the quantity record list to be the number of log file output points of the actual outputted log files.

16. The computer program product as described in claim 12, wherein the retrieving of the log files actually outputted during the execution of the program comprises reading the actually outputted log files from a predetermined storage location or retrieving the actually output log files transmitted by the program from a predetermined interface.

17. The computer program product as described in claim 12, wherein the analyzing of the theoretical number of the log file output points included in the source code comprises:

analyzing the theoretical number of log file output points of the source code included in different folders, wherein the number of log file output points of log files actually outputted during the running of the program corresponds to the number of log file output points of actual outputted log files included in each folder.

18. A computer program product for determining a code coverage rate, the computer program product being embodied in a non-transitory computer readable storage medium and comprising computer instructions for:

determining log file output points of actual outputted log files included in source code
5 based on location information in the source code of the corresponding log file output points recorded in the actual outputted log files;

determining log file output points of un-outputted log files based on the log file output points of the actual outputted log files and theoretical log file output points; and

determining uncovered code based on the log file output points of the un-outputted log files.

10 19. A computer program product for determining a code coverage rate, the computer program product being embodied in a non-transitory computer readable storage medium and comprising computer instructions for:

retrieving log files actually outputted by a program in different runtime environments;

determining log file output points which are outputted in one runtime environment and not
15 in another runtime environment based on location information in source code of corresponding log file output points recorded in the log files actually outputted in each runtime environment; and

determining uncovered code when the program is executed in the corresponding runtime environments, the determination being based on the log file output points of the un-outputted log files.

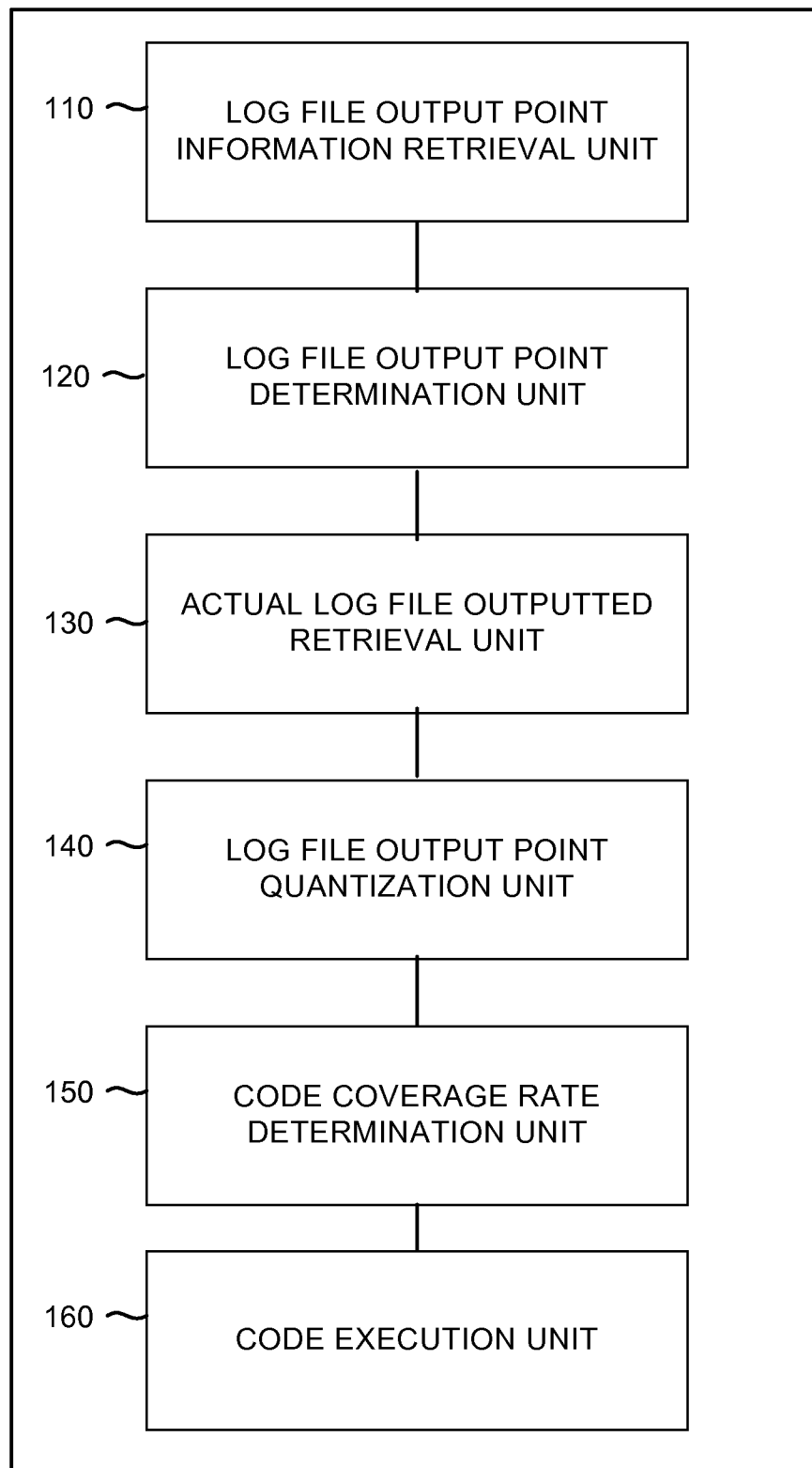
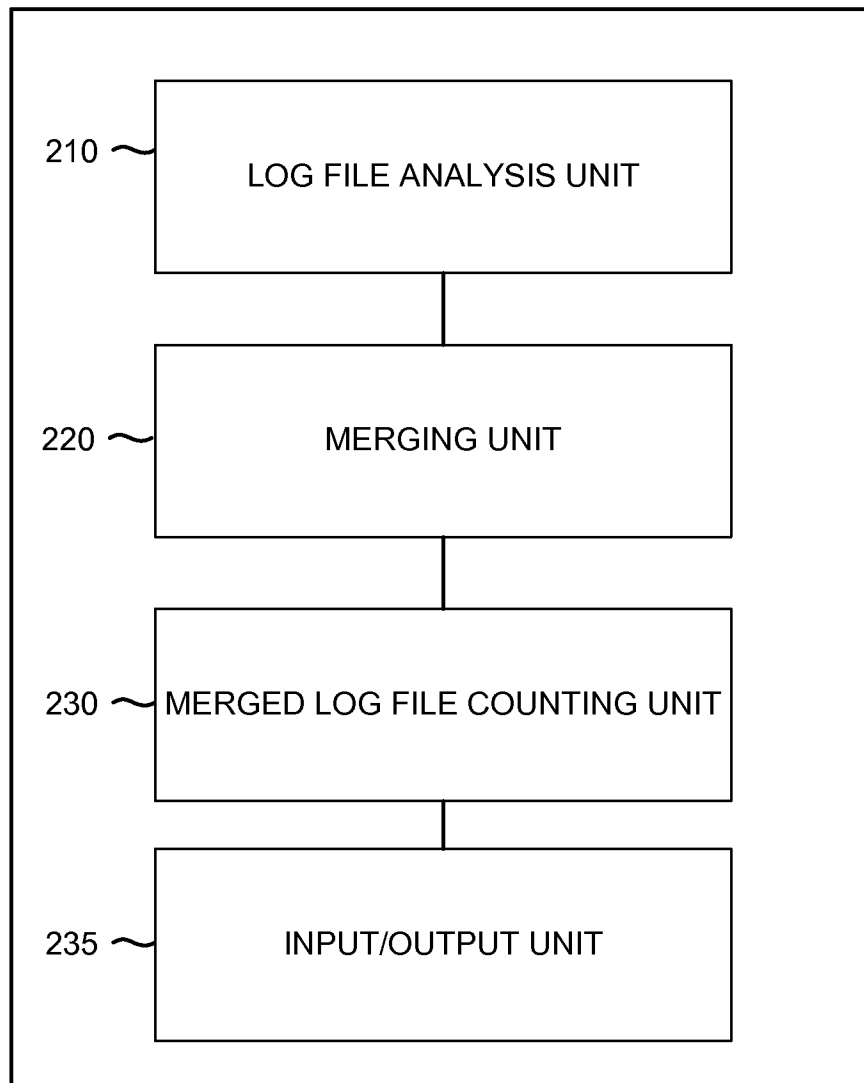
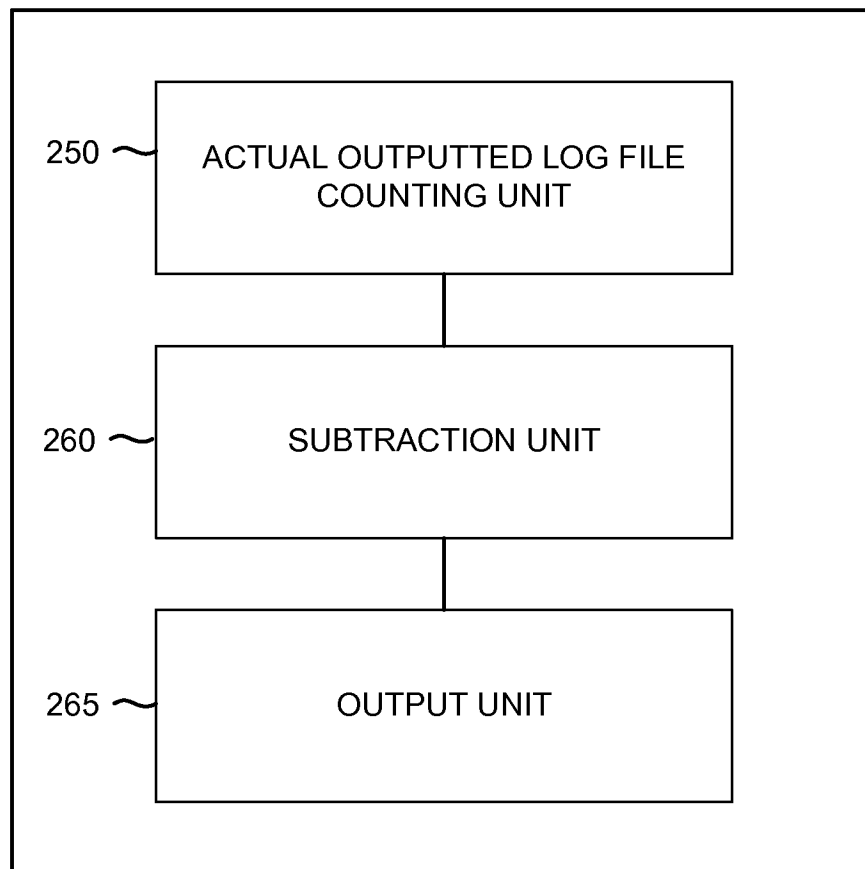
100

FIG. 1



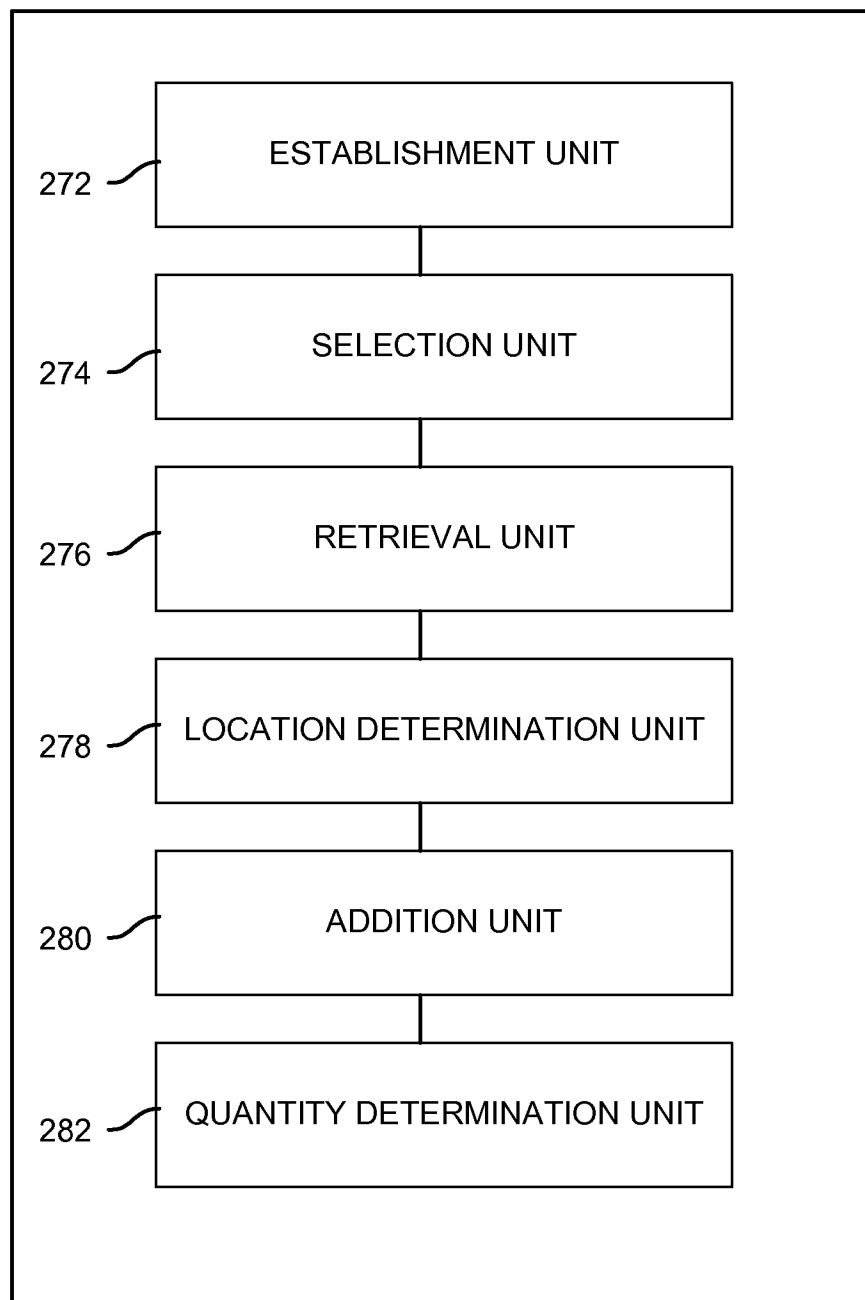
200

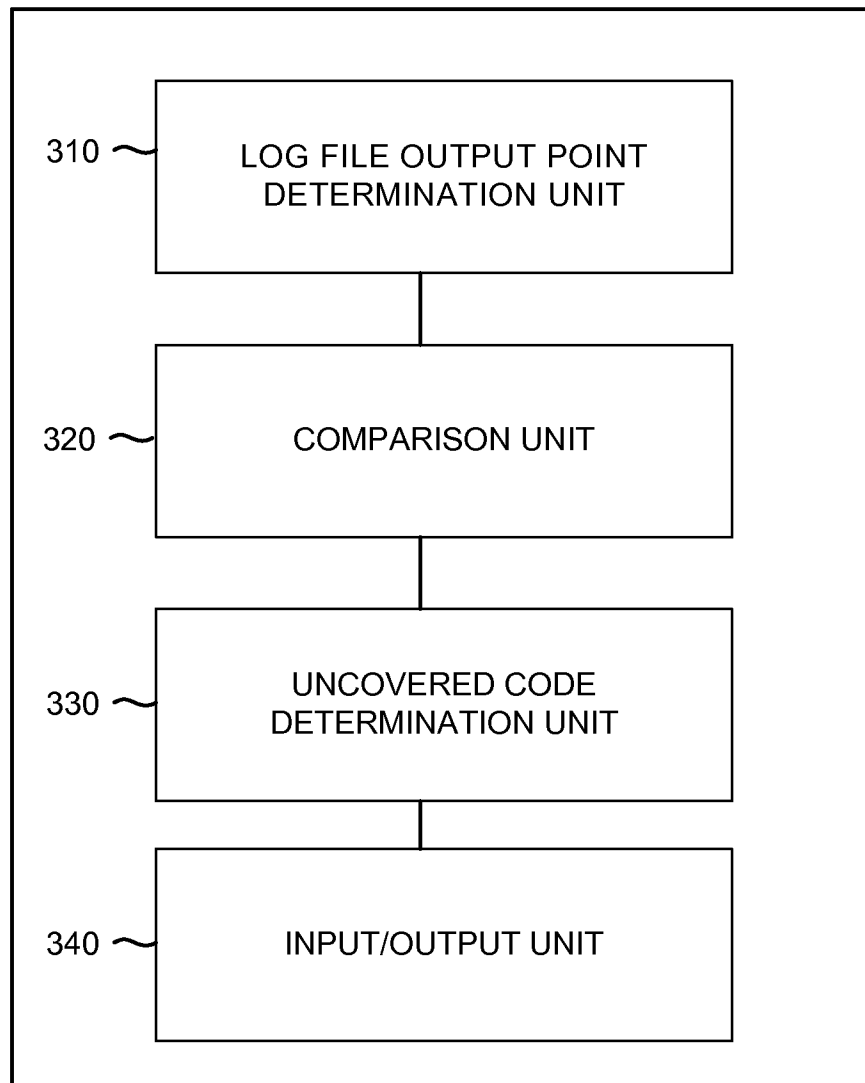
FIG. 2A



240

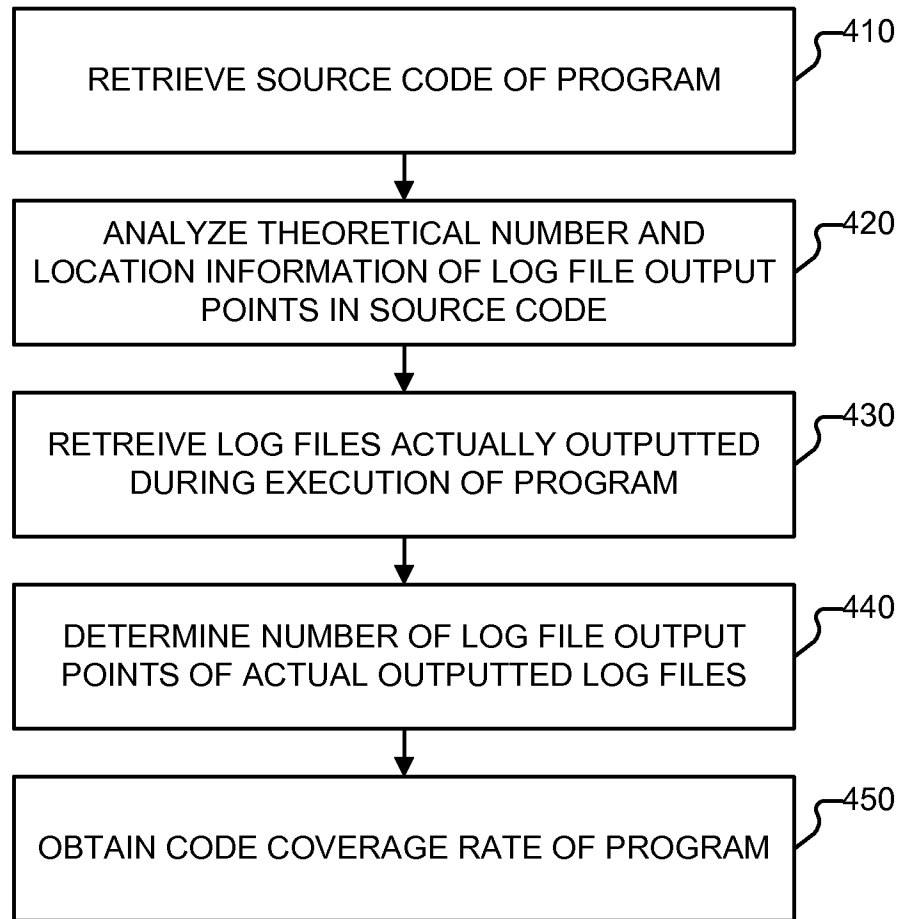
FIG. 2B

**270****FIG. 2C**



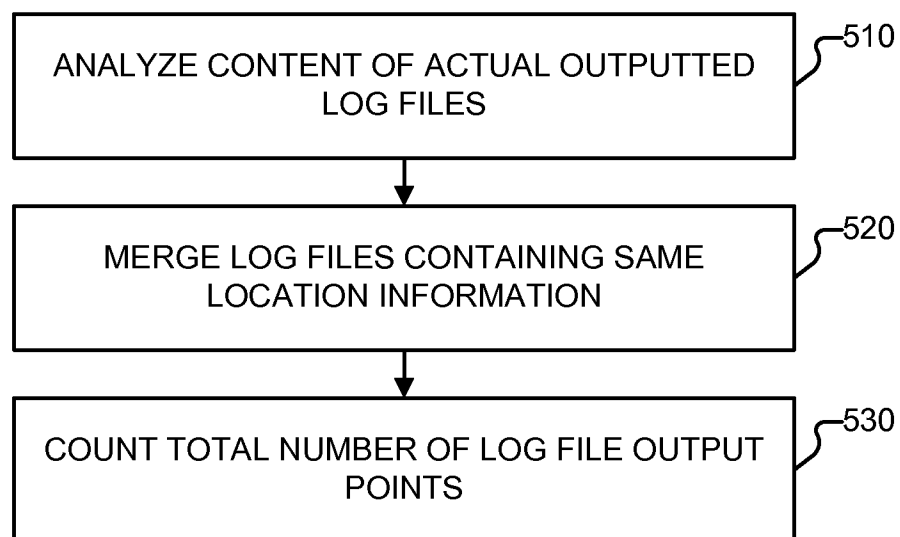
300

FIG. 3



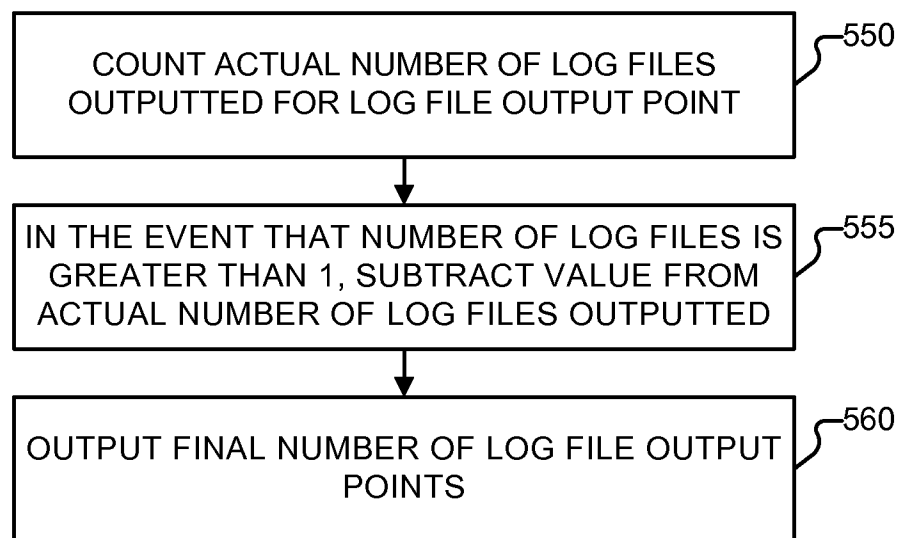
400

FIG. 4



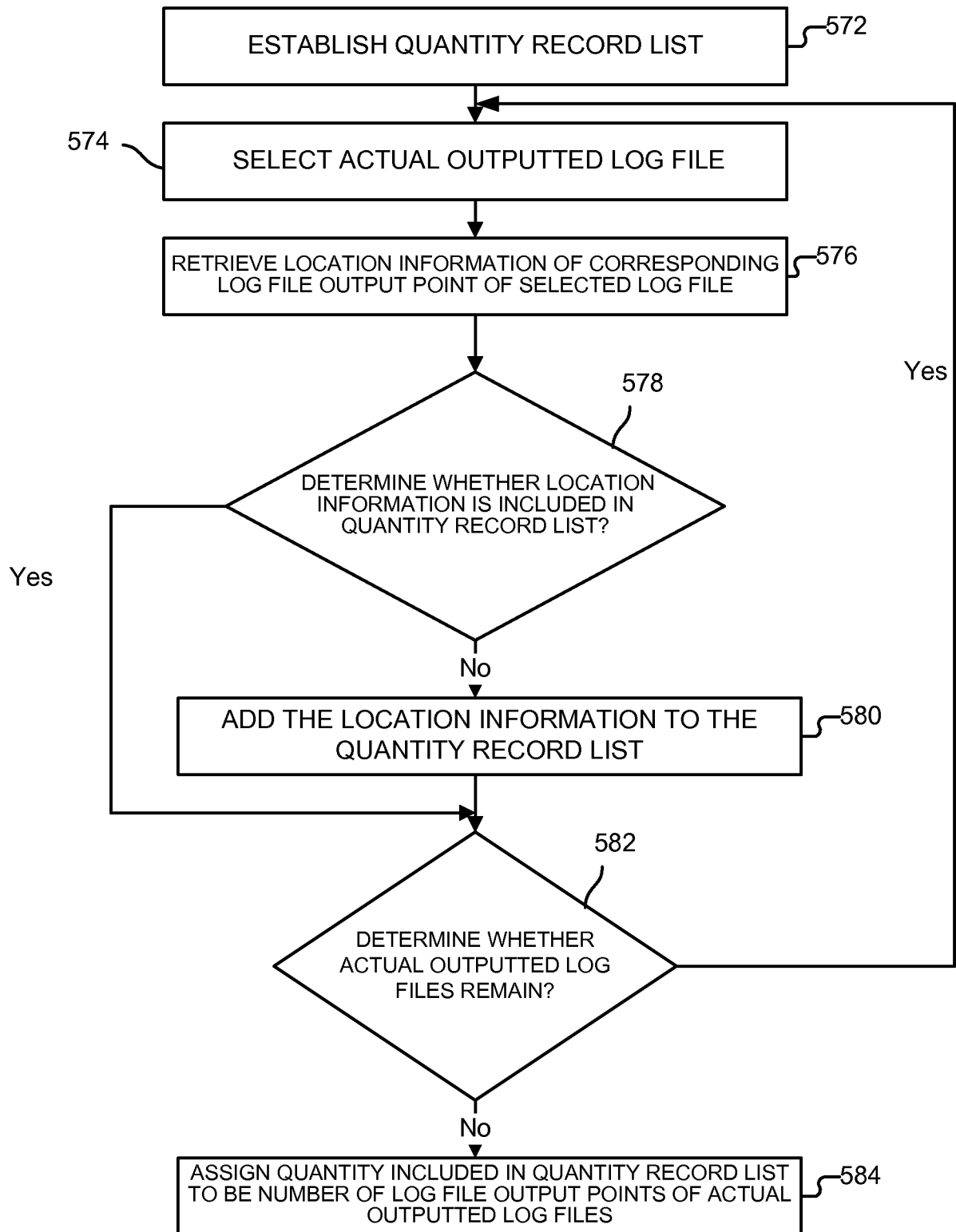
500

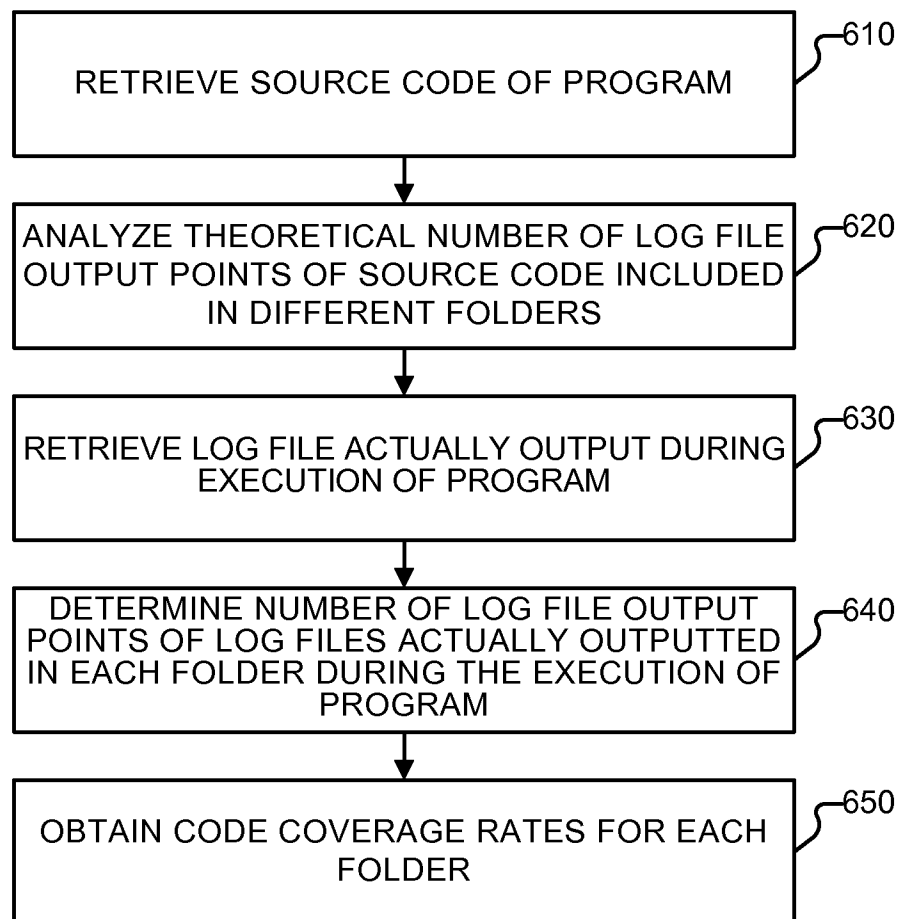
FIG. 5A

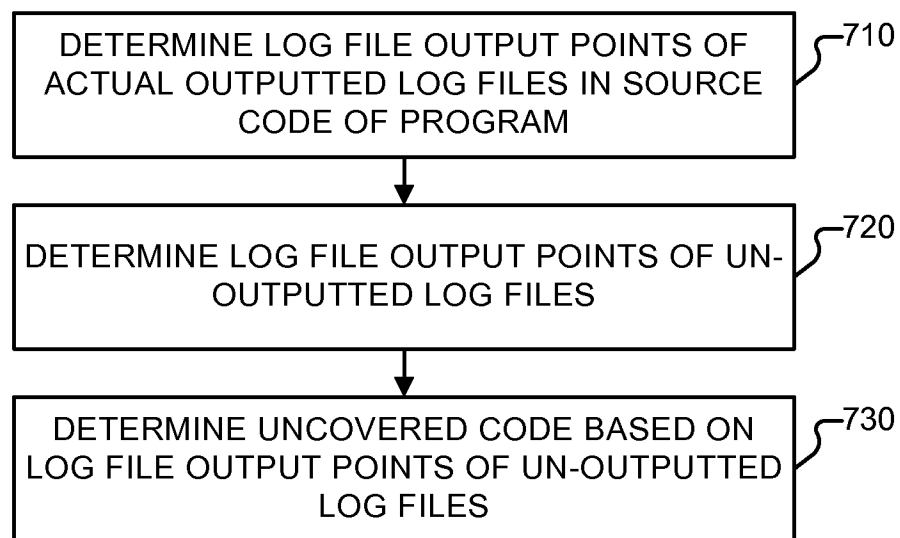


540

FIG. 5B

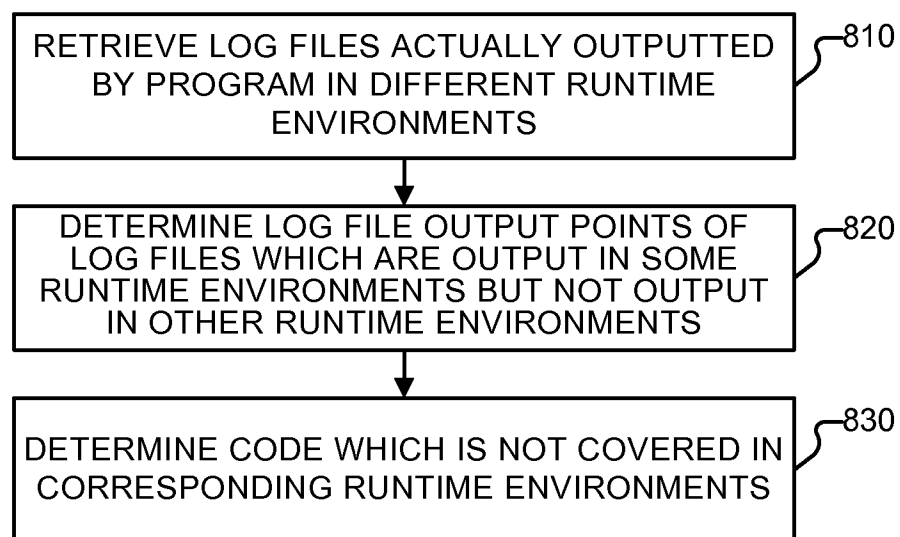
**570****FIG. 5C**

**600****FIG. 6**



700

FIG. 7



800

FIG. 8

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2013/025585

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06F11/36
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data, IBM-TDB

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 7 617 415 B1 (KADAKIA NIRAVE D [US]) 10 November 2009 (2009-11-10) the whole document	1-19
X	----- MECHELLE GITTENS, KERI ROMANUFA, DAVID GODWIN, JASON RACICOT: "All Code Coverage Is Not Created Equal: A Case Study in Prioritized Code Coverage", ACM, 2 PENN PLAZA, SUITE 701 - NEW YORK USA, 2006, XP040053799, the whole document ----- -/--	1-19



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

8 May 2013

Date of mailing of the international search report

23/05/2013

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Salsa, Francesco

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2013/025585

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	HORGAN J R ET AL: "Achieving software quality with testing coverage measures", COMPUTER, IEEE SERVICE CENTER, LOS ALAMITOS, CA, US, vol. 27, no. 9, 1 September 1994 (1994-09-01), pages 60-69, XP011420675, ISSN: 0018-9162, DOI: 10.1109/2.312032 the whole document -----	1-19

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2013/025585

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 7617415	B1	10-11-2009	NONE
