Program-Delta Oriented Debugging Supporting Method DMET

Makoto Matsushita, Masayoshi Teraguti, and Katsuro Inoue

1Graduate School of Engineering Science, Osaka University, Toyonaka, 560-8531 Japan
2Tokyo Research Laboratory, IBM Japan, Ltd., Yamato, 242-8502 Japan

SUMMARY

Research has been conducted on methods of automatically performing tests that focus on the delta (difference) between a revision that is known to function normally in advance and a revision that contains defects to identify the causes of the defects. However, since software had to be created from source programs for each test, an extremely long test time had been required. Also, since the emphasis had only been on the testing tasks and no consideration was given to debugging tasks, these methods were not practical. Therefore, in the current research, the authors propose the DMET debugging technique, which can be used for actual software maintenance. DMET, which reduces the test execution time compared with conventional techniques, supports the entire sequential processing flow from testing to debugging. Also, to verify the effectiveness of the proposed technique, the authors implemented the DSUS development support environment, which uses DMET, and performed comparison experiments. From the results, it was apparent that using DMET enabled the total debugging time to be shortened. © 2006 Wiley Periodicals, Inc. Syst Comp Jpn, 37(11): 35-43, 2006; Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/scj.20351

Key words: revision management; delta; debugging.

1. Introduction

Software maintenance is processing that is performed to enable active software to continue operating or to improve that software [3]. The proportion of the total cost of software devoted to maintenance has risen to approximately 80% [4], and approximately 65% of the total time spent by people involved with software is devoted to maintenance or tasks related to it [2].

To understand the essence of software maintenance, Swanson classified the main causes for which software maintenance is required into three basic types [15]:

- Defects that cause errors within the software
- Changes to the environment surrounding the software
- Requests by users or personnel responsible for maintenance

In addition, Swanson defined the maintenance activities that are executed corresponding to these basic causes as follows:

- Corrective maintenance: Identifies and corrects errors
- Adaptive maintenance: Makes corrections according to changes in the environment
- Perfective maintenance: Improves performance and alters or adds functions

Generally, among these software maintenance activities, the ones that tend to be considered to occur most often are activities for corrective maintenance. However, when
Lientz and Swanson conducted surveys in the 1980s for the first time, they reported that corrective maintenance comprises no more than 20% of maintenance tasks, and preventive maintenance accounts for 55% [11].

During these maintenance activities, many changes are made to existing software. However, research by Hetzel showed that the probability that errors will be introduced when software is modified was from 50% to 80% [7]. Therefore, during corrective maintenance and preventive maintenance, which comprise 75% of maintenance activities, operation must be verified not only for functions in the parts that were modified, but also for functions in parts that were not modified.

Regression testing [6, 10, 13] has been widely used for these operation verification activities that occur during development. Also, the software that is developed is often managed by using a software management system such as a revision management system [1, 16]. Attention given to these two points has led to previous research on means of supporting preventive maintenance of software that was being managed by a software management system [12, 17]. However, the test tools [9] that were to be used when executing tests had to be established and testing had to be performed manually in certain circumstances. Also, support for the debugging tasks that are performed after testing was insufficient.

Therefore, the current research targets preventive maintenance tasks that are performed to modify or extend functions. During these kinds of maintenance tasks, parts other than the functions for which the tasks were performed may be affected, and defects may be detected by the tests. In this paper, we propose the DMET debugging technique, which uses regression testing, to efficiently eliminate these kinds of defects.

This technique first chooses product revisions sequentially and uses a test tool to automatically perform testing. When a revision that outputs defects and a revision that outputs normally are detected as a result of testing, the corrections that were performed between these revisions are assumed to contain an error. Next, a check is performed to determine which parts in the current revision correspond to these correction contents (hereafter, referred to simply as the delta). The user uses these results to perform debugging and correct the defects. By also reflecting those changes in the old revision after the defects were corrected, they can be used for subsequent tests.

To verify the effectiveness of DMET, we created a trial debugging support system named DSUS, which uses DMET, and performed comparison experiments using DSUS. From the experimental results, it was apparent that DMET was able to support the sequential processing flow from testing to debugging and that it could shorten the time required for debugging tasks during actual maintenance.

## 2. Software Maintenance

### 2.1. Maintenance tasks

The tasks that are performed to modify software during maintenance activities basically can be classified into the following three types.

1. Understanding the software and the modifications that must be made
2. Modifying the software
3. Verifying the operation of the modified software

Of these, if the type (1) tasks are performed correctly, they are useful for investigating the causes of defects that are discovered. Conversely, if the software is poorly understood, a great deal of time will be required to investigate the causes of defects and correct them. As a result, one technique for improving software understandability and maintainability is to perform software configuration management or revision management.

Software configuration management refers to tasks for identifying, controlling, and understanding the states of various products that are created during the software development and maintenance processes [5]. Also, revision management consists of tasks for correctly recognizing, organizing, and managing the various corrections made to the products (software or accompanying documents) that were created by the development team. Models of various management techniques have been proposed. In addition, revision management systems based on these models have been implemented [1, 16]. Normally, products are revised multiple times, and the product that is created for each alteration is called a revision.

When performing type (3) tasks, one technique for testing whether or not the software operates with the performance required by the specifications after the software was modified is to perform regression testing [6, 10, 13]. Regression testing, which has been put to practical use during actual software maintenance, is used for examining whether corrections were performed properly.

### 2.2. Problems with existing techniques

During the maintenance stage for software that was developed by using a revision management system, there exists a revision (hereafter, referred to as the base revision) that operates with the performance required by the specifications. With conventional software maintenance, the program of the base revision is changed and the operation of the software is verified by using regression testing during the types (2) and (3) maintenance tasks described above.
If a defect is discovered during testing, the types (1) and (2) maintenance tasks are performed repeatedly to eliminate the error that is the cause of that defect (to perform debugging). However, if the defect is found in a function that has not been changed, it is generally difficult to find the error that caused it.

Therefore, research has been conducted [12, 17] to identify the error that is the cause of the defect by automatically performing testing using the corrections that were made from the base revision until the current revision. However, a great deal of time was required for testing because each time a test was performed, software for applying the test was created by modifying and compiling the base revision. Education and training related to the test tool [9] that was used for testing were also required since each time a defect was discovered, test cases had to be sequentially created corresponding to that defect. In addition, since only the testing task for identifying the cause of the defect was emphasized and no consideration was given to the subsequent tasks up to debugging, the technique could not be directly applied to actual maintenance tasks.

3. DMET Debugging Technique

This section describes our proposed DMET debugging technique, which uses regression testing and program delta information that is maintained by a revision management system.

3.1. Overview

DMET, which derives its name from Debugging METhod, is a technique for improving the efficiency of debugging tasks when all functions work normally in the base revision but defects are created and incorporated in some functions during maintenance tasks. DMET eliminates defects by repeatedly applying a testing technique, display technique, and reflection technique.

The testing technique automatically performs tests by using a test tool to identify the causes of defects between revisions. DMET uses regression testing to verify whether or not each revision contains defects. If a defect is detected during regression testing, a test (hereafter, referred to as a localization test) is performed to verify which revision that defect was created in and incorporated.

The display technique displays the delta with the revision of the executable program that was identified by the testing technique on the latest revision of the source program. Since source program correction tasks are generally performed on the latest revision, this technique automatically obtains the modifications that were made from the identified revision until the latest revision.

The reflection technique reflects the modifications that were made during the correction tasks in the latest revision while going back through the revisions to the revision that was identified by the testing technique. In this paper, we will refer to the collection of revisions in which defects occurred during tests and for which corrective modifications must be reflected in the latest revision as reflection-required revisions. By applying the corrections to the reflection-required revisions, subsequent testing tasks can continue more efficiently.

In the following sections, we will first describe the software development environment that is predicated for the proposed technique and then explain specific procedures for the testing technique, display technique, and reflection techniques.

3.2. Prerequisite environment for applying DMET

This section describes prerequisite conditions for the software, its development and maintenance, and the personnel responsible for software maintenance when our proposed debugging technique is used during software maintenance. These types of prerequisites are widely used such as in the development of open source software, which has reached a certain degree of maturity. Therefore, we believe that these prerequisite conditions can be applied without problem in the software development environments that have been used in recent years.

- Development is performed by using a revision management system

We assume that development tasks are performed using a revision management system to understand the work history during development in terms of relatively detailed units. The revision management targets are software source programs and compiled files. When compiled files are registered as revisions, relationships between the relevant files and source program revisions are also registered as collateral information.

- A base revision exists

The target of this technique is software maintenance tasks. Therefore, we assume that by letting the software at the time that the maintenance tasks started be the base revision, a revision always exists for which all functions are guaranteed to operate before modifications are made.

- Some output is produced in response to input

Since this technique uses a testing tool to judge whether the output is correct for the entered test data, the target of this technique is software that produces output. We assume here that the output is generally some visible result such as the output of a character string to a terminal, for example.

- The input and usage environment are not changed
Since large-scale specification changes are not generally made during software maintenance activities, we assume that the input to the software does not change. In other words, the input parameters themselves do not increase or decrease. For similar reasons, the proposed technique also is based on the assumption that the software usage environment does not change. In other words, the hardware or operating environment does not change.

In this kind of software development environment, a developer generally registers development results in the revision management system in terms of small work units. When a single developer is carrying out small-scale development, the regression testing can also be executed each time a revision is registered. However, in general, it is often the case that multiple developers perform individual tasks while cooperating or that a single developer simultaneously performs multiple tasks in parallel. Also, development results in which corrections are incomplete are also often registered to record the intermediate progress of a collection of tasks. Since it is meaningless under these circumstances for regression testing to be executed each time the developer registers a revision, regression testing is used to verify whether the corrections were made properly at the time that a certain collection of tasks has ended.

3.3. Testing technique

The testing technique automatically performs regression testing for executing all prepared tests each time any change is made to the software.

Let $T_i$ to $T_n$ denote the set of tests used for performing perfective maintenance, which were prepared for the software to be tested. These tests are prepared in advance when the maintenance tasks begin. Also, when $V_i$ denotes the $i$-th revision for a certain file, let $V_B$ represent the base revision and $V_L$ ($B \leq L$) represent the latest revision. Now, consider the case when test $V_i$ ($1 \leq i \leq n$) is applied to the software at a certain time. First, for each file, let the normal output of this software be denoted by output result $O_B$, when $I_h$ which is the input of $T_h$ is assigned to $V_B$, which is the revision that is known in advance to operate correctly.

The test technique first applies all $T_i$ ($1 \leq i \leq n$) to $V_l$ to verify whether the corresponding output $O_{Li}$ matches $O_{B_i}$. If a defect is detected when a certain test $T_j$ is applied, the localization test $LclT(i, L - 1, L)$ is performed.

The localization test $LclT(i, j, k)$ judges the result when $T_j$ is applied to $V_j$ ($j \leq k$) and returns the two revisions between which the cause of the defect exists (Fig. 1). However, we know here that when $T_j$ is used to perform the test for $V_k$, its output $O_{Li}$ is the same as $O_{Li'}$. Figure 2 shows the algorithm for the localization test $LclT(i, j, k)$.

(1) If $j$ matches the base revision $B$, decide that the "defect exists between $V_B$ and $V_k$" and end.
(2) Apply test $T_j$ to $V_j$. Specifically, assign $I_h$ to $V_j$ to obtain output $O_{ji}$.
(3) If $O_{ji}$ matches $O_{Bi}$, since a defect already exists in revision $j$, decide that the test execution result is "X" (the same defect is included). In addition, execute $LclT(i, j - 1, j)$ recursively and let that result be the result that is obtained.
(4) If $O_{ji}$ matches $O_{Bi}$, since no defect had been included at the time of revision $j$, decide that the test execution result is "O" (no defect is included). Also, decide that the "defect exists between $V_j$ and $V_k$" and end.

![Localization test](image)

Fig. 1. Localization test.

![Localization test algorithm](image)

Fig. 2. Localization test algorithm.
(5) If some other $O_j$ is obtained, decide that the test execution result is "-" (a new defect was detected). Also, if no execution result was obtained because the software was abnormally terminated, set "x" (no result was obtained). In either case, execute $Let(T(i, j - 1), j)$ recursively and let that result be the result that is obtained.

This technique identifies the location that contains the defect by linearly searching each revision sequentially starting from a new location. If, for example, the results are only "O" and "x" and a transition from "O" to "x" occurs only once, then efficient searching can be done by using a binary search. However, if another output is included as a prerequisite or if multiple transitions from "O" to "x" or vice versa occur for some reason, then a binary search cannot be used here.

3.4. Display technique

To display the delta with the revision that was detected by the testing technique [this is assumed here to be between revision $C$ and revision $E$ ($C < E \leq L$)] in the latest revision $L$, first the software and source program relationship information is used to check which revisions of which source program correspond to the delta that is highlighted. The highlight is displayed as follows according to whether code was deleted or inserted between the two revisions.

[Deletion] Nothing appears in the source program of the current revision, and the pre-modified source program cannot be highlighted (Fig. 3). As a result, this situation is highlighted in the latest revision by a mark indicating the location and contents of the deleted source code (dotted line in $V_L$ in Fig. 3).

[Insertion] The part that appears in the source program of the current revision is highlighted (see Fig. 4). Although some modifications are also made after the insertion, the part that is the fatal cause of the defect is considered to have not been changed.

3.5. Reflection technique

If a new defect is detected by the localization test, that defect is first eliminated for the latest revision $L$ (hereafter, the delta in which this is done will be referred to as the correction $\Delta L$). Next, to continue the localization tests, $\Delta L$ is reflected in past revisions. First, a patch program is used to apply $\Delta L$ to past revisions, and the new software that should be created is compiled. If compilation is successful, the result is used as a new program to be tested. If the $\Delta L$ patch fails or if the $\Delta L$ patch succeeds and compilation fails, that revision is excluded from subsequent regression testing.

4. Evaluation Experiments

This section describes DMET evaluation experiments using the DSUS debugging support system that is based on DMET.

4.1. Trial system DSUS

To construct a debugging support system based on DMET, we paid careful attention to the following points when creating the system.

- Independence from a programming language

  Generally, constructing a system that depends on a programming language enables more detailed debugging support to be provided. However, since many programming languages are currently used for program development and since DMET itself is a technique that is not dependent on a language, we decided that DSUS should not have functions that are dependent on a specific programming language.

- Support for both testing and error correction

  DSUS provides an environment for correcting errors that were detected by tests, not just for automatically executing tests. As a result, development tasks can be completed within DSUS.

- Automatic execution of tools

  Sequentially executing each of the tasks that were described for DMET beginning with test execution can impose an even greater load on the developer. With DSUS, we aim to reduce the load incurred by introducing DMET by automatically performing the DMET procedure and the tests that accompany it.
The document contains a description of a software development process involving the creation of an interface for a system. The process includes steps for defining and collecting an experimental platform, developing and collecting data, and testing. The text is accompanied by a diagram illustrating the steps and components involved. The goal is to create a functional interface that meets specific requirements.
functions due to this extension are the software modules that are to be debugged in step 2. Table 1 shows a list of the software modules that were collected and their development history.

For example, software A has a total of 28 revisions where the software that was provided first was the first version. Localization tests were executed 25 times according to DMET for software A, and as a result, it was apparent that the deltas between revisions that were detected by DMET contained defects.

4.4. Step 2

Groups $G_1$ and $G_2$ were each given three software modules that contain defects, which were obtained in step 1, and asked to use DSUS to perform debugging tasks. Test data that had been prepared in advance was used as the test data used for debugging, and debugging ended at the time that the test subject verified that the processing could be performed correctly for the assigned test data. The time taken by each test subject from when debugging started until it ended was totaled. Tables 2 and 3 show the debugging times (in minutes) obtained by this experiment for the test subjects (T1 to T10).

4.5. Discussion

When we used a Welch test with a significance level of 5% for the debugging time totals for software A and B, we verified that a significant difference was seen between $G_1$ and $G_2$. It is apparent from this experiment that using DMET can shorten the debugging time.

5. Related Research

This section considers research related to techniques that had been used previously for identifying errors that cause defects and describes problems involved in those existing techniques.

5.1. Research of Ness and Ngo

Ness and Ngo used a technique called regression containment for compiler development at Cray Research [12]. The technique of Ness and Ngo first performs regression testing automatically. If regression testing fails, then using the fact that the base revision operates correctly, testing is repeatedly performed while executing the corrections that were obtained from configuration management in the order that the corrections were actually applied. At the stage when the test failed, the correction that was applied at that time is identified as the error that caused a bug.

However, although the technique of Ness and Ngo operates well in certain situations, it does not operate well when the test first fails because multiple deltas are applied instead of just a single delta or when an inconsistency occurs so that the program cannot be compiled when a change is applied.

5.2. Research of Zeller

Zeller proposed a technique [17] that can also deal with defects due to multiple errors or inconsistencies due to the application of changes, which are problems with the technique of Ness and Ngo. Zeller’s technique considers a correction as one element of a set without taking into consideration the order in which corrections are applied. Therefore, if the number of corrections that were made is $n$, then the number of sets that are considered is $2^n$. Zeller uses an algorithm that finds the set having the minimum
number of elements for which a defect occurs from among the sets that are considered. By using sets of corrections, this algorithm succeeds in identifying the causes of defects that occur due to multiple errors. The algorithm, which can also deal with inconsistencies due to the application of changes, can detect errors that could not be found by using the technique of Ness and Ngo.

However, although making the algorithm more complex increases the precision of identifying the delta that caused the bug, the number of sets that must be tested increases exponentially relative to the number of corrections. Therefore, the number of times tests are performed increases, and an enormous amount of time is required for testing.

5.3. Problems

The techniques of both Ness and Ngo and Zeller consider the source program of the base revision as a baseline, and corrections must be applied to that source code each time and the corrected source code must be compiled to perform testing. Also, since compilation must be performed while accurately managing which corrections were made to which source programs, the task is complex even if only simple compilation is being performed.

To find the cause of a defect, you must create test cases and perform testing yourself, and you must study about or undergo training for the test tool that is being used for each technique. In addition, since these techniques emphasize only testing tasks for identifying the causes of defects and do not take into consideration debugging tasks, it is difficult to apply them to actual maintenance tasks.

DMET enables compilation tasks to be easily performed by using a support environment such as DSUS since information indicating which source programs are to be compiled at which times are managed within the framework of DMET itself. Also, since test cases are first created in DMET and used for debugging, you need not create test cases corresponding to defects yourself, and the time required for creating test cases is reduced. Since DMET is intended not only for identifying defects but also for all tasks that are included up to debugging, it can easily be applied even in actual maintenance tasks, which we showed by the current experiments.

6. Conclusions

In the current research, we proposed the DMET debugging technique, which is intended for maintenance tasks. We also showed the effectiveness of DMET through experiments using the DSUS trial system. Using the proposed technique can be expected to enable debugging tasks to be performed more easily during actual maintenance.

Some future topics of research are as follows. First, we plan to improve the delta detection algorithm so that deltas that contain defects are detected more accurately. When the deltas that are detected by DMET become relatively large, even if the results that are obtained by using DMET are used, they will not contribute to a reduction of the working time. For such cases, we plan to investigate methods of having DSUS judge this situation in advance and alert the developer instead of just simply exhibiting the results.

REFERENCES


**AUTHORS (from left to right)**

Makoto Matsushita graduated in 1993 with a specialty in Information Science from the Department of Engineering Science at Osaka University, took a leave from the second half of his doctoral course in 1998, and became an assistant in the Department of Engineering Science. He has been an assistant in the Graduate School of Engineering Science since 2002 and an associate professor since 2005. He is engaged in research on software development environments, software development processes, and open source development. He holds a D.Eng. degree.

Masayoshi Teraguti graduated in 1993 with a specialty in Information Science from the Department of Engineering Science at Osaka University, completed the first half of his doctoral course in 2000, and joined the Tokyo Research Laboratory, IBM Japan, Ltd. While at Osaka University, he was engaged in research on debugging support environments. He holds an M.E. degree.

Katsuro Inoue (member) graduated in 1979 with a specialty in Information Science from the Department of Engineering Science at Osaka University, completed his doctoral course in 1984, and became an assistant in the Information Division of the Department of Engineering Science there. From 1984 to 1986, he was an assistant professor in the Computer Science Course at the University of Hawaii Manoa Campus. In 1988, he became a lecturer in the Information Science Division of the Department of Engineering Science at Osaka University, an assistant professor in 1991, and a professor in 1995. Since 2002, he has been a professor in the Graduate School of Engineering Science. He holds a D.Eng. degree. He is engaged in research on software engineering.