

Model Checking with Program Slicing Based on Variable Dependence Graphs

Masahiro Matsubara Kohei Sakurai Fumio Narisawa

Hitachi Research Laboratory, Hitachi, Ltd.
Japan

[masahiro.matsubara.td|kohei.sakurai.cp|fumio.narisawa.ks]@hitachi.com

Masushi Enshoiwa Yoshio Yamane Hisamitsu Yamanaka

Hitachi Advanced Digital, Inc.
Japan

Hitachi Automotive Systems, Ltd.
Japan

In embedded control systems, the potential risks of software defects have been increasing because of software complexity which leads to, for example, timing related problems. These defects are rarely found by tests or simulations. To detect such defects, we propose a modeling method which can generate software models for model checking with a program slicing technique based on a variable dependence graph. We have applied the proposed method to one case in automotive control software and demonstrated the effectiveness of the method. Furthermore, we developed a software tool to automate model generation and achieved a 35% decrease in total verification time on model checking.

1 Introduction

In embedded control systems, potential risks to system safety have been increasing because programs are getting larger due to electrification, enhancement of diagnosis, etc. Functional safety standards such as IEC 61508 or ISO 26262 have been established to ensure these systems do not fall into dangerous situations. However, the problem of test coverage still remains. Furthermore, there are corner cases which are difficult to be found with usual tests or simulations.

For example, hardware malfunctions could cause software faults, so it is not sufficient to test the software only. Inspection techniques to test combinations of hardware and software such as HILS (Hardware in-the-Loop Simulation) have been in practical use, but failures caused by hardware malfunctions or timing problems caused by software interruptions are difficult to be detected because there could be large number of test cases.

To solve these problems, model checking is applied. In model checking, all of the state transitions of the system are fully searched to detect corner cases including timing problems. Implementation bugs can also be found if models are made out of the source code. However, the problem in applying model checking is well known: a state explosion that verification does not complete. To avoid a state explosion, the scope of verification has to be limited. However, too limited a model misses some causes of malfunctions, so that the part of the source code to be verified has to be determined appropriately.

We are trying to adopt model checking to the development of automotive control software, using it to find software defects which occur rarely on real systems by means of modeling both the software from the source code and the hardware. To solve the state explosion problem mentioned above, we introduced a program slicing technique [10] based on a variable dependence graph which enables an flexible adjustment of the boundaries between the source code to be modeled as software model in detail and the external environment to the software model. Also we introduced external environment

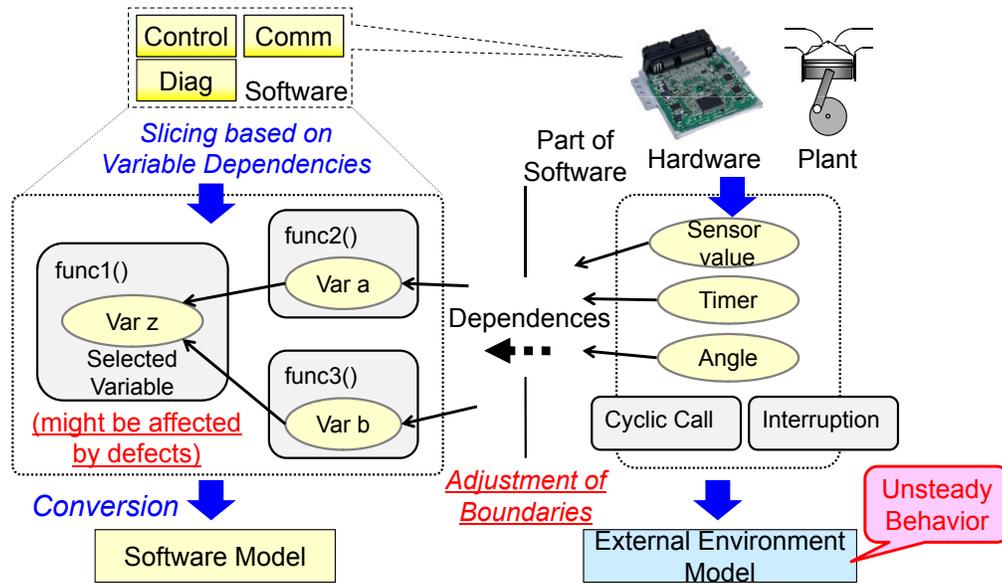


Figure 1: Modeling Method for Embedded Control System

models which include unsteady behavior or faults in them to detect malfunctions caused by interactions of hardware and software.

The contributions of this paper are as follows:

- A practical modeling method to verify a large scale embedded control program. The main point of the method is the slicing technique based on a variable dependence graph.
- The algorithm for analysis of variable dependence graph, which could be adapted to the analysis of system dependence graph.
- An experimental evaluation of the tool which automates the above method.

The rest of this paper is structured as follows: Section 2 presents the modeling method and the slicing technique. Section 3 presents an example case where the method is applied, to show the usefulness of the method. Section 4 presents the tool which automates the method. Section 5 presents the related work. Finally, Section 6 concludes this paper by summarizing its main points.

2 Modeling Method

2.1 System Modeling

The way of modeling an embedded control system is shown in Figure 1. To find defects in implementations, the software model of the target system is converted from the source code which is related to the selected variables by a verifier. Those selections are done in accordance with a verification point. The code to be converted is sliced out from the whole program to avoid a state explosion, using a slicing technique based on a variable dependence graph (see 2.2).

On the other hand, the hardware, controlled instruments (if needed for verification), and a part of the software are modeled as external environments. These external environment models can have unsteady behavior or faults if they are concerned with the content to be verified.

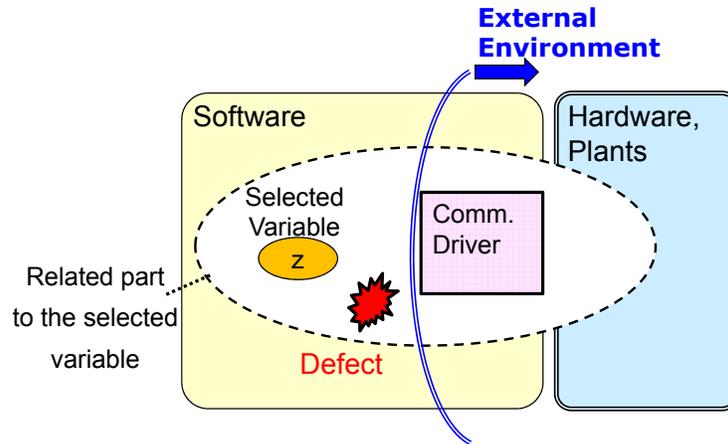


Figure 2: Abstraction of Boundary between Software and Hardware

The slicing technique based on the variable dependence graph makes it easy to adjust the boundaries between the software models and the external environment models. The reason that such adjustments are needed is that the appropriate size of the software model differs depending on the verification points. Too large a model leads to a state explosion, and too small a model can miss the cause of a malfunction.

An example of a boundary adjustment is shown in Figure 2. There is a communication driver between the software and the hardware, and this driver has complex processing. It is better to omit the driver to reduce the state number of the software model, if the driver has no relation to a malfunction. In this case, it is appropriate to treat the driver as a part of the external environment, and simplify its processing. This method could make the modeling highly automated but not fully automated. The use of design knowledge is expected for modeling, so that the scope of the software model has to be changed according to the verification point.

2.2 Slicing Technique

The data to be used for slicing is usually a program dependence graph (PDG) or a system dependence graph (SDG) [7]. The criteria for slicing include a start statement, an end statement, and variables in the end statement.

In this method, the variable dependence graph (VDG) in Figure 3 is used as data for slicing, which is a kind of data flow to connect dependence among the variables. VDG is equivalent to PDG (or SDG) except for the data unit (a statement in PDG), but it is more suitable for expressing dependencies in a tree format. Tree format is better to pursue related variables and code sequentially, and to adjust the boundaries between code converted to the software model and code treated as external environments. In PDG, data dependence and control dependence are used as an edge. On the other hand, in VDG, dependence brought by an assignment which is similar to data dependence is also used as an edge. Nodes in VDG are variables distinct with positions and execution paths in the source code.

There are two directions in VDG trees. In one, the root node is the goal of the dependencies in the tree (we call it the goal tree): in the other, the root node is the start (the start tree). These root nodes are one side of the slicing criteria, and leaf nodes including those which are on a boundary adjusted by a verifier are another side.

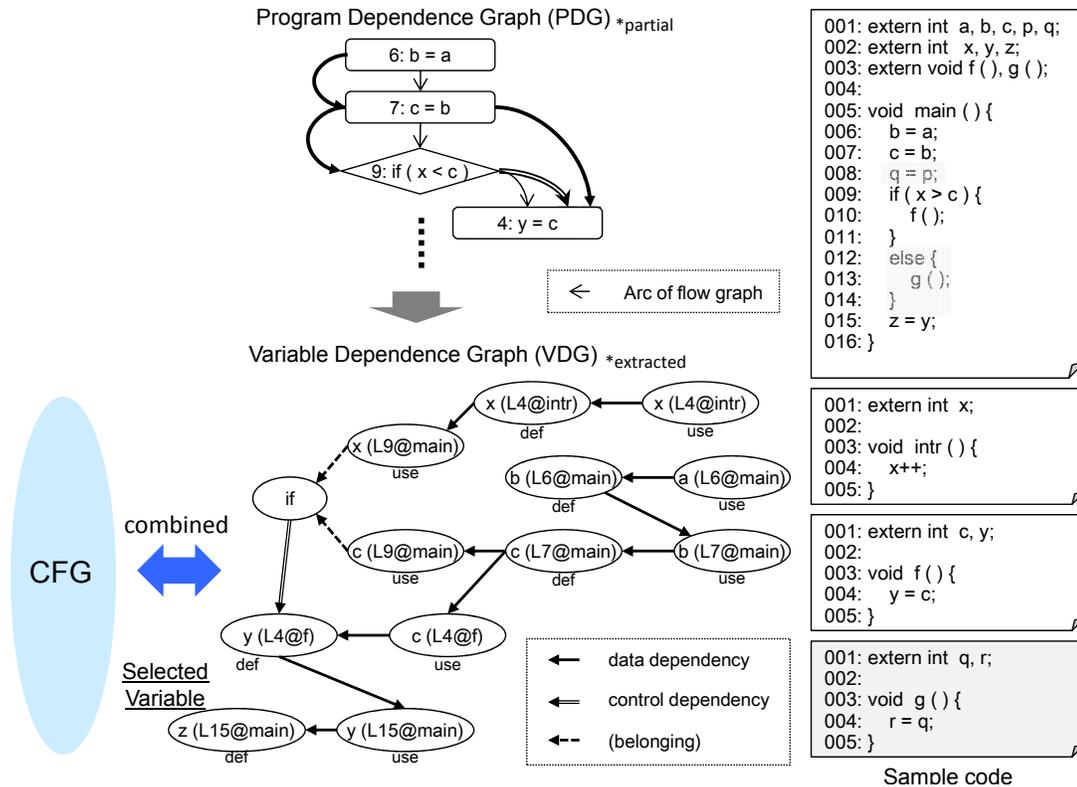


Figure 3: Variable Dependence Graph of Sample Code

The steps of slicing are as follows:

1. Select variables which are related to the verification point in the source code of the target system. These variables are those which could be affected by defects, and are written in properties or assertions.
2. Analyze and extract the VDG goal / start trees from the source code, where the root nodes are the variables selected in step 1.
3. Adjust the boundaries in the VDG trees. (Figure 4)
4. Execute slicing to extract code which has a relation to the variables in the VDG trees extracted in Step 2 and adjusted in Step 3.

In step 3, a variable on a boundary becomes interface between the software model and a external environment model. It is possible to set a boundary between functions by setting the boundary between variables which belong to different functions.

In step 4, code to be extracted are not only statements which include a variable in VDG directly, but also declarations of the variables, functions and control statements which include or relates to the above statements in control flow.

Statements (except function calls) which do not include a variable cannot be extracted, so that these statements have to be picked up additionally on a control flow if needed.

The benefits of this slicing technique are as follows:

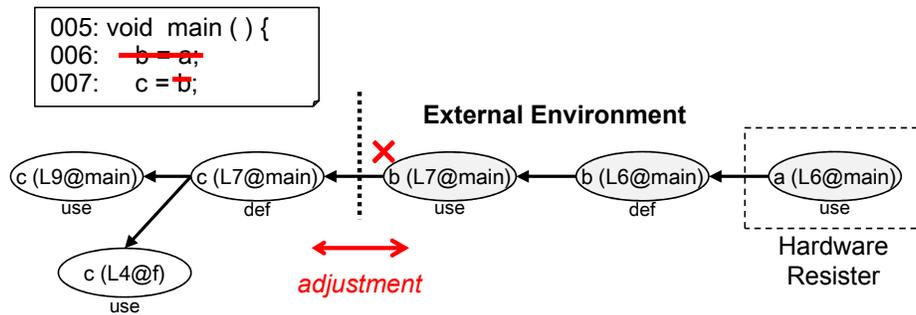


Figure 4: Adjustment Boundary on VDG

- High accuracy of slicing, equal to PDG-based slicing.
- Easy to adjust boundaries between software models and external environments, especially in the tree expression of variable dependencies.
- Interface variables between software models and external environments are clear. Information on the interface is important to make or select external environment models.
- Variables whose range has to be changed are clear when input / output values have data mapping.

The sliced code is converted into software models in the language of the model checker, such as Promela of SPIN [5]. This conversion is done via AST (Abstract Syntax Tree) according to the predetermined rules.

2.3 External Environment Models

External environment models are necessary to detect malfunctions caused by interactions between the hardware and software. Specifically, they set every possible combination of input values to software models, or limit those combinations to reduce the state number according to the specification of the external environment. They also simulate the behavior outside of the software so that it becomes possible to judge whether a property for the external environment is held or not.

External environment models are written manually. When a verifier makes a new model or selects one from the existing models of the external environment, information about the interface variable given in VDG is useful. Usually, input values to a software model or determinants of them are set with randomness. Processes are added to external environment models.

Furthermore, a verifier can put unsteady behavior or faults into the external environment models which occur at random realized with non-determinism. If there is a malfunction when an external environment model has unsteady behavior, and vice versa, that behavior seems to be one of the factors which cause the malfunction, so that a verifier could analyze a counter example from that point of view.

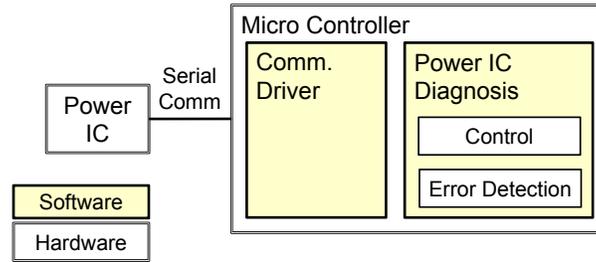


Figure 5: System Structure of Sample Case

3 Case Study

In this section, we apply the modeling method described in section 2 to show its usefulness.

3.1 Example Case

The target to be verified is automotive software for a diagnosis of a power IC in a controller, as shown in Figure 5. The purpose of the verification was to investigate the reason for a malfunction. The malfunction was an incorrect error detection in the diagnosis on a long-term test. The diagnosis reports an error, but no fault was found in the power IC. Standard test methodology was not enough to reproduce the symptom and to find the cause of the malfunction, so model checking was applied.

The verification was done in two steps. In the first step, we focused on the serial communication driver between the micro controller and the power IC, but no defects were found. In the second step, we modeled the software for the diagnosis and the power IC. The communication driver was simplified as an external environment model, because it is proved the driver has no defect in the first step.

The modeling method was applied in the second step. Figure 6 shows the model. For the slicing, the variable which shows the result of the diagnosis was selected. The boundaries were adjusted so that the communication driver and some other part of the software were omitted. One of the inputs given by external environment models is the power IC status to be diagnosed, and another is the controller mode which is decided in the omitted software depending on the power supply. In the transition of the controller mode, a path was added that transits from the power-on state to the power-off state suddenly and abnormally. This path is intended to express the effect of an instantaneous drop of the voltage in the power supply.

In model checking, it is important to reduce the number of concurrent processes in order to reduce the state number. Although the micro controller and power IC run concurrently, the software model and the external model was set in the same one process, and the external environment models change its states before the diagnosis which is a cyclic task. If no malfunction was detected in this model, the external environment models could be separated into a different process. Furthermore, the data mapping was applied to the power IC status to be diagnosed. The sizes of the source code and models are shown in Table 1; logical LOC refers to non-comment lines of code.

The result was that one defect was detected in the software which could cause a malfunction when the controller mode turns to the power-off state suddenly while the diagnosis software is in the specific state. It is confirmed that this malfunction could happen in the actual controller.

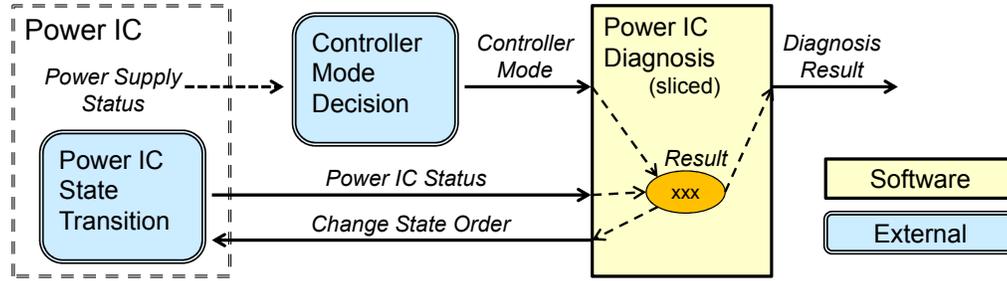


Figure 6: Model of Sample Case

Item	Size (approx.)
Whole source code (.c, .h)	130 k Logical LOC
Software model (Promela)	600 LOC
System model (Promela)	750 LOC

Table 1: Sizes of Source code and Model

The mechanism of the malfunction is as follows. Although the diagnosis is stopped by power off, the power IC continues to change its state while the power is off. After the power supply is restored, an inconsistency between the actual power IC status and the recognition for that status by the diagnosis software occurs. This symptom happens only in a specific situation, so this is a timing problem. It is difficult to detect this defect with usual tests or simulations like HILS because a verifier can hardly set conditions to make this malfunction occur.

This case shows that model checking is effective against timing problems, and the modeling method can treat large-scale control software. An obstacle to applying this modeling method to a development process is that it is time consuming. To improve the efficiency of the method, we developed a software tool to automate the method of generating verification models.

4 Tool

4.1 Functionalities

The tool we developed has two main functions as shown in Figure 7. One is slicing C language code based on a variable dependence graph; another is conversion from C language code to Promela.

When a user indicates one or more functions as execution start points of the program, the tool will analyze the variable dependencies. A variable dependence graph in a tree format will be extracted and displayed according to the variables that the user has selected. The tool can extract variable dependencies which extend to other execution paths, so that the effects of interruptions or other tasks can be considered.

The user can limit the variable dependencies to adjust the boundaries between the software models and external environment models. Such an operation is reflected as a color classification of the source code. In the tool, variable dependence trees, function call trees, and source code files are displayed in association with each other, as shown in Figure 8. When a user selects a variable node on the VDG trees, related function node and code will be highlighted to help the user understand the structure of the program, and to help the user to decide the boundaries.

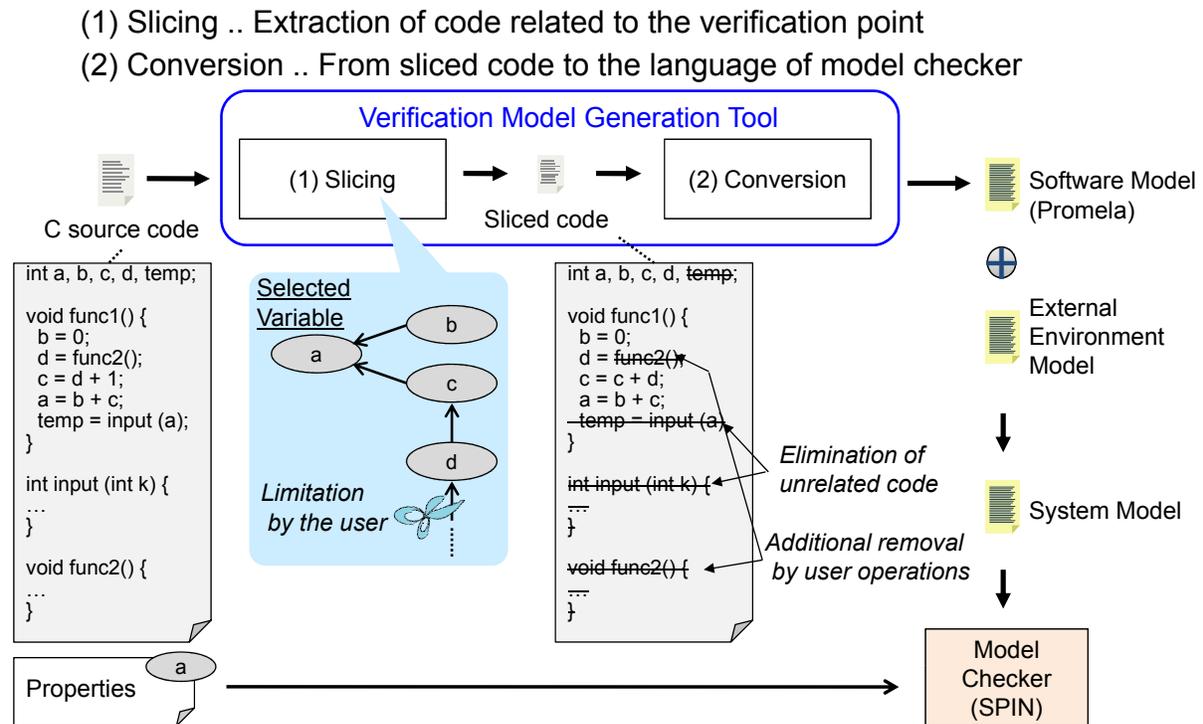


Figure 7: Verification Model Generation Tool

In a tree format, the VDG could become larger than in a network format, so there are some refinements in the projection of the tree. Redundant tree paths (the same partial trees as other parts) are replaced with a node which indicates redundancy. Every node can be gathered to the function node on which the variable is written, so that a boundary adjustment between functions becomes easier. Moreover, the tool has search functions for variables on VDG trees or execution paths to help operations of the user.

At present, the limitations of the tool are as follows. Source code has to be pre-processed before the input. Pointers can be analyzed, but only one variable is pointed by a pointer, and pointed variables given by loop controls are ignored. Function calls by function pointers are also ignored. Slicing and model conversion are separated into different tools, and model conversion is done for each C code file. This tool is not open to the public.

4.2 Algorithm for Analysis of VDG

The algorithm to analyze VDG used in this tool is shown in Figure 9. Dependencies are connected one by one along the control flow graph. Every node, which is a variable distinguished with the position in code and the execution path, is assigned information about the stack trace and associated to the control flow graph (CFG). Whether the nodes connect or not is judged using that information of the dependency source node, the sink node, and the nodes that have been already connected to the sink node. Candidate nodes to connect obtained in A and C can be reduced in consideration of the attributes of the node to be connected, such as the substance (memory area) and the position in the CFG. Dependencies due to loop

statements are analyzed in the similar way. Different from [7] [8], initialization vertexes or finalization vertexes are not used. The advantage of this algorithm we think is scalability. Actually, analysis was successful for code of which the size is over 350k logical LOC.

The process of the analysis for the sample code is shown in Figure 10, where nodes in a conditional statement belong to a “if” node for the simplicity of appearance. There are no function parameters and arguments in the sample code, but they can be treated in the same way.

After the analysis, a VDG for an execution path is obtained. This VDG becomes a network if there is a loop statement. Goal/start trees are extracted from VDG with a selection of the variables, considering shared variables between different execution paths.

4.3 Evaluation

Figure 11 shows the verification process with model checking we employed. Besides the slicing tool and the model converter, the verification assist tool was used in the process. This tool actuates checking on more than one property simultaneously, and shows a variable value table of a counter example in time series to help the user understand the mechanism that the violation of the property happens. If a false positive / false negative error was detected in the verification, the step “Model Revision” is taken additionally to refine the model, and another verification will be done.

Table 2 shows the condition of the experiment. The target is automotive control software, and it had never been verified with model checking. Verifications were done without the tools first, and with the

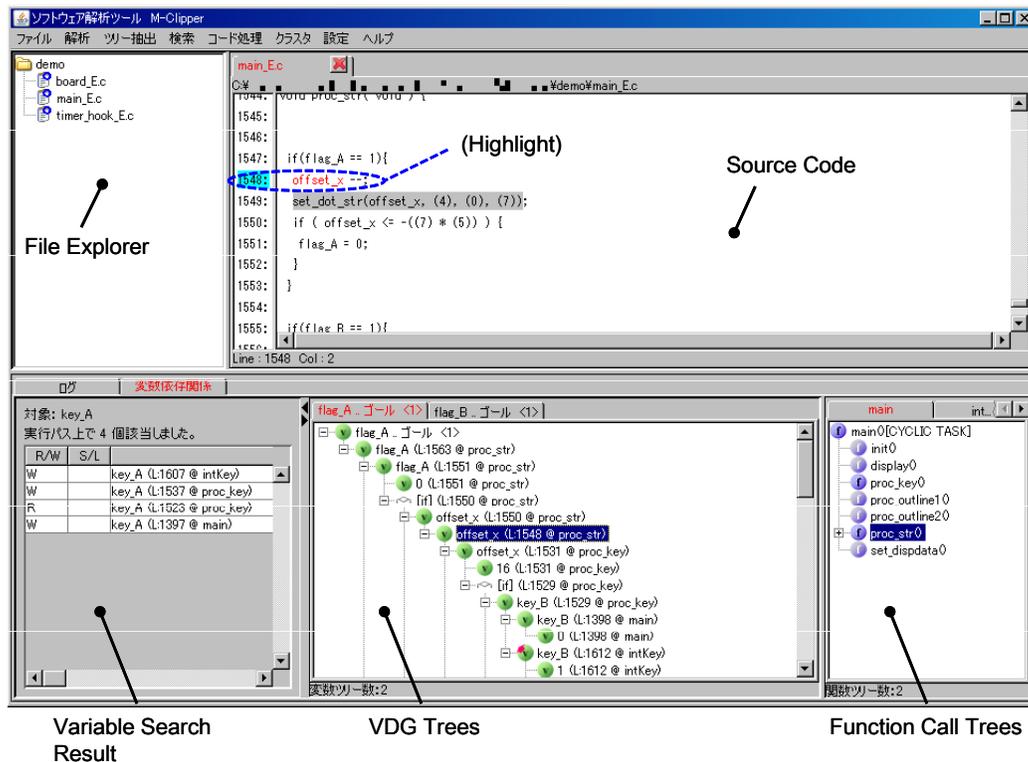


Figure 8: Screen Shot

```

A .. nodes already acquired
C .. nodes already acquired, in conditional statements

construct CFG
for all start point  $p$  of execution paths on CFG
  statement  $s \leftarrow$  statement of  $p$ 
  while  $s$  exists do
     $T \leftarrow$  stack trace of  $s$ 
    get variable nodes  $V$  corresponding to  $s$ , and assign  $T$  to  $V$ 
    connect use to def in  $V$ 
    for all node  $vi$  in  $V$  do
      for all def  $d$  in  $A$  do
        if  $d$  is able to connect to  $vi$  then
          connect  $d$  to  $vi$ 
        end if
      end for
    for all use  $u$  in  $C$  do
      if  $u$  is able to connect to  $vi$  then
        connect  $u$  to  $vi$ 
      end if
    end for
    add  $V$  to  $A$  (and  $C$ )
     $s \leftarrow$  next statement on CFG
  end while
end for

```

Figure 9: Algorithm for Analysis of VDG

tools next. Table 3 shows measured time, and those values are normalized, as the total time in "without tools" is 1. Highlighted rows are the steps to which the tools were applied. The same measured time as the first try is set for some steps in the second try which should have no effects of the tools, because the verifier could work faster the second time. The step "Model Revision" was taken because there was a false positive error at first.

Item	Description
Model checking experience of verifier	Approx. 1 year
Target	Diagnosis of input signal
Source code size (.c, .h)	94 k Logical LOC
Number of properties	2

Table 2: Condition of Experiment

When tools are used, the whole verification time on model checking is reduced 35% compared to the verification without the tools. Although no tool is used in the step "Simulation Check" in which the verifier confirms the correctness of the model with simulation, the time is reduced because the model conversion by the tool is precise. There is no reduction of time in the step "Verification" because the number of properties is only two in this case. For the four steps in which the time reduction is obtained, it is expected that more reduction will be achieved when the tools are improved.

Except for the user operations, the analysis and the slicing were completed within 5 minutes. (PC Spec: Intel(R) Core(TM)2 Duo E6550 @2.33GHz)

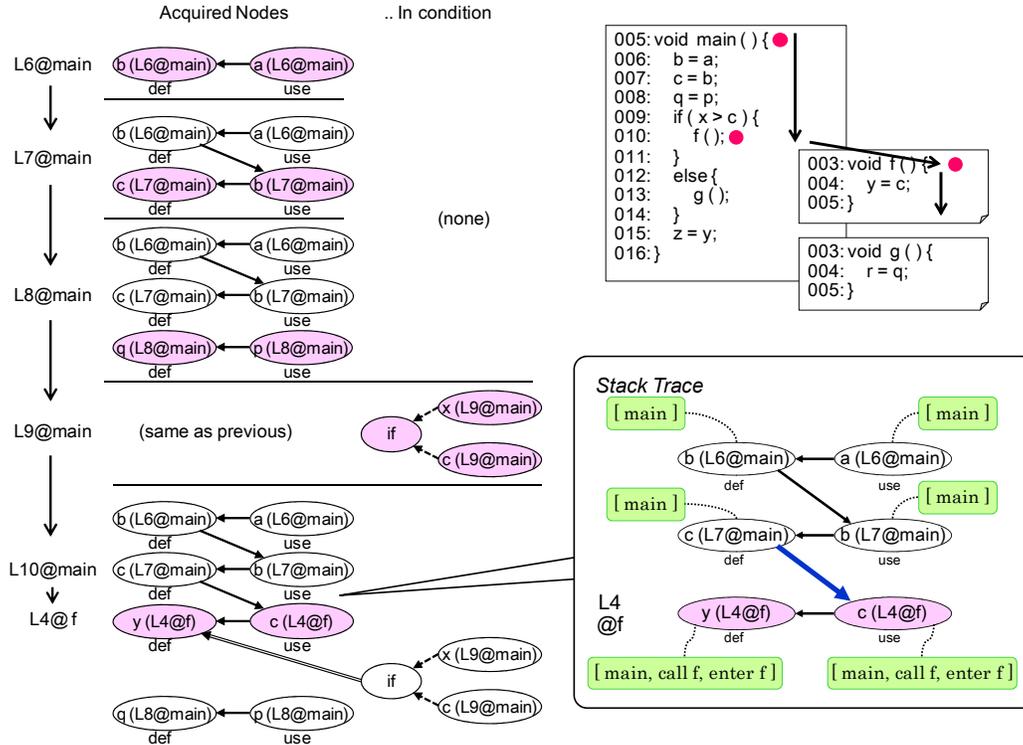


Figure 10: VDG Analysis Steps of Sample Code

#	Step	(1) without tools	(2) with tools	Difference	Decrease	Remarks
1	Planning	0.032	0.032	0.00	0.0%	same as (1)
2	Slicing	0.193	0.098	0.09	49.2%	
3	Architecture Design	0.148	0.148	0.00	0.0%	same as (1)
4	Model Conversion	0.131	0.030	0.10	77.4%	
5	Impl. External Model	0.102	0.102	0.00	0.0%	same as (1)
6	Model Synthesis	0.006	0.006	0.00	0.0%	same as (1)
7	Simulation Check	0.181	0.074	0.11	59.1%	
8	Setting Properties	0.008	0.008	0.00	0.0%	same as (1)
9	Verification	0.002	0.002	0.00	0.0%	2 properties
10	Analysis of CE	0.119	0.074	0.05	38.0%	2 counter examples
11	Model Revision	0.078	0.078	0.00	0.0%	same as (1)
	Sum	1.000	0.652	0.35	34.8%	

Table 3: Result of Evaluation (execution times are normalized)

5 Related Work

Bandera [4] slices Java code and converts them into languages for model checker such as SPIN. Modex / Feaver [6] slices C code using CFG, and converts it to Promela with a convert table and a test harness. However, these tools provide no functionality to adjust an extent of the software to be verified. The Wisconsin Program-Slicing Tool [8] [9] can analyze PDG / SDG, but it provides no functionality to adjust boundaries on PDG/SDG either. CBMC [3], BLAST [2] and SLAM [1] input source code directly

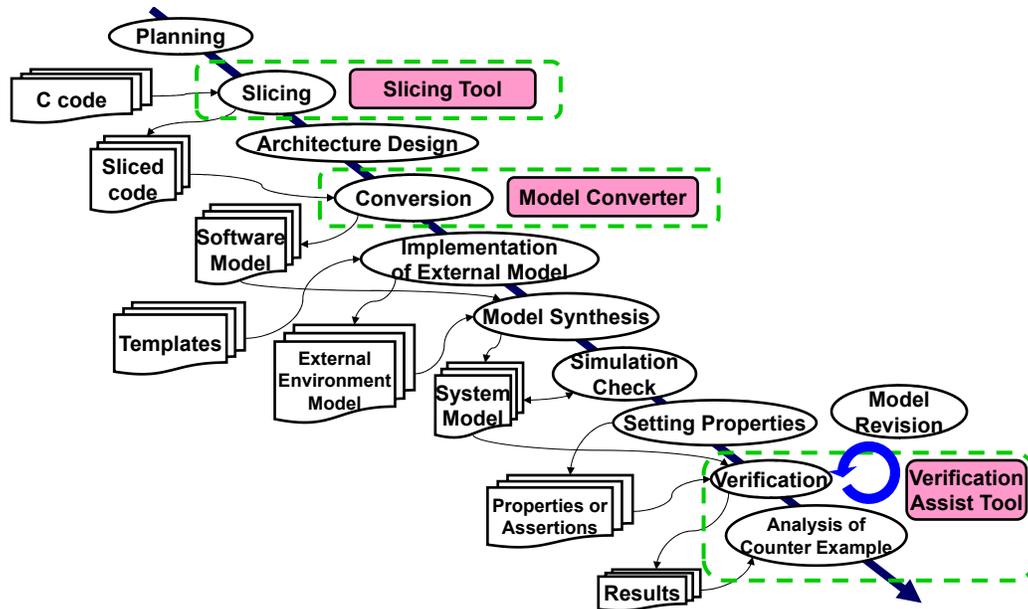


Figure 11: Verification Process

for model checking. SLAM verifies programs with software stubs as external environments. Such a static analyzer can detect defects automatically, but verification items are fixed. It is important for a flexible verification to limit a target to be verified in the software, and the slicing technique in this paper enables it with less time than the ordinal slicing.

6 Conclusion

In embedded control systems, the potential risks of software defects have been increasing because of growing software complexity. To detect software defects which are difficult to be found with usual tests or simulations, we proposed a modeling method which can generate software models from source code for model checking, with a program slicing technique based on a variable dependence graph to avoid a state explosion. VDG provides the measure to adjust a boundary between the software model converted from the source code and the external environment model, so that a large scale program can be verified with flexibility. VDG also provides the information of interface variables between software models and external environment models, so that it becomes easier for a verifier to make or select external environment models to the software models.

We applied the proposed method to one case in automotive control software to find the cause of a non-reproducible malfunction, and showed the effectiveness of the method by clarifying that it was a timing problem caused by concurrent operations of the different hardware. Furthermore, we developed the algorithm to analyze VDG, and we also developed a software tool to automate the generation of the model. We achieved a 35% decrease in total verification time on model checking. It is expected that more time reduction will be achieved when the tool is improved.

References

- [1] T. Ball, V. Levin & S. K. Rajamani (2011): *A Decade of Software Model Checking with SLAM*. *Communications of the ACM* 54(7), pp. 68–76, doi:10.1145/1965724.1965743.
- [2] D. Beyer, T. A. Henzinger, R. Jhala & R. Majumdar (2007): *The software model checker BLAST*. *International Journal on Software Tools for Technology Transfer (STTT)* 9(5), pp. 505–525, doi:10.1007/s10009-007-0044-z.
- [3] E. Clarke, D. Kroening & F. Lerda (2004): *A tool for checking ANSI-C programs*. In: *Proceedings of Tools and Algorithms for the Construction and Analysis of Systems*, pp. 168–176, doi:10.1007/978-3-540-24730-2_15.
- [4] J. C. Corbett & M. B. Dwyer et al. (2000): *Bandera: Extracting Finite-state Models from Java Source Code*. In: *Proceedings of the 22nd International Conference on Software Engineering*, pp. 439–448, doi:10.1109/ICSE.2000.870434.
- [5] G. J. Holzmann (1997): *The model checker SPIN*. *IEEE Transaction on Software Engineering* 23(5), pp. 279–295, doi:10.1109/32.588521.
- [6] G. J. Holzmann & T. C. Ruys (2005): *Effective Bug Hunting with SPIN and Modex*. In: *Proceedings of the 12th International SPIN Workshop*, pp. 24–24, doi:10.1007/11537328_3.
- [7] S. Horwitz, T. Reps & D. Binkley (1988): *Interprocedural slicing using dependence graphs*. In: *Proceedings of the ACM SIGPLAN 1988 conference on Programming Language design and Implementation*, pp. 35–46, doi:10.1145/53990.53994.
- [8] T. Reps, S. Horwitz, M. Sagiv & G. Rosay (1994): *Speeding up Slicing*. In: *Proceedings of the 2nd ACM SIGSOFT Symposium on the Foundations of Software Engineering*, pp. 11–20, doi:10.1145/193173.195287.
- [9] T. Reps & G. Rosay (1995): *Precise interprocedural chopping*. In: *Proceedings of the 3rd ACM SIGSOFT Symposium on the Foundations of Software Engineering*, pp. 41–52, doi:10.1145/222124.222138.
- [10] Mark Weiser (1981): *Program Slicing*. In: *Proceedings of the 5th International Conference on Software Engineering*, pp. 439–449.