DIAGNOSTIC REASONING IN SOFTWARE FAULT LOCALIZATION

Robert L. Sedlmeyer*
William B. Thompson
Paul E. Johnson

University of Minnesota

ABSTRACT

We present a diagnostic model of software fault localization. A *diagnoatic.* approach to fault localization has proven effective In the domains of medicine and digital hardware. Applying this approach to the software domain requires two extensions: a heuristic abstraction mechanism which infers program function from structure using recognition and transformation tactics; and a search mechanism which integrates both prototypic and causal reasoning about faults.

I. INTRODUCTION

In this paper we present a model of fault localization for program debugging based on a trouble-shooting paradigm [J], Within a diagnostic framework we define the. software fault localization task as follows: A fault exists In a program whenever its output differs from that expected by the user. These descrepancies are called fault *manifestations.* The task is to identify the cause of each manifestation which is specific enough to effect program repair.

Application of the trouble-shooting strategy has proven effective in constructing intelligent systems for hardware [2,3,4] and medical [5,6,7] diagnosis. In contrast to ve rification-based approaches (cf. [8,9,10,11] this strategy concentrates computational resources on suspect components. Applying the diagnostic approach to the software domain requires two extensions: a heuristic abstraction mechanism that infers function from structure using recognition and transformation tactics; and a search mechanism that integrates both prototypic and causal reasoning to localize faults.

Several program debugging systems evidence trouble-shooting tactics [12,13,14], but none are based on a comprehensive diagnostic theory. Sussman [1] was primarily interested in developing a theory of skill acquisition and examined the role of debugging in that context. While Miller and Goldstein [13] addressed the debugging task directly, the faults which MYCROFT analyzed were tightly constrained by the simplicity of the programs Involved. Shapiro [14] addresses more realistic debugging environments, but formulates a theory of faults rather than of fault localization.

*Current address: Department of Computer Tech., Purdue University at Fort Wayne, Indiana 46805

Our model partitions diagnostic knowledge for the software fault localization task into knowledge uspd to locate *known* and *novel* faults. For diagnosing *known* faults the knowledge base contains a hierarchy of faults which are known to occur in programs from a particular applications domain, and a set of empirical associations which relate fault manifestations to possible causes. For diagnosing *novel* faults the knowledge base contains models of implementation alternatives and execution behavior of functions indigenous to the domain. Both knowledge sources are utilized by a set of localization tactics to generate, select and test fault hypotheses.

II. KNOWLEDGE OF PROGRAM STRUCTURE AND FUNCTION

Knowledge of program structure and function is necessary for diagnosis of both known and novel faults. In the former case it is used to confirm expectations associated with a given fault hypothesis; in the latter case it is used to trace violations of expected behavior to their source. In our model this knowledge is captured in functional *phototypes.*

A functional. *phototype,* consists of four components: a set of recognition triggers, a list of pre- and post-conditions describing the execution behavior, a description of prototype components and their topology, and a List of constraints among components that must hold for recognition. Prototypes are defined at three levels of abstraction: the language, programming, and applications levels; and are hierarchically organized. A portion of the functional prototype hierarchy describing the component topology of the "master file priming read" (MFPR) function is given in Figure 1.

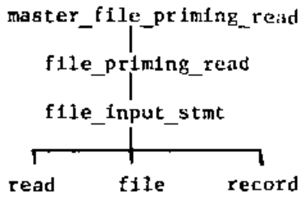

Figure 1. Topological Hierarchy for Master File Priming Read Functional Prototype

The figure shows that the MFPR consists of a single component, the file priming read. A file_priming read is only recognized as a MFPR if two constraints are met: the input record is used as a master record and the input file is used as a master file. Recognition triggers include the file priming_read and file input stmt prototypes as well as the language keyword "read".

III. KNOWLEDGE OF FAULTS

While functional prototypes describe expected program structure and function; fault *modals* describe expected defects in program structure and function. A fault *model* embodies knowledge of a particular implementation error or class of errors. This knowledge consists of conditions under which the model is applicable, a set of fault hypotheses, and a set of related fault models. A fault hypothesis is an expected defect in a functional prototype. Defects are defined as missing prototype components or violated constraints. Like functional prototypes fault models are hierarchically organized. Figure 2 details a fault model for a master file input error.

```
(master_file_priming_read error
   (triggered_by
      1.  invalid first master record
      2.  invalid first new master record
      3.  invalid master record count)
   (expected_defects
      1.  missing MFPR)
   (related_faults
      1.  insert_error)
```

Figure 2. Sample Fault Model

IV. LOCALIZATION TACTICS

Localization is performed using three tactics: *fault-driven, function-driven* and computation-*dntvan. Vault-driven.n localization* directs search at finding a particular kind of fault, such as an unexpected end_of file. *Function-driven localizatA.on* analyzes a particular functional prototype for an error. *Computation-driven localization* concentrates search on the computation of a set of program variables. Each makes use of an abstraction mechanism (Section 5) to generate and test fault hypotheses.

Multiple localization tactics are desirable for three reasons. First, they enhance the probability of finding faults since different tactics analyze the source code from different perspectives. Secondly, tactics which best fit the available information can be chosen. Thirdly, each tactic is optimal, in the sense of minimizing computational resources, for a particular localization task.

Fault-driven localization uses the fault model hierarchy to generate fault hypotheses. Output discrepancies serve as triggers for selecting a particular fault model. The more salient the trigger set the more specific the fault model and the associated fault hypotheses. Given a fault model, localization continues by first choosing one of these fault hypotheses and then invoking the abstraction mechanism to test it. If the hypothesis fails then a related hypothesis can be proposed, a different fault model can be applied, or the tactic can be abandoned.

Function-driven localization uses output discrepancies to suggest that a particular functional prototype is improperly implemented. The abstraction mechanism is then invoked to identify the code which implements the prototype. Tf the prototype cannot be found then the source code producing the most feasible match is considered as the intended implementation. The fault is identified by the statement causing failure in the matching process. This tactic can also be directed to examine a particular abstraction level for a fault. All prototypes defined at the chosen level are considered suspect. This application of the function-driven tactic is less directed and usually less desirable. Only when the output discrepancies offer no basis for prototype selection does this alternative become attractive.

Computation-driven localization traces the computation of a particular set of program variables. The computation is analyzed by extracting from the program those statements that directly affect the values of variables in the set. This extraction process is known as slicing [15]. The computation can be followed in either a forward or reverse order of execution flow. Faults are found by using the abstraction mechanism to interpret statements in the slice as members of more abstract functional prototypes and comparing the expected computational results to those derived.

V. PROGRAM ABSTRACTION

Program abstraction is performed by matching functional prototypes to the source code. Abstraction may either be expectation-driven or data-driven depending upon the localization tactic selected. Expectation-driven abstraction matches prototypes to code in a top-down manner, recursively matching components until a direct match can be made against the source code or previously recognized prototypes. Data-driven abstraction employs language-level triggers to select prototypes for matching. Prototypes matched at lower levels of abstraction serve as triggers for matching at higher levels. Once a functional prototype is identified the corresponding code is bound to it.

Recognition-based abstraction is an efficient technique, but it has limitations. Recognition of a functional prototype may fail for one of three reasons: a defect exists in the source code implementing the prototype, the wrong prototype is selected, or the source code represents an unfamiliar but correct implementation. The effectiveness of the recognition mechanism depends on the exactness of the triggering process and the richness of the alternative implementation set. Purely structural matching is augmented by functional matching. The behavior of a program segment which is inferred from language semantics, can also be compared to the functional descriptions in the prototype hierarchy.

VI. AN EXAMPLE

We have implemented an initial version of the model in a program named FALOSY (FAult Localization SYstem). FALOSY addresses faults in master file update programs [16]. In this section we illustrate FALOSY's reasoning for the master file priming read error.

FALOSY is presented with a discrepancy list and a list representation of the program's source code. The discrepancy list formally describes differences between expected and observed output. An abridged trace of FALOSY's reasoning is given in the Appendix. Numbers in parentheses refer to line numbers in the Appendix.

A production system, whose antecedents are sets of discrepancies, is used to select the initial localization tactic. In this case l.he-fault-driven tactic is chosen and the master file priming read fault model is triggered (1). FALOSY hypothesizes that the priming, read for the master file is missing (2). The abstraction mechanism is invoked to identify the corresponding prototype (4). A check is first made to determine if it has been previously identified. Since it has not, the recognition mechanism is invoked recursively to find the components oi the master file priming read prototype.

Eventually search is carried out at the source level, and three read statements (6, 26) are selected for further matching. Constraints are now checked starting with those at the lowest level in the abstraction hierarchy. The first candidate is rejected since the file identifier is not used as the master file (20). The second and third candidates are rejectee because they do not precede the update loop (24,28). No candidate satisfies all constraints and recognition fails. The original hypothesis is thus verified.

APPENDIX

```
1. Applicable fault models (m_f_pr)
2. Expected defects ((m_f_pr missing))
3. Fault hypothesis (m_f_pr missing))
4. Trying to recognize: m_f_pr
5. Triggers are (read f_i_s f_p_r
6. Matching m_f_pr to (s6 read transfile transbuf)
7. m_f_pr has role f_p_r
8. Matching f_p_r to (s6 read transfile transbuf)
9. f_p_r has role f_i_s
10. Matching f_i_s to (s6 read transfile transbuf)
11. f_i_s is primitive
12. Checking constraint (data-type file pascal_fid)
13. Checking constraint (data-type record pascal_rec)
14. Match for f_i_s succeeds
15. Checking constraint (before f_i_s update_loop)
16. Trying to recognize: update-loop
17. Recognition of update-loop succeeds
18. Match for f_p_r succeeds
19. Checking constraint (use file m_f)
20. Role file, which is bound to transfile, is used
    as a t_f whereas m_f was expected
21. Match for m_f_pr fails
22. Matching m_f_pr to (s24 read oldfile oldbuf)
23. Checking constraint (before f_i_s update-loop)
24. Hole f_i_s, which is bound to g 00014, does not
    precede role update loop, which is bound to
    g0005
25. Match for f p r fails
26. Matching m f pr to (s52 read transfile transbuf)
27. Checking constraint (before f_i_s update loop)
28. Role f i s, which is bound to gOOO18, does not
    precede role update loop which is bound to
    g()0005
29. Match for f_p_r fails
30. Recognition of m f pr fails
31. Fault hypothesis verified
32. Fault is: (m f pr missing)
```

BIBLIOGRAPHY

[11 J. DeKleer, "Local methods for localization of faults in electonic circuits," M1T-A! Memo 394, 1976.

[2] M.Genesereth, "Diagnosis using hierarchic design models," *Proceeding* of the Second* AAA] *Conference-,* August, 1982.

[3] R. Davis, H. Shrobe, w. Hamscher, K. Wieckert, M. Shirley, and S. Pol it, "Diagnosis based on description of structure and function," proceeding of *the Second* AAA! *Coference,* August, 1982.

[4] R. Hartlev, "How expert should an expert system be?", *Proceeding** IJCAI-81, August 1981.

[5] R. Patil, P. Szolovits, W. Schwartz, "Causal understanding of patient illness in medical diagnosis," *Proceedings of 1JCA1-81,* August, 1981.

[6] E. Short life, *Computer Based Medical Consultations' MYCIN,* American Elsevier, New York, New York, 1976.

[71 B. Chandrasekaran and S. Mittal, "Deep versus compiled knowledge approaches to diagnostic problem-solving," *Proceeding* of the Second* AAAI *Conference,* August, 1982.

[81 H. Wertz, "Understanding and improving LISP programs," 7 TCAI-77 *Proceeding*.*

[91 R. Ruth, "Intelligent program analysis," artifical *Intelligence,* Vol. 7, No. l, 1976

[101 C. Adam and M. "LAURA, a system to debug student programs," *Artificial Intelligence,* November, 1980.

[11] F. "Understanding and debugging computer programs," International *Journal of Man-Machine. Studies,* Vol. 12, 19S0, pp. 189-202.

[12] G Sussman, A *Computational Model of Skill Acquisition,* American Elsevier, New York, 1975.

[13]Miller, "A structured planning and debugging environment," *International Journal of Man-Machine Studies,* Vol. 11. 1978.

[14] D. Shapiro, "Sniffer: a system that understands bug," MIT-AI Memo 638, June, 1981

[151M. Weiser, "Program slices: formal, psychological, and practical investigations of an automatic program abstraction method," University of Michigan Ph.D. Thesis, 1979.

[16]B. Dwyer, "One more time - How to update a master file," *Communications of the ACM,* Vol. 24, No. 1, January 1981.