Application of Clustering Analysis to Coverage Testing for Large KBS *

Valerie Barr
Department of Computer Science
Hofstra University
Hempstead, NY 11550
vbarr@magic.hofstra.edu

Mala Mehrotra
Pragati Synergetic Research Inc.
145 Stonelake Court
Yorktown, VA 23693
mm@infi.net

TRUBAC (Testing with RUle-BAse Coverage) is a rule-base testing method which evaluates the effectiveness of test sets in covering relationships within the rule-base and identifies sections of the rule-base which have not been tested. This facilitates selection of additional test cases and can lead to a more accurate prediction of system reliability in real use (Barr 1995; 1996; 1997; 1998). However, it is not clear that the approach used in TRUBAC, which involves constructing a graph representation of the rule-base, will scale-up to larger systems. MVP-CA(Multi-ViewPoint Clustering Analysis) Tool (Mehrotra & Wild 1995), on the other hand, is an analysis tool to semi-automatically partition a rule-base system into clusters of related rules so as to expose its semantic underpinnings. MVP-CA tool is geared towards attacking the scalability problem for large rule-based systems by exposing mini-models in the underlying software architecture of the rule-base. It is for this reason that combining the techniques used in TRUBAC with the approach used in MVP-CA may lead to a very powerful rule-base testing tool.

In the TRUBAC representation each node of the graph represents a single antecedent element, a single consequent, or a logical operator such as AND or OR. Edges represent connections from antecedent elements to operators and from operators to consequents. In the MVP-CA tool, clustering a rule-base with the data-flow metric breaks up the rule-base into related sections based on the antecedent-consequent relations connecting the rules. Thus it is conceivable that TRUBAC could be modified to operate over composite nodes formed from the MVP-CA clusters, providing a high-level view of the system being tested. The advantage of providing such a clustering facility for TRUBAC is two-fold: First, it would form a comprehension-aid base for testing a large rulebase by facilitating a pictorial representation of the directed acyclic graph (DAG) generated for the rulebase, which is not possible currently when the rulebase exceeds about 50 rules. Secondly, it would provide a semantics-base for interpreting the results obtained from TRUBAC.

The Multi-ViewPoint-Clustering Analysis (MVP-CA) methodology, is geared towards understanding large knowledge-based software systems by enabling one to discover multiple, significant software structures within large knowledge-based systems (Mehrotra 1995b; 1995a; 1996; Mehrotra & Wild 1993). The current MVP-CA prototype tool is able to extract various views of the software architecture of flat knowledge-based systems through clustering rules. These clusters are suggestive of various rule-models of the system. These models can form a basis for understanding the system incrementally as well as suggesting different choices of hierarchical structures to be adopted for software engineering purposes. In the MVP-CA methodology both syntactic and semantic criteria are used for obtaining meaningful partitionings. Several distance metrics have been defined in the MVP-CA tool to capture different types of information from different types of expert systems. The MVP-CA tool provides a mechanism to structure both hierarchically (from detail to abstract) and orthogonally (from different perspectives). It provides cluster-level and pattern-based information that can be used for analyzing the knowledge-base with partial views of the subdomains. The computational complexity of test-generation is attacked through the MVP-CA tool by identification of suitable stable variables in the rule-base which can serve as focal points for the formation of subknowledge bases to be tested. These stable variables can help with equivalence partitioning of the rulebase for testing purposes. These variables surface through the clustering techniques used in MVP-CA tool. By isolating different subdomains in the range of the variable, through the generated clusters, it can provide a semantics-based handle to both generate test sets, or judge the coverage resulting from any given test set in a meaningful manner.

A coverage-based testing approach allows us to assess how well test data covers a rule-base, identifying weaknesses in the test set as well as sections of the rule-base which have not been tested. The coverage data, along with information about how representative the test data is of the general population, how likely each kind of test case is in the general population, and how well the system performs on the test data can be used to compute a reliability prediction value for the system (Barr 1998). But the reliability prediction is based on being able to put the test data into equivalence classes such that the cases in an equivalence class all execute the same inference chain through the rule-base. Thus a test-generation tool to take advantage of this knowledge can be constructed which assures that each generated test case executes a unique inference chain through the rulebase. A method for testing such a system will be to cluster the rulebase using the data-flow metric; identify equivalence regions in the rule-base; generate test data for each such region; test the clusters and generate a reliability prediction measure for the clusters; Finally combine these into a reliability prediction value for the whole system.

Alternatively, given a set of test cases, we can generate a reliability prediction measure for the system. By combining this approach with MVP-CA, we would change the level of granularity represented within the graph. Instead we could construct a meta-graph in which each meta-node represents a cluster, as identified by MVP-CA. Processing of a test case would mark edges from clusters to operators and operators to clusters. Mapping the test coverage data back to the individual rules that comprise each cluster will allow us to then generate detailed information, in a final analysis step, of logical relationships that have not been exercised by the test data.

In this way, a great deal of initial testing could be carried out over a graph which is smaller in size than the graph constructed by TRUBAC, thereby speeding up much of the testing and analysis process. Only the final testing steps would have to be carried out over the detailed graph that TRUBAC constructs. If research proves that there is no information lost by testing initially over the clusters and mapping back to the rules, the combined approaches will serve as a very powerful testing approach for large rule-based systems.

Preliminary experiments have been carried out by using the MVP-CA tool to combine the two approaches. We have successfully shown that system input to TRUBAC could first be processed by MVP-CA. Work is in progress to generate cluster-based nodes for TRUBAC which would attack the scalability problem in TRUBAC and could then be processed by a modified version of TRUBAC.

References

Barr, V. 1995. TRUBAC: A tool for testing expert systems with rule-base coverage measures. In *Proceedings of the Thirteenth Annual Pacific Northwest Software Quality Conference.*

Barr, V. 1996. *Applications of Rule-Base Coverage Measures to Expert System Evaluation.* Ph.D. Dissertation, Rutgers University.

Barr, V. 1997. Rule-base coverage analysis applied to test case selection. *Annals of Software Engineering* 4.

Barr, V. 1998. Quantitative performance prediction for rule-based expert systems. In *Workshop on Validation and Verification of Knowledge-Based Systems.* Madison, WI: Sixteenth National Conference on Artificial Intelligence.

Mehrotra, M., and Wild, C. 1993. Multi-view point clustering analysis. In *Proceedings, 1993 Goddard Conference on Space Applications of Artificial Intelligence*, 217–231.

Mehrotra, M., and Wild, C. 1995. Analyzing knowledge-based systems using multi-viewpoint clustering analysis. *Journal of Systems and Software* 29:235–249.

Mehrotra, M. 1995a. Application of Multi-Viewpoint Clustering Analysis to a Highway Maintenance System. Technical Report FHWA-RD-95087, Federal Highway Administration, McLean, VA.

Mehrotra, M. 1995b. Requirements and capabilities of the multi-viewpoint clustering analysis methodology. In *Notes for the IJCAI-95 Workshop on Verification, Validation and Testing of Knowledge-Based Systems.*

Mehrotra, M. 1996. Application of multi-viewpoint clustering analysis to an Expert Systems Advocate Advisor. Technical Report FHWA-RD-97022, Federal Highway Administration, Pragati Final Report, Yorktown, VA.