Simple, Unobtrusive Metrics For ASSET's Software Reuse Library

Larry A. Jacowitz

IBM Federal Systems Company
2611 Cranberry Sw.
Morgantown, WV 26505
Tel: (304) 594-1762
Email: jacowitz@source.asset.com

David Irving

SAIC
2611 Cranberry Sq.
Morgantown, WV 26505
Tel: (304) 594-1762
Email: irvingd@source.asset.com

Abstract

Current experience at ASSET provides insight into the way users navigate through a software reuse library. Simple, unobtrusive metrics are described for measuring library effectiveness and library efficiency. These metrics are used to analyze typical library usage.

Keywords: metrics, reuse libraries, networks

1 Background

The evolution of a network of collaborating software reuse libraries implies the need for a common set of metrics relating to user fulfilment. These libraries include domain-specific libraries at several levels (project, department, enterprise) and non-domain specific (or reference) libraries. A wide range of subject matter, user communities, and library mechanisms are represented. Yet it is highly desirable to seek a common set of metrics to be applied across the board. Such a common set will necessarily be a "lowest common denominator" and simplistic in nature. However, this simplicity is advantageous because it enables one to easily gauge how well each library is serving its users. It is general practice to employ user surveys to provide this measure. It is obvious that this may be perceived as highly intrusive to the user and it is labor intensive for the library staff. Response rates are less than 100Therefore, it is desirable to have an alternative approach, such as a preliminary assessment made by tracking the user's search and retrieval path within the library mechanism without contacting the user. The results of analyses to date at ASSET show promise for this approach.

2 Position

When a user enters his software reuse library of choice he will follow a typical sequence of commands which will permit him to identify the parameters of his problem, perform a search of the library, browse candidate assets, and extract those which meet his criteria. By tracking the user's actions in this process one can synthesize the user's objective and determine the degree of success in achieving that objective. Two indicators are sought: first, the effectiveness of the library in meeting the user's requirements; second, the efficiency with which the library does this. Define an extraction ratio, ER = number of user extractions per search; this is an indicator of the overall effectiveness of the library. Next, define an extraction index, EI = ratio of the number of user extractions to the number of search candidates found by the library mechanism; this is an indicator of the efficiency with which the mechanism finds candidates. There are other possible intermediate parameters which can add insight to the process, notably the ratio of browses to searches or extractions.

When this approach was applied to ASSET's user activity over a four-month period certain modalities became apparent. The typical user search is an iterative process; the user may be unsuccessful on the first try in matching his needs with the search mechanism schema. In this case the typical user will converge on a matching path within two or three tries. It has been noted that this same user learns from the experience and converges more quickly on subsequent searches. An often chosen alternative approach is to scan the catalog headings and call out specific candidates by their unique library identification number. It is certainly true that the extraction index and extraction ratio are a function of the classification scheme, ease of use, search mechanism peculiarities, extent of the library's holdings, and so forth. However, the simple metric technique has merit for the reasons stated.

Analysis of the cumululative statistics of ASSET usage since inception of this approach shows some expected, but some unexpected results. First, the ER, the number of extractions per search on a monthly basis has been quite consistently within a band of 0.9 to 1.5, with a cumulative average of 1.3. So users have been finding somewhat more than one component per search. Second, EI, the ratio of extractions to candidates turned up by the search, varies over a wide range. Results are erratic and not statistically significant. This is because at one extreme users may invoke the entire catalog (hundreds of candidates) just to browse through it; at the other extreme a user extracts a component each time he calls for a known component by identification number. Third, we have found that the ratio of searches:browses:extractions tends to be stable from month to month about roughly 1:3:1

3 Comparison

Traditionally user surveys are the only means of determining the effectiveness of a library to meet user needs. This method is time- consuming for the library staff and obtrusive to the user. Often low response rates pertain. An indirect approach is desirable.

The approach described in this paper offers an alternative which is simple and unobtrusive. In addition to providing information about the users' functional requirements, this technique is useful in evaluating the ease of use of the library, the effectiveness of the classification scheme, etc.

This technique should be extended to other software reuse libraries. Comparative analysis of different, heterogeneous libraries may provide ideas for improving the state of practice in software reuse libraries.

4 Biography

Larry Jacowitz is the ASSET Director for IBM's Federal Systems Company, the prime contractor for ASSET. The ASSET facility is located in Morgantown, WV, and is implemented under ARPA's STARS program. Dr. Jacowitz previously held management positions in a variety of federal civil programs, including DOE, Dept. of Agriculture and Immigration & Naturalization Service, and NASA's space shuttle and space station programs at locations in Gaithersburg, MD, Downey, CA, Owego, NY, and Huntsville, AL. He holds BS, MS and PhD. degrees in engineering from MIT and OSU.

David Irving is a Senior Software Engineer at ASSET. He is responsible for the maintenance and improvement of the software reuse library mechanism. He also developed the tracking technique described in this paper. Mr. Irving's previous assignments at SAIC have been in the area of Ada code development and IV&V on projects for DOE and missile projects for the Navy. Mr. Irving holds a BA degree in mathematics from Univ. of Pennsylvania and M.S. in physics from Univ. of Tennessee.: