MVP-CA Methodology for the Expert System Advocate's Advisor (ESAA)

PUBLICATION NO. FHWA-RD-97-022 NOVEMBER 1997

U.S. Department of Transportation
Federal Highway Administration

Research and Development
Turner-Fairbank Highway Research Center
6300 Georgetown Pike
McLean, VA 22101-2296

FOREWORD

This report describes the application of Multi-ViewPoint Clustering Analysis for structuring an expert system rule base in various ways to capture explicit knowledge and implicit information about the knowledge contained in the system. The methodology was applied to the relatively small, but ill-structured, expert system—Expert System Advocate's Advisor. The motivation for clustering, an overview of the clustering tool, and the results of applying the methodology are covered in this report.

The audience for this report are the developers of transportation-related artificial intelligence (AI) programs since the report describes the application of a sophisticated approach to a recognized problem in the development of AI-based tools. Researchers will also be interested in the application of the methodology.

A. George Ostensen
Director, Office of Safety and Traffic
Operations Research and Development

NOTICE

Technical Report Documentation Page

1. Report No. FHWA-RD-97-022	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle MVP-CA METHODOLOGY FOR THE EXPERT SYSTEM ADVOCATE'S ADVISOR (ESAA)		5. Report Date November 1997
		6. Performing Organization Code
7. Author(s) Mala Mehrotra		8. Performing Organization Report No.
9. Performing Organization Name and Address Pragati Synergetic Research, Inc. 145 Stonelake Ct. Yorktown, VA 23693		10. Work Unit No. (TRAIS) M6
		11. Contract or Grant No. DTFH61-95-P-00216
12. Sponsoring Agency Name and Address Office of Safety and Traffic Operations R&D Federal Highway Administration 6300 Georgetown Pike McLean, VA 22101-2296		13. Type of Report and Period Covered Final Report 12/1/94 - 3/31/95
		14. Sponsoring Agency Code

15. Supplementary Notes

Contracting Officer's Technical Representative (COTR): James A. Wentworth, HSR-2

16. Abstract

The Multi-ViewPoint Clustering Analysis (MVP-CA) tool is a semi-automated tool to provide a valuable aid for comprehension, verification, validation, maintenance, integration, and evolution of complex knowledge-based software systems. In this report, the contractor has shown the value of using their MVP-CA tool for analyzing a poorly structured rule base such as the Expert System Advocate's Advisor (ESAA). In particular, they have shown that MVP-CA methodology is capable of exposing the current underlying software architecture of the knowledge base. Such capability can pave the way towards showing alternate ways of restructuring the system while the system is evolving. Also, the MVP-CA technology juxtaposes rules with similar content and structure so that incomplete specifications on various variables and qualifier values in the rules surface quite easily. Through grouping of rules from multiple perspectives, this report also shows how redundancies, inconsistencies, and anomalies become apparent easily. MVP-CA methodology stresses the proper structuring of a complex system first, so that understanding the system becomes a basis for verification, validation, testing, and maintenance.

17. Key Words Clustering, Software Engineering, Anomalies, Verification, Software Architecture, Incompleteness	18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22161.

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 89	22. Price

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

TABLE OF CONTENTS

LIST OF FIGURES

Figure No. Page

LIST OF TABLES

Table No. Page

INTRODUCTION

Software systems have become too complex to be understood through manual inspection alone. It is imperative to build software tools that can attack the problem of complexity by exposing the mini-models in the underlying software architecture. For these codes to transition into operational environments, the software has to be validated and verified. Proper structuring of these systems is essential as a first step towards understanding, which can then become a basis for verification and validation (V&V), testing, and maintenance. In order to build reliable systems, it is important that the knowledge in the system be suitably abstracted, structured, and otherwise clustered in a manner that facilitates its understanding. Development and maintenance of complex systems will require the ability to abstract overall concepts in the system at various levels of detail and to consider the system from different points of view. A semi-automated tool, such as the Multi-ViewPoint-Clustering Analysis (MVP-CA) tool, which allows the user to focus attention on different aspects of the problem, can provide a valuable aid for comprehension, verification, validation, maintenance, integration, and evolution of such complex software systems.

Existing approaches to structuring systems in the conventional framework are limited in a major way; they only provide a single viewpoint of a system. The contractor believes that **no one single structuring viewpoint is sufficient to comprehend a complex system.** However, note that in the research, "multiple viewpoints" do not mean different representational forms of the same software, such as state transition views or functional flow views. Rather, it means **different meaningful ways of organizing the same information in the knowledge base such that interrelationships between various conceptual aspects of the domain are made more explicit.** This is similar to organizing a deck of 52 cards in at least four valid ways. One arrangement asserts the suit aspect of the cards, giving 4 clusters with 13 cards in each cluster. Another asserts the rank perspective, giving 13 clusters with 4 cards per cluster. An equally meaningful viewpoint can divide the deck of cards into two clusters, numerical in one cluster and face cards in another. Also, if all the red cards were separated from the black ones, another viewpoint would emerge. Similarly, in presenting a software system from different perspectives, a better understanding of the system as a whole can be achieved because subtle interrelationships can become evident only when the system is clustered in various meaningful ways.

By exposing the underlying domain knowledge in knowledge-based software systems from multiple viewpoints, various subtle interrelationships in the domain become evident that were not possible through manual inspection alone. Focus can be shifted from the syntax and representational aspects of the software to the more basic semantic aspects of its design. In fact, one of the significant potential benefits of clustering based on the multi-viewpoint methodology is to reveal to the user previously unseen structures in the knowledge base that either give additional insight into the verification and validation aspects of the system or indicate problems in its organization and suggest alternative reorganizational choices.

This report outlines a feasibility study performed with the contractor's MVP-CA tool on a small, but poorly structured knowledge base — the Expert System Advocate's Advisor (ESAA) — that was known to have multiple errors. The intent was to see how many and what type of analytical information the MVP-CA tool was capable of providing for such a knowledge base. Earlier, the contractor had used this tool to study PAMEX (Pavement Maintenance Expert System), a well-structured knowledge base. Results are available in reference 5.

OVERVIEW OF MVP-CA

The Multi-ViewPoint-Clustering Analysis (MVP-CA) methodology, developed by the contractor, is geared towards understanding large knowledge-based software systems by enabling the user to discover multiple, significant structures within them. The contractor's 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 or mini-models of the system. These models can then suggest different choices of hierarchical structures that could be adopted during development or evolution of the software system.

In this subsection, a brief overview of the current status of the software is provided. The current MVP-CA tool is divided into two phases: the *Cluster Generation Phase* and the *Cluster Analysis Phase*. In the *Cluster Generation Phase*, focus is on generating meaningful clusters through statistical and semantics-based measures. Statistics are generated in terms of *cohesiveness* internal to a cluster and *dispersion* of various patterns across clusters, as well as *coupling* across clusters. In the *Cluster Analysis Phase*, focus is on performing a statistical and functional analysis of the generated clusters. Statistical output generated from the previous phase aids in the analysis and forms the basis for formulating better constraints in order to improve the quality of subsequent clusterings. For example, iterating on different values for *distance metric* and choice of appropriate *grouping range* generates various meaningful viewpoints. Functional analysis of the clusters in the Cluster Analysis Phase captures the key concepts that underlie the generated clusters.

Some of the details of the MVP-CA tool are graphically represented in figure 1. The current MVP-CA software consists of a couple of programs that have to be manually called by the user in order to perform the clustering and to analyze the generated clusters. In this phase, the existing rule base, together with the concept focus list, feeds into the front-end interpreter. A *concept focus list* is formed from the pool of all patterns present in the knowledge base. It provides the semantic basis for clustering the rules in the knowledge

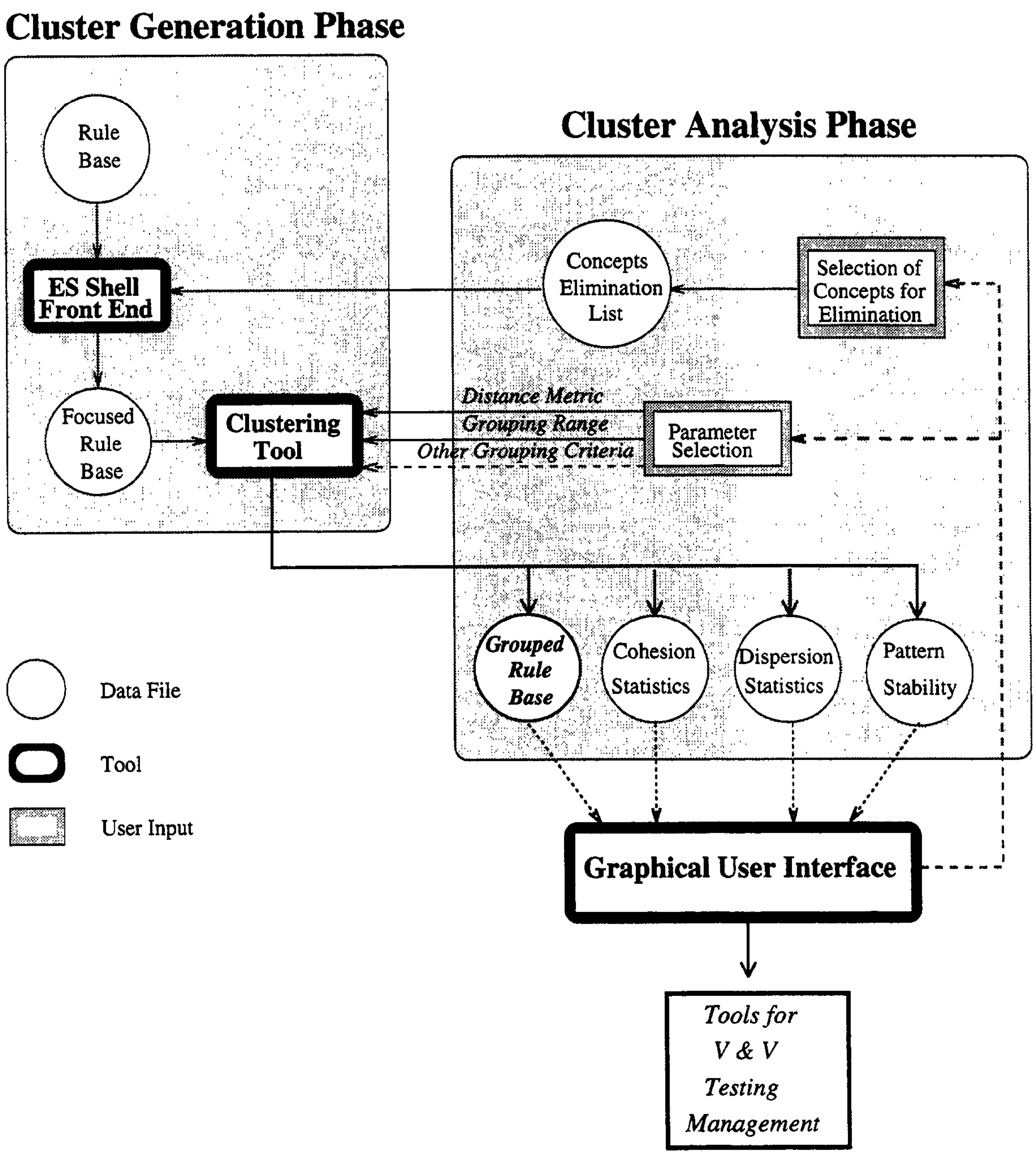

Figure 1: Software architecture of MVP-CA.

base by acting as a filter for patterns that are allowed to play a role in the clustering process. In general, patterns with a frequency of one do not contribute to the clustering process in a meaningful way and can be eliminated safely. Similarly, patterns with very high frequency can also be perceived as noise for the clustering algorithm. However, they have to be carefully weeded out so that the semantics of the rule base are not altered in any substantial manner.

Next, the interpreter parses the modified rule base and transforms it into an internal form required by the clustering tool. It should be noted that the **tool itself is language-independent**. The clustering algorithm is an agglomerative one with the most similar clusters being merged at each iteration. The definition of "similarity" varies at each run as it is defined by the distance metric chosen for that run. This pattern of mergings forms a hierarchical cluster from the single-member rule clusters to a cluster containing all the rules. In order to aid in the analysis of this hierarchy and to highlight high- and low-dispersion concept patterns, various statistics are recorded during the Cluster Generation Phase. They are detailed in the following paragraphs:

Distance metric measures the relatedness of two rules in a rule base by capturing different types of information for different classes of expert systems. There are four distance metrics that have been implemented so far. Classification systems yield easily to a data-flow grouping and, hence, information is captured from the consequent of one rule to the antecedent of other rules. This defines our *data-flow* metric. In a monitoring system, since the bulk of domain information required for grouping is present in the antecedents of rules, the *antecedent* distance metric captures information only from the antecedents of rules. Alternatively, grouping the rule base by information from the consequents only, gives rise to the *consequent* metric. The *total* metric is general enough and captures information from both sides of rules to take care of systems where a combination of the above programming methodologies exist. The kind of distance metric to be used is a function of both the nature of the task performed by the rule base (classification, diagnosis, control) and the nature of the analysis required by the user (restructuring, testing, comprehension, reuse, etc.), as we shall see from the later discussions.

For a given clustering, C, the *cohesiveness measure* is an index of the similarity of rules belonging to the same group and it measures the number of concept patterns shared among all rule pairs in the group. Overall cohesiveness of a clustering is the cohesiveness for each group averaged over all groups for a given clustering. In the preliminary stages of exploration of a knowledge-based system, cohesion plots give valuable insight into the range for optimal partitioning regions, G_{max} to G_{min}, to be examined.

Dispersion measure indicates the degree to which a single pattern is dispersed among the clusters. At the onset of clustering, each rule belongs in its own group and dispersion measure gives essentially the *frequency* of occurrence of each pattern. It gives a snapshot of the relative importance of patterns and is very useful in spotting non-domain-related patterns that need to be weeded out. This helps in forming the initial concept focus list that is input to the MVP-CA tool. Removing these patterns helps define the clusters more distinctly — a process that we call "sharpening." This is also reflected in the cohesion plots as the drop points for the plateau regions become much sharper.

As the user generates different clusterings of the rule base, dispersion statistics continue to guide the removal of select patterns from the rule base by providing a measure for shared concept patterns in a group. In general, the dispersion of patterns that relate to key concepts will decrease and stabilize as the mergings progress. Such patterns, therefore, are likely to represent concepts characterizing the clusters to which they belong. These patterns can be suppressed in the subsequent runs of the clustering algorithm, revealing subtle alternate viewpoints — a concept that has been termed "multi-viewpoint clustering."

At any point in the clustering process, the MVP-CA tool keeps track of the dispersion of each pattern, i.e., the number of groups in which it occurs. The merging of two groups may cause a pattern or a combination of patterns to occur only in the new group. When this happens, the pattern or pattern combination is flagged as having become **stable** with respect to this group. Stability is a very important criterion as it alone determines the degree to which a cluster has become "firewalled." The stable pattern in this group is analogous to a local variable in a conventional procedure. The user can perform checks on its consistency with respect to its use in other rules. Incompletely specified patterns also surface through these clusters. Redundancy conditions can be flagged if the rules have been overspecified with respect to the stable pattern. Examples of these conditions are illustrated later on.

In the MVP-CA methodology, both syntactic and semantic criteria are used for obtaining meaningful partitionings. This methodology can be summarized as follows. First, form a preliminary clustering for a particular distance metric. Identify the "noise" patterns that are interfering with the formation of perfect groups by examining the dispersion statistics generated from the clustering. Sharpen the current viewpoint by identifying very highly or very sparsely dispersed "noise" patterns and by weeding them out of the clustering process. Currently, this semantic criteria of identifying and weeding out patterns are performed manually in the form of a concept focus list. Next, identify the primary viewpoint for this clustering by noting the stable and dominant pattern in the groups along with its associated attributes(6,7). Note that a dominant pattern may not necessarily be the stable pattern in the cluster as it may occur across several clusters. Remove the dominant concept pattern(s) and the associated attributes that are responsible for the primary viewpoint from all the rules before reclustering the rule base. The new clustering will reveal secondary (tertiary, etc.) viewpoints. Cluster the knowledge base with other distance metrics, applying the above procedure to get additional viewpoints on the rule base.

ESAA

The Expert System Advocate's Advisor (ESAA) is an expert system written in EXSYS (expert system shell). It consists of 68 EXSYS rules, 36 declared qualifiers (out of which only 27 were used in the rules), 27 variables, and 11 choices or conclusions. Even though ESAA is a small expert system, unlike the previous much larger expert system known as PAMEX (also written in EXSYS with 327 rules), this expert system did not exhibit careful upfront software designing. The contractor used the MVP-CA tool to uncover any software architectural flaws, as well as inconsistencies and anomalies among the rules in ESAA. A listing of the ESAA rules is provided in appendix A.

Since ESAA is a small knowledge base, some procedures could be sidestepped to conserve time. These fall into two broad categories of pattern numbering and rule labeling conventions.

Pattern Numbering Conventions

Since the qualifiers are in the form of context-sensitive questions, the contractor has manually abbreviated them and assigned pattern numbers to them. The front-end interpreter can only take rule bases written in context-free grammar in order to do automatic pattern number generation for them. However, the procedure for assigning pattern numbers manually has been performed in the same spirit and methodology of an automatic algorithm. The general guidelines followed for generation of pattern numbers in ESAA are as follows:

A pattern is any qualifier (abbreviated), variable, number, operator, or choice (abbreviated). Each pattern in the knowledge base is assigned a unique number. All abbreviated qualifiers are numbered from 1 to 37; the values of the qualifiers are numbered from 100 to 208. This design choice allows all the stable qualifiers to be grouped close together because the stable patterns are sorted numerically before their presentation. Also, it is important to note that the same value for different qualifiers is mapped to unique pattern numbers. In other words, *yes* for ID_NEED has a different pattern number than a *yes* for the M_NEED qualifier. It is important to distinguish the two *yes* values, because otherwise there would be a false relationship between the rules.[1] This type of context-oriented numbering scheme also helps substantially with automatic detection of incomplete qualifier value specifications, as will be evident from our experimental results later on.

[1] It is not difficult to automate this process for EXSYS rule bases as these values are declared with their qualifiers in the Qualifier section. In fact, most expert system shells provide this type of declaration facility. Thus, the numbering system can work based on the Declarations section instead of generating pattern numbers from the patterns in the rules directly. In the spirit of emulating an automatic algorithm, the contractor even numbered the qualifiers (and their values) that were declared, but never used in any of the rules.

The 27 variables declared in the Variables section of ESAA have been numbered from 40 to 66. The 10 conclusions listed in the Choices section have also been abbreviated and numbered from 70 to 80. A total of 23 arithmetic, logical, and relational operators (identified in the Formulas section) have been numbered from 220 to 242. Since logical and relational operators do not mean anything without taking into consideration the context in which they occur, each of these operators has been concatenated with their numerical counterpart before a unique pattern number is assigned to the combination. Thus, the logical operators and arithmetic operations are combined with the number and are taken together, i.e., "<> 0" (not equal to zero) is mapped to pattern number 220, etc. This facility, where two or more patterns can be concatenated to be treated as a single-pattern entity, is being provided in the user interface of future versions of the tool.

Providing an infrastructure that allows the capture of contextual information is the first step towards making the cluster-generation process in the tool semantically oriented. It also allows analysis of the rule base at various conceptual units of informational chunks. Thus, if each clause or formula is mapped to a pattern number, a higher unit of information can be used to analyze the knowledge base. Bellman and Landauer's incidence matrices with clauses and formulas can be automatically generated through this infrastructure (see references 1, 2, 3, and 4). Thus, the semantic basis for clustering could be any one of the basic entities, patterns, formulas, or clauses.

Rule Labeling Conventions

In displaying groups on paper (and on the computer screen), it is not possible to print out the whole rule, so the contractor labeled each rule with a name. Since the rules did not have rule names (to signify their semantic content), the contractor abbreviated the various formulas in the rules and used them as rule labels. Labels that reflect the content of the rules were used. The labeled rules are as follows:

```
Rule #        Label

rule  1   ID_NEED=Y   ->  EST_BEN=10
rule  2   M_NEED=NN   ->  EST_BEN(+0)
rule  3   M_NEED=LIT   ->  EST_BEN(+2)
rule  4   M_NEED=SM   ->  EST_BEN(+6)
rule  5   M_NEED=LOT   ->  EST_BEN(+10)
rule  6   IMPACT=NN   ->  EST_BEN(+0)
rule  7   IMPACT=LIT   ->  EST_BEN(+2)
rule  8   IMPACT=SM   ->  EST_BEN(+6)
rule  9   IMPACT=MAJ   ->  EST_BEN(+10)
rule  10  MAJ_IMPR=NN   ->  EST_BEN(+0)
rule  11  MAJ_IMPR=LIT   ->  EST_BEN(+2)
rule  12  MAJ_IMPR=SM   ->  EST_BEN(+6)
```

```
rule  13  MAJ_IMPR=LOT  ->  EST_BEN(+10)
rule  14  COMPL=HG  ->  EST_BEN(+10)
rule  15  COMPL=MED  ->  EST_BEN(+6)
rule  16  COMPL=LOW  ->  EST_BEN(+2)
rule  17  COMPL=NN  ->  EST_BEN(+0)
rule  18  TRNG_TOOL=LOT  ->  EST_BEN(+10)
rule  19  TRNG_TOOL=SM  ->  EST_BEN(+5)
rule  20  TRNG_TOOL=NNVLIT  ->  EST_BEN(+0)
rule  21  END_USER=CL  ->  EST_BEN(+4)
rule  22  END_USER=TECHVPRO  ->  EST_BEN(+8)
rule  23  END_USER=ADM  ->  EST_BEN(+10)
rule  24  C_AN_T<>0  ->  AN_T_SAV(C_AN_T, ES_AN_T)
rule  25  YRLY_ANAL_EXEC_T>1000  ->  EST_BEN(+10)(0)
rule  26  C_EXEC_T<>0  ->  EXEC_T_SAV(C_EXEC_T,ES_EXEC_T)
rule  27  C_AN_T<>0  ->  RAW_AN_EXEC(AN_T_SAV,EXEC_T_SAV)
rule  28  NUM_REDO=ES_NUM_REDO  ->  REDO_SAV(RAW_AN_EXEC,NUM_REDO,ES_NUM_REDO)
rule  29  YRLY_FREQ>0  ->  YRLY_AN_EXEC_T(YRLY_FREQ,REDO_SAV)
rule  30  ADV=Y  ->  ORG_RSK_F(+0)(+10)
rule  31  MGMT_SUP=Y  ->  ORG_RSK_F(+0)(+10)
rule  32  MGMT_LV=1V2  ->  ORG_RSK_F(+0)
rule  33  MGMT_EXP=Y  ->  ORG_RSK_F(+0)(+10)
rule  34  MDL=Y_ALG  ->  DOM_RSK_F=1
rule  35  MDL=Y_MAN  ->  DOM_RSK_F=5
rule  36  MDL=N  ->  DOM_RSK_F=10
rule  37  REQ_PERF=50  ->  DOM_RSK_F(+4)
rule  38  REQ_PERF=80  ->  DOM_RSK_F(+6)
rule  39  REQ_PERF=100  ->  DOM_RSK_F(+10)
rule  40  INTER=Y  ->  DOM_RSK_F(+10)(+0)
rule  41  USER_ENTH=NN  ->  USER_RSK_F=10
rule  42  USER_ENTH=LTL  ->  USER_RSK_F=6
rule  43  USER_ENTH=SM  ->  USER_RSK_F=2
rule  44  USER_ENTH=LOT  ->  USER_RSK_F=0
rule  45  COMP_PROF=NN  ->  USER_RSK_F(+10)
rule  46  COMP_PROF=LTL  ->  USER_RSK_F(+6)
rule  47  COMP_PROF=SM  ->  USER_RSK_F(+2)
rule  48  COMP_PROF=LOT  ->  USER_RSK_F(+0)
rule  49  MGMT_LV=3  ->  ORG_RSK_F(+2)
rule  50  MGMT_LV>3  ->  ORG_RSK_F(+5)
rule  51  ORG_RSK_F,USER_RSK_F,DOM_RSK_F  ->  EST_RSK(ORG_RSK_F,USER_RSK_F,DOM_RSK_F)
rule  52  EST_BEN>-1,EST_RSK>-1  ->  RATIO(EST_BEN/EST_RSK)
rule  53  EST_BEN<40,EXP_WHY_BEN  ->  exp_gap
rule  54  EST_RSK>40,EXP_EST_RSK  ->  exp_sz
rule  55  RATIO<1,EXP_LOWRATIO  ->  exp_low
rule  56  RATIO>1  ->  go_ahd,LK_GOOD
rule  57  END=Y  ->  inp_comp
```

```
rule  58  BEG_BEN=Y  ->  beg_ben
rule  59  DISP_BEN+1000>EST_BEN  ->  conc_ben
rule  60  ORG_RSK_F<DISP_ORG_RSK+1000  ->  conc_org_rsk
rule  61  DOM_RSK_F<DISP_DOM_R_F+1000  ->  conc_dom_rsk
rule  62  USER_RSK_F<DISP_USER_R_F+1000  ->  conc_user_rsk
rule  63  ID_NEED=N  ->  EST_BEN=0
rule  64  M_NEED=U,EXP_AVL,EXP_LVL,RES,END_USERS  ->  M_NEED=LOT
rule  65  COMPL=U,REG=N,EXPED_COMPL=Y  ->  COMPL=H
rule  66  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y,EXP_STF=Y  ->  ID_NEED=Y
rule  67  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y  ->  ID_NEED=Y
rule  68  ID_NEED=U,PRBLM=CMPLX,EXP_STF=Y  ->  ID_NEED=Y
```

- In the rule labels only, the abbreviated form of the qualifiers appears. For example, ID_NEED is an abbreviated form for

  "Qualifier 1: Is there a clearly identified need that can best be addressed by an expert system? Note: existing conditions must be known and described to establish a need."

  Similar abbreviations for the rest of the qualifiers in ESAA are evident by comparing the rule labels with the actual rules.

- In labeling the rules, the left-hand side and the right-hand side of rules are delineated by the symbol `->` .

- The equal sign ("=") on the left-hand side of a labeled rule is a check for equality. On the right-hand side, it is used for assignment.

- In labeling the rules, all arithmetic operators were ignored so as to not clutter up the presentation of clustered rules.

- Increment of a value was represented as follows, e.g., EST_BEN(+4) in the rule label means the variable "EST_BEN" is incremented by four. On the other hand, "EST_BEN=4" in the rule label means "EST_BEN" has been assigned the value 4.

- The abbreviation "ORG_RSK_F(+0)(+10)" means, in one case, "ORG RISK FACTORS" is given the value zero and, in the other case, it is assigned the value 10. This is indicative of an if-then-else construct in the rule.

- Conclusions or choices are abbreviated and labeled in lower case to distinguish them from the qualifiers.

- Confidence factors have been ignored in the rule labelings.

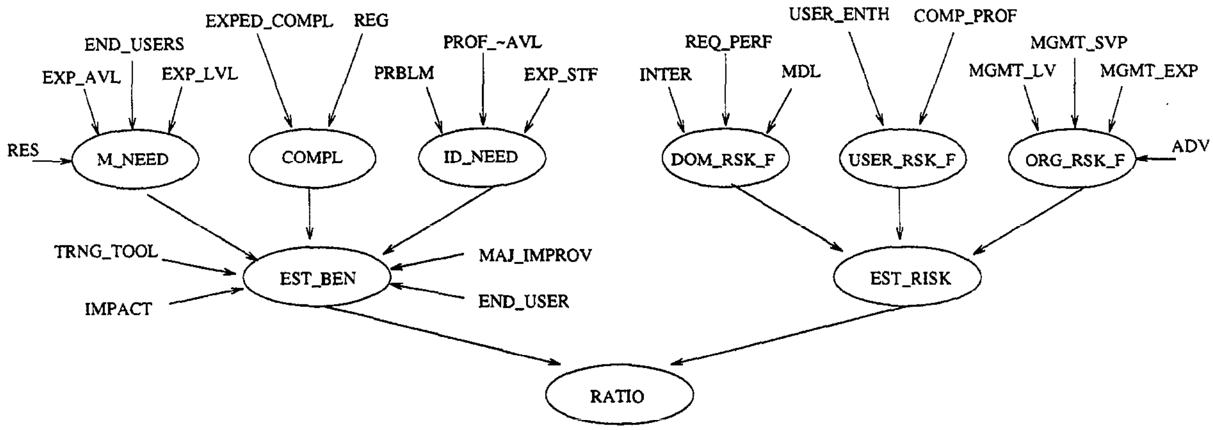

Figure 2: Software architecture of ESAA.

EXPERIMENTAL RESULTS

Having generated the pattern numbers and labeling all the rules, ESAA could be clustered with all four distance metrics. Since ESAA is a small rule base, the MVP-CA tool was allowed to generate clusters all the way from 67 clusters to 1 cluster, and the contractor could do a comprehensive analysis of all the clusters generated. Each distance metric provides useful information for studying a particular aspect of the knowledge base. The data-flow metric is useful in generating the underlying software architecture of the rule base. The antecedent metric is useful in exposing the incomplete regions as well as the anomalous regions in the rule base. The consequent metric is useful in studying the general trends under which various conclusions were being asserted in ESAA. Since dominant patterns are found in the antecedents of the rules in ESAA, information provided by the total metric is very similar to that provided by the antecedent metric.

Data-Flow Metric

The data-flow metric produces clusters of rules that are chained through a right to left dependency. In particular, it helps in understanding the "def-use" aspect of the rule base, thus providing a window into the software architecture of the rule base. Figure 2 exemplifies this use of the MVP-CA tool. In performing the analysis, the contractor tracked the cluster in which each pattern that had a frequency of more than one stabilized first. The contractor then looked for the dominating patterns in that cluster and formed the subtree with the dominating pattern as the parent and the stable pattern as the child. Since qualifiers mostly occur on the left-hand side of rules and choices on the right-hand side, the variables can be assumed to carry the dependency information. Typically, therefore, the variables will assume parent roles and qualifiers will feed into them.

However, in instances where some qualifier values were unknown, other qualifier values were checked to determine a value for the unknown qualifier. This gave us some qualifiers that played a subsidiary role to other unknown qualifiers. Such qualifiers are exposed in this section as well, with their associated clusters.

It is easy to spot the patterns with a frequency of one as they are stable in the initial state, when each rule is in its own group.

The following is an elaboration of the results from the data-flow metric and a highlighting of the important clusters as they form during the clustering process. There is also an indication of how these clusters contribute to our knowledge of the rule base so as to justify their importance. As noted before, the initial clustering has 68 clusters, with each rule in its own cluster. At each step, two clusters are merged to form a new cluster, while the other clusters are not changed. Thus, when cluster n is mentioned below, it refers to the newly formed cluster at that step.

In Cluster 56, when MDL stabilized, the software architecture of the different subtrees in ESAA started emerging. MDL is an abbreviation for the qualifier question "Is there a procedure to be used as a model for the expert system?" As can be seen from this cluster, MDL is a stable pattern in a cluster where Domain Risk Factor is the dominating pattern. The latter has not stabilized as early in the clustering process as there are other dependent variables for this concept. However, MDL can be assigned to be a child of the Domain Risk Factor, thus forming the first step in the software architectural analysis of ESAA.

```
Cluster No.: 56
Properties of newly merged group:
Number Rules in Group = 4
  RuleNo  Description
      34  MDL=Y_ALG  ->  DOM_RSK_F=1
      61  DOM_RSK_F<DISP_DOM_R_F+1000  ->  conc_dom_rsk
      35  MDL=Y_MAN  ->  DOM_RSK_F=5
      36  MDL=N  ->  DOM_RSK_F=10
Number Stable Patterns in Group = 6
  PattNo  Description
       5  MDL
      64  DISP_DOM_R_F
      73  conc_dom_risk
     117  yes, an algorithmic system(MDL)
     118  yes, a manual system(MDL)
     119  no(MDL)
```

USER_ENTH in Cluster 53 is the next one to stabilize. USER_R_F is the dominating pattern for this cluster. It can be concluded from this observation that USER_ENTH is feeding into the concept of USER_R_F and, hence, USER_ENTH is a child of USER_R_F. Also note the presence of Rule 62, which does not have any instance of USER_ENTH. However, it is connected to this rule group through chaining. That is, if A – > B and B – > C, then A and C are linked in the data-flow clustering. Thus, this cluster reveals how our clustering algorithm is successful in placing rules in their appropriate context.

```
Cluster No.: 53
Properties of newly merged group:
Number Rules in Group = 5
  RuleNo  Description
      41  USER_ENTH=NN  ->  USER_RSK_F=10
      62  USER_RSK_F<DISP_USER_R_F+1000  ->  conc_user_rsk
      42  USER_ENTH=LTL  ->  USER_RSK_F=6
      43  USER_ENTH=SM  ->  USER_RSK_F=2
      44  USER_ENTH=LOT  ->  USER_RSK_F=0
Number Stable Patterns in Group = 8
  PattNo  Description
      18  USER_ENTH
      65  DISP_USER_R_F
      74  conc_user_risk
     157  none(USER_ENTH)
     158  a little(USER_ENTH)
     159  some(USER_ENTH)
     160  a lot(USER_ENTH)
     232  10
```

The next pattern to stabilize is IMPACT in Cluster 43. Through this group, it can be deduced that IMPACT feeds into the EST_BEN concept. Again, note the presence of RULES 52, 53, and 59 due to the transitive dependency relationship.

```
Cluster No.: 43
Properties of newly merged group:
Number Rules in Group = 12
  RuleNo  Description
       1  ID_NEED=Y  ->  EST_BEN=10
      59  DISP_BEN+1000>EST_BEN  ->  conc_ben
      52  EST_BEN>-1,EST_RSK>-1  ->  RATIO(EST_BEN/EST_RSK)
       2  M_NEED=NN  ->  EST_BEN(+0)
      53  EST_BEN<40,EXP_WHY_BEN  ->  exp_gap
       3  M_NEED=LIT  ->  EST_BEN(+2)
       4  M_NEED=SM  ->  EST_BEN(+6)
```

```
      5  M_NEED=LOT  ->  EST_BEN(+10)
      6  IMPACT=NN  ->  EST_BEN(+0)
      7  IMPACT=LIT  ->  EST_BEN(+2)
      8  IMPACT=SM  ->  EST_BEN(+6)
      9  IMPACT=MAJ  ->  EST_BEN(+10)
Number Stable Patterns in Group = 13
  PattNo  Description
     29  IMPACT
     59  EXP_WHY_BEN
     62  DISP_BEN
     71  conc_ben
     77  Exp_gap
    103  none(M_NEED)
    104  a little (M_NEED)
    105  some(M_NEED)
    182  major improvement(IMPACT)
    183  some improvement(IMPACT)
    184  a little improvement(IMPACT)
    185  no improvement(IMPACT)
    236  <40
```

REQ_PERF stabilizes in Cluster 41 around the concept of DOM_RSK_F. The former is therefore assigned as a child of the latter.

```
Cluster No.: 41
Properties of newly merged group:
Number Rules in Group = 8
  RuleNo  Description
     34  MDL=Y_ALG  ->  DOM_RSK_F=1
     61  DOM_RSK_F<DISP_DOM_R_F+1000  ->  conc_dom_rsk
     35  MDL=Y_MAN  ->  DOM_RSK_F=5
     36  MDL=N  ->  DOM_RSK_F=10
     51  ORG_RSK_F,USER_RSK_F,DOM_RSK_F  ->  EST_RSK(ORG_RSK_F,USER_RSK_F,DOM_RSK_F)
     37  REQ_PERF=50  ->  DOM_RSK_F(+4)
     38  REQ_PERF=80  ->  DOM_RSK_F(+6)
     39  REQ_PERF=100  ->  DOM_RSK_F(+10)
Number Stable Patterns in Group = 10
  PattNo  Description
      5  MDL
      7  REQ_PERF
     64  DISP_DOM_R_F
     73  conc_dom_risk
    117  yes, an algorithmic system(MDL)
```

```
118  yes, a manual system(MDL)
119  no(MDL)
125  50% as good as senior experts(REQ_PERF)
126  80% as good as senior experts(REQ_PERF)
127  as good as senior experts(REQ_PERF)
```

ADV stabilizes in Cluster 39 around the concept of ORG_RSK_F, thus giving us another child-parent relationship.

```
Cluster No.: 39
Properties of newly merged group:
Number Rules in Group = 4
  RuleNo  Description
      30  ADV=Y  ->  ORG_RSK_F(+0)(+10)
      49  MGMT_LV=3  ->  ORG_RSK_F(+2)
      60  ORG_RSK_F<DISP_ORG_RSK+1000  ->  conc_org_rsk
      50  MGMT_LV>3  ->  ORG_RSK_F(+5)
Number Stable Patterns in Group = 6
  PattNo  Description
      14  ADV
      63  DISP_ORG_RSK
      72  conc_org_risk
     147  yes(ADV)
     155  three(MGMT_LV)
     156  more than three(MGMT_LV)
```

COMP_PROF stabilizes in Cluster 35 around the dominating concept of USER_RSK_F. Hence, the former is assigned to be a child of the latter.

```
Cluster No.: 35
Properties of newly merged group:
Number Rules in Group = 9
  RuleNo  Description
      41  USER_ENTH=NN  ->  USER_RSK_F=10
      62  USER_RSK_F<DISP_USER_R_F+1000  ->  conc_user_rsk
      42  USER_ENTH=LTL  ->  USER_RSK_F=6
      43  USER_ENTH=SM  ->  USER_RSK_F=2
      44  USER_ENTH=LOT  ->  USER_RSK_F=0
      45  COMP_PROF=NN  ->  USER_RSK_F(+10)
      46  COMP_PROF=LTL  ->  USER_RSK_F(+6)
      47  COMP_PROF=SM  ->  USER_RSK_F(+2)
```

```
     48  COMP_PROF=LOT  ->  USER_RSK_F(+0)
Number Stable Patterns in Group = 13
  PattNo  Description
      18  USER_ENTH
      19  COMP_PROF
      65  DISP_USER_R_F
      74  conc_user_risk
     157  none(USER_ENTH)
     158  a little(USER_ENTH)
     159  some(USER_ENTH)
     160  a lot(USER_ENTH)
     161  none(COMP_PROF)
     162  a little(COMP_PROF)
     163  some(COMP_PROF)
     164  a lot(COMP_PROF)
     232  10
```

In the next iteration, MAJ_IMPR stabilizes around the concept of EST_BEN.

```
Cluster No.: 34
Properties of newly merged group:
Number Rules in Group = 16
  RuleNo  Description
       1  ID_NEED=Y  ->  EST_BEN=10
      59  DISP_BEN+1000>EST_BEN  ->  conc_ben
      52  EST_BEN>-1,EST_RSK>-1  ->  RATIO(EST_BEN/EST_RSK)
       2  M_NEED=NN  ->  EST_BEN(+0)
      53  EST_BEN<40,EXP_WHY_BEN  ->  exp_gap
       3  M_NEED=LIT  ->  EST_BEN(+2)
       4  M_NEED=SM  ->  EST_BEN(+6)
       5  M_NEED=LOT  ->  EST_BEN(+10)
       6  IMPACT=NN  ->  EST_BEN(+0)
       7  IMPACT=LIT  ->  EST_BEN(+2)
       8  IMPACT=SM  ->  EST_BEN(+6)
       9  IMPACT=MAJ  ->  EST_BEN(+10)
      10  MAJ_IMPR=NN  ->  EST_BEN(+0)
      11  MAJ_IMPR=LIT  ->  EST_BEN(+2)
      12  MAJ_IMPR=SM  ->  EST_BEN(+6)
      13  MAJ_IMPR=LOT  ->  EST_BEN(+10)
Number Stable Patterns in Group = 18
  PattNo  Description
      29  IMPACT
      36  MAJ_IMPR
```

```
 59  EXP_WHY_BEN
 62  DISP_BEN
 71  conc_ben
 77  Exp_gap
103  none(M_NEED)
104  a little (M_NEED)
105  some(M_NEED)
182  major improvement(IMPACT)
183  some improvement(IMPACT)
184  a little improvement(IMPACT)
185  no improvement(IMPACT)
201  none(MAJ_IMPROV)
202  a little(MAJ_IMPROV)
203  some(MAJ_IMPROV)
204  a lot(MAJ_IMPROV)
236  <40
```

After a few mergings, MGMT_LV stabilizes around the concept of ORG_RSK and provides another child-parent relationship.

```
Cluster No.: 25
Properties of newly merged group:
Number Rules in Group = 5
  RuleNo  Description
      30  ADV=Y  ->  ORG_RSK_F(+0)(+10)
      49  MGMT_LV=3  ->  ORG_RSK_F(+2)
      60  ORG_RSK_F<DISP_ORG_RSK+1000  ->  conc_org_rsk
      50  MGMT_LV>3  ->  ORG_RSK_F(+5)
      32  MGMT_LV=1V2  ->  ORG_RSK_F(+0)
Number Stable Patterns in Group = 9
  PattNo  Description
      14  ADV
      17  MGMT_LV
      63  DISP_ORG_RSK
      72  conc_org_risk
     147  yes(ADV)
     153  one(MGMT_LV)
     154  two(MGMT_LV)
     155  three(MGMT_LV)
     156  more than three(MGMT_LV)
```

TRNG_TOOL in Cluster 24 stabilizes around the concept of EST_BEN. Thus, EST_BEN is a parent of TRNG_TOOL.

```
Cluster No.: 24
Properties of newly merged group:
Number Rules in Group = 25
  RuleNo  Description
       1  ID_NEED=Y  ->  EST_BEN=10
      59  DISP_BEN+1000>EST_BEN  ->  conc_ben
      52  EST_BEN>-1,EST_RSK>-1  ->  RATIO(EST_BEN/EST_RSK)
       2  M_NEED=NN  ->  EST_BEN(+0)
      53  EST_BEN<40,EXP_WHY_BEN  ->  exp_gap
       3  M_NEED=LIT  ->  EST_BEN(+2)
       4  M_NEED=SM  ->  EST_BEN(+6)
       5  M_NEED=LOT  ->  EST_BEN(+10)
       6  IMPACT=NN  ->  EST_BEN(+0)
       7  IMPACT=LIT  ->  EST_BEN(+2)
       8  IMPACT=SM  ->  EST_BEN(+6)
       9  IMPACT=MAJ  ->  EST_BEN(+10)
      10  MAJ_IMPR=NN  ->  EST_BEN(+0)
      11  MAJ_IMPR=LIT  ->  EST_BEN(+2)
      12  MAJ_IMPR=SM  ->  EST_BEN(+6)
      13  MAJ_IMPR=LOT  ->  EST_BEN(+10)
      14  COMPL=HG  ->  EST_BEN(+10)
      15  COMPL=MED  ->  EST_BEN(+6)
      16  COMPL=LOW  ->  EST_BEN(+2)
      17  COMPL=NN  ->  EST_BEN(+0)
      18  TRNG_TOOL=LOT  ->  EST_BEN(+10)
      19  TRNG_TOOL=SM  ->  EST_BEN(+5)
      21  END_USER=CL  ->  EST_BEN(+4)
      23  END_USER=ADM  ->  EST_BEN(+10)
      20  TRNG_TOOL=NNVLIT  ->  EST_BEN(+0)
Number Stable Patterns in Group = 28
  PattNo  Description
      29  IMPACT
      36  MAJ_IMPR
      37  TRNG_TOOL
      59  EXP_WHY_BEN
      62  DISP_BEN
      71  conc_ben
      77  Exp_gap
     103  none(M_NEED)
     104  a little (M_NEED)
     105  some(M_NEED)
     108  clerical(END_USER)
     111  administrative(END_USER)
     140  medium(COMPL)
     142  low(COMPL)
```

```
143  none(COMPL)
182  major improvement(IMPACT)
183  some improvement(IMPACT)
184  a little improvement(IMPACT)
185  no improvement(IMPACT)
201  none(MAJ_IMPROV)
202  a little(MAJ_IMPROV)
203  some(MAJ_IMPROV)
204  a lot(MAJ_IMPROV)
205  none(TRNG_TOOL)
206  a little(TRNG_TOOL)
207  some(TRNG_TOOL)
208  a lot(TRNG_TOOL)
236  <40
```

In the next merge, END_USER stablizes around the concept of EST_BEN.

```
Cluster No.: 23
Properties of newly merged group:
Number Rules in Group = 26
  RuleNo  Description
       1  ID_NEED=Y  ->  EST_BEN=10
      59  DISP_BEN+1000>EST_BEN  ->  conc_ben
      52  EST_BEN>-1,EST_RSK>-1  ->  RATIO(EST_BEN/EST_RSK)
       2  M_NEED=NN  ->  EST_BEN(+0)
      53  EST_BEN<40,EXP_WHY_BEN  ->  exp_gap
       3  M_NEED=LIT  ->  EST_BEN(+2)
       4  M_NEED=SM  ->  EST_BEN(+6)
       5  M_NEED=LOT  ->  EST_BEN(+10)
       6  IMPACT=NN  ->  EST_BEN(+0)
       7  IMPACT=LIT  ->  EST_BEN(+2)
       8  IMPACT=SM  ->  EST_BEN(+6)
       9  IMPACT=MAJ  ->  EST_BEN(+10)
      10  MAJ_IMPR=NN  ->  EST_BEN(+0)
      11  MAJ_IMPR=LIT  ->  EST_BEN(+2)
      12  MAJ_IMPR=SM  ->  EST_BEN(+6)
      13  MAJ_IMPR=LOT  ->  EST_BEN(+10)
      14  COMPL=HG  ->  EST_BEN(+10)
      15  COMPL=MED  ->  EST_BEN(+6)
      16  COMPL=LOW  ->  EST_BEN(+2)
      17  COMPL=NN  ->  EST_BEN(+0)
      18  TRNG_TOOL=LOT  ->  EST_BEN(+10)
      19  TRNG_TOOL=SM  ->  EST_BEN(+5)
```

```
      21  END_USER=CL  ->  EST_BEN(+4)
      23  END_USER=ADM  ->  EST_BEN(+10)
      20  TRNG_TOOL=NNVLIT  ->  EST_BEN(+0)
      22  END_USER=TECHVPRO  ->  EST_BEN(+8)
Number Stable Patterns in Group = 32
  PattNo  Description
       3  END_USER
      29  IMPACT
      36  MAJ_IMPR
      37  TRNG_TOOL
      59  EXP_WHY_BEN
      62  DISP_BEN
      71  conc_ben
      77  Exp_gap
     103  none(M_NEED)
     104  a little (M_NEED)
     105  some(M_NEED)
     108  clerical(END_USER)
     109  technicians(END_USER)
     110  professional(END_USER)
     111  administrative(END_USER)
     140  medium(COMPL)
     142  low(COMPL)
     143  none(COMPL)
     182  major improvement(IMPACT)
     183  some improvement(IMPACT)
     184  a little improvement(IMPACT)
     185  no improvement(IMPACT)
     201  none(MAJ_IMPROV)
     202  a little(MAJ_IMPROV)
     203  some(MAJ_IMPROV)
     204  a lot(MAJ_IMPROV)
     205  none(TRNG_TOOL)
     206  a little(TRNG_TOOL)
     207  some(TRNG_TOOL)
     208  a lot(TRNG_TOOL)
     230  + 8
     236  <40
```

Both INTER and DOM_RSK_F stabilize in Cluster 21 and now the subtree with DOM_RSK_F as the parent can be completed. This is the first variable to stabilize; hence, this cluster is important as it contains all the information on the qualifiers that are dependent on this variable.

```
Cluster No.: 21
Properties of newly merged group:
Number Rules in Group = 9
  RuleNo  Description
      34  MDL=Y_ALG  ->  DOM_RSK_F=1
      61  DOM_RSK_F<DISP_DOM_R_F+1000  ->  conc_dom_rsk
      35  MDL=Y_MAN  ->  DOM_RSK_F=5
      36  MDL=N  ->  DOM_RSK_F=10
      51  ORG_RSK_F,USER_RSK_F,DOM_RSK_F  ->  EST_RSK(ORG_RSK_F,USER_RSK_F,DOM_RSK_F)
      37  REQ_PERF=50  ->  DOM_RSK_F(+4)
      38  REQ_PERF=80  ->  DOM_RSK_F(+6)
      39  REQ_PERF=100  ->  DOM_RSK_F(+10)
      40  INTER=Y  ->  DOM_RSK_F(+10)(+0)
Number Stable Patterns in Group = 13
  PattNo  Description
       5  MDL
       7  REQ_PERF
      13  INTER
      57  DOM_RSK_F
      64  DISP_DOM_R_F
      73  conc_dom_risk
     117  yes, an algorithmic system(MDL)
     118  yes, a manual system(MDL)
     119  no(MDL)
     125  50% as good as senior experts(REQ_PERF)
     126  80% as good as senior experts(REQ_PERF)
     127  as good as senior experts(REQ_PERF)
     145  yes(INTER)
```

MGMT_SUP stabilizes in Cluster 20, with ORG_RSK_F as the dominating concept.

```
Cluster No.: 20
Properties of newly merged group:
Number Rules in Group = 6
  RuleNo  Description
      30  ADV=Y  ->  ORG_RSK_F(+0)(+10)
      49  MGMT_LV=3  ->  ORG_RSK_F(+2)
      60  ORG_RSK_F<DISP_ORG_RSK+1000  ->  conc_org_rsk
      50  MGMT_LV>3  ->  ORG_RSK_F(+5)
      32  MGMT_LV=1V2  ->  ORG_RSK_F(+0)
      31  MGMT_SUP=Y  ->  ORG_RSK_F(+0)(+10)
```

```
Number Stable Patterns in Group = 11
  PattNo  Description
      14  ADV
      15  MGMT_SUP
      17  MGMT_LV
      63  DISP_ORG_RSK
      72  conc_org_risk
     147  yes(ADV)
     149  yes(MGMT_SUP)
     153  one(MGMT_LV)
     154  two(MGMT_LV)
     155  three(MGMT_LV)
     156  more than three(MGMT_LV)
```

PRBLM, PROF_ AVL, and EXP_STF all stabilize around a subsidiary concept of ID_NEED, showing that these are finer concepts associated with ID_NEED. Cluster 18 highlights this aspect of ID_NEED:

```
Cluster No.: 18
Properties of newly merged group:
Number Rules in Group = 3
  RuleNo  Description
      66  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y,EXP_STF=Y  ->  ID_NEED=Y
      67  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y  ->  ID_NEED=Y
      68  ID_NEED=U,PRBLM=CMPLX,EXP_STF=Y  ->  ID_NEED=Y
Number Stable Patterns in Group = 7
  PattNo  Description
      22  PRBLM
      23  PROF_~AVL
      24  EXP_STF
     102  unknown(ID_NEED)
     168  complex(PRBLM)
     170  yes(PROF_~AVL)
     172  yes(EXP_STF)
```

MGMT_EXP stabilizes in Cluster 17 around the concept of ORG_RSK_F:

```
Cluster No.: 17
Properties of newly merged group:
Number Rules in Group = 7
  RuleNo  Description
      30  ADV=Y  ->  ORG_RSK_F(+0)(+10)
      49  MGMT_LV=3  ->  ORG_RSK_F(+2)
```

```
      60  ORG_RSK_F<DISP_ORG_RSK+1000  ->  conc_org_rsk
      50  MGMT_LV>3  ->  ORG_RSK_F(+5)
      32  MGMT_LV=1V2  ->  ORG_RSK_F(+0)
      31  MGMT_SUP=Y  ->  ORG_RSK_F(+0)(+10)
      33  MGMT_EXP=Y  ->  ORG_RSK_F(+0)(+10)
Number Stable Patterns in Group = 13
  PattNo  Description
      14  ADV
      15  MGMT_SUP
      16  MGMT_EXP
      17  MGMT_LV
      63  DISP_ORG_RSK
      72  conc_org_risk
     147  yes(ADV)
     149  yes(MGMT_SUP)
     151  yes(MGMT_EXP)
     153  one(MGMT_LV)
     154  two(MGMT_LV)
     155  three(MGMT_LV)
     156  more than three(MGMT_LV)
```

USER_RSK_F stabilizes in Cluster 13, providing the second subtree, with variable USER_RSK_F as the parent.

```
Cluster No.: 13
Properties of newly merged group:
Number Rules in Group = 18
  RuleNo  Description
      34  MDL=Y_ALG  ->  DOM_RSK_F=1
      61  DOM_RSK_F<DISP_DOM_R_F+1000  ->  conc_dom_rsk
      35  MDL=Y_MAN  ->  DOM_RSK_F=5
      36  MDL=N  ->  DOM_RSK_F=10
      51  ORG_RSK_F,USER_RSK_F,DOM_RSK_F  ->  EST_RSK(ORG_RSK_F,USER_RSK_F,DOM_RSK_F)
      37  REQ_PERF=50  ->  DOM_RSK_F(+4)
      38  REQ_PERF=80  ->  DOM_RSK_F(+6)
      39  REQ_PERF=100  ->  DOM_RSK_F(+10)
      40  INTER=Y  ->  DOM_RSK_F(+10)(+0)
      41  USER_ENTH=NN  ->  USER_RSK_F=10
      62  USER_RSK_F<DISP_USER_R_F+1000  ->  conc_user_rsk
      42  USER_ENTH=LTL  ->  USER_RSK_F=6
      43  USER_ENTH=SM  ->  USER_RSK_F=2
      44  USER_ENTH=LOT  ->  USER_RSK_F=0
      45  COMP_PROF=NN  ->  USER_RSK_F(+10)
```

```
     46  COMP_PROF=LTL  ->  USER_RSK_F(+6)
     47  COMP_PROF=SM  ->  USER_RSK_F(+2)
     48  COMP_PROF=LOT  ->  USER_RSK_F(+0)
Number Stable Patterns in Group = 27
  PattNo  Description
       5  MDL
       7  REQ_PERF
      13  INTER
      18  USER_ENTH
      19  COMP_PROF
      56  USER_RSK_F
      57  DOM_RSK_F
      64  DISP_DOM_R_F
      65  DISP_USER_R_F
      73  conc_dom_risk
      74  conc_user_risk
     117  yes, an algorithmic system(MDL)
     118  yes, a manual system(MDL)
     119  no(MDL)
     125  50% as good as senior experts(REQ_PERF)
     126  80% as good as senior experts(REQ_PERF)
     127  as good as senior experts(REQ_PERF)
     145  yes(INTER)
     157  none(USER_ENTH)
     158  a little(USER_ENTH)
     159  some(USER_ENTH)
     160  a lot(USER_ENTH)
     161  none(COMP_PROF)
     162  a little(COMP_PROF)
     163  some(COMP_PROF)
     164  a lot(COMP_PROF)
     232  10
```

The next variable to stabilize is ORG_RSK_F in Cluster 12. Rule 51 is especially important in this group because it brings all three risk factors together to calculate an estimated risk. Since three variables have stabilized in this cluster, the variable EST_RSK that connects these three variables should be the parent. However, the parent of EST_RSK is unclear at this point and will be decided only after a few more merges make this variable stable.

```
Cluster No.: 12
Properties of newly merged group:
Number Rules in Group = 25
  RuleNo  Description
      30  ADV=Y  ->  ORG_RSK_F(+0)(+10)
      49  MGMT_LV=3  ->  ORG_RSK_F(+2)
      60  ORG_RSK_F<DISP_ORG_RSK+1000  ->  conc_org_rsk
      50  MGMT_LV>3  ->  ORG_RSK_F(+5)
      32  MGMT_LV=1V2  ->  ORG_RSK_F(+0)
      31  MGMT_SUP=Y  ->  ORG_RSK_F(+0)(+10)
      33  MGMT_EXP=Y  ->  ORG_RSK_F(+0)(+10)
      34  MDL=Y_ALG  ->  DOM_RSK_F=1
      61  DOM_RSK_F<DISP_DOM_R_F+1000  ->  conc_dom_rsk
      35  MDL=Y_MAN  ->  DOM_RSK_F=5
      36  MDL=N  ->  DOM_RSK_F=10
      51  ORG_RSK_F,USER_RSK_F,DOM_RSK_F  ->  EST_RSK(ORG_RSK_F,USER_RSK_F,DOM_RSK_F)
      37  REQ_PERF=50  ->  DOM_RSK_F(+4)
      38  REQ_PERF=80  ->  DOM_RSK_F(+6)
      39  REQ_PERF=100  ->  DOM_RSK_F(+10)
      40  INTER=Y  ->  DOM_RSK_F(+10)(+0)
      41  USER_ENTH=NN  ->  USER_RSK_F=10
      62  USER_RSK_F<DISP_USER_R_F+1000  ->  conc_user_rsk
      42  USER_ENTH=LTL  ->  USER_RSK_F=6
      43  USER_ENTH=SM  ->  USER_RSK_F=2
      44  USER_ENTH=LOT  ->  USER_RSK_F=0
      45  COMP_PROF=NN  ->  USER_RSK_F(+10)
      46  COMP_PROF=LTL  ->  USER_RSK_F(+6)
      47  COMP_PROF=SM  ->  USER_RSK_F(+2)
      48  COMP_PROF=LOT  ->  USER_RSK_F(+0)
Number Stable Patterns in Group = 41
  PattNo  Description
       5  MDL
       7  REQ_PERF
      13  INTER
      14  ADV
      15  MGMT_SUP
      16  MGMT_EXP
      17  MGMT_LV
      18  USER_ENTH
      19  COMP_PROF
      55  ORG_RSK_F
      56  USER_RSK_F
      57  DOM_RSK_F
      63  DISP_ORG_RSK
      64  DISP_DOM_R_F
```

```
 65  DISP_USER_R_F
 72  conc_org_risk
 73  conc_dom_risk
 74  conc_user_risk
117  yes, an algorithmic system(MDL)
118  yes, a manual system(MDL)
119  no(MDL)
125  50% as good as senior experts(REQ_PERF)
126  80% as good as senior experts(REQ_PERF)
127  as good as senior experts(REQ_PERF)
145  yes(INTER)
147  yes(ADV)
149  yes(MGMT_SUP)
151  yes(MGMT_EXP)
153  one(MGMT_LV)
154  two(MGMT_LV)
155  three(MGMT_LV)
156  more than three(MGMT_LV)
157  none(USER_ENTH)
158  a little(USER_ENTH)
159  some(USER_ENTH)
160  a lot(USER_ENTH)
161  none(COMP_PROF)
162  a little(COMP_PROF)
163  some(COMP_PROF)
164  a lot(COMP_PROF)
232  10
```

Finally, Cluster 11 is formed by the merging of two major clusters — one with the primary concept around EST_BEN and the other with the parent EST_RSK. At this point, the role of Rule 52 for calculating the RATIO of EST_BEN and EST_RSK can be noted. This provides the parent nodes for the various subtrees as shown in figure 2.

```
Cluster No.: 11
Properties of newly merged group:
Number Rules in Group = 55
  RuleNo  Description
      1   ID_NEED=Y  ->  EST_BEN=10
     59   DISP_BEN+1000>EST_BEN  ->  conc_ben
     52   EST_BEN>-1,EST_RSK>-1  ->  RATIO(EST_BEN/EST_RSK)
      2   M_NEED=NN  ->  EST_BEN(+0)
     53   EST_BEN<40,EXP_WHY_BEN  ->  exp_gap
      3   M_NEED=LIT  ->  EST_BEN(+2)
```

```
4  M_NEED=SM  ->  EST_BEN(+6)
5  M_NEED=LOT  ->  EST_BEN(+10)
6  IMPACT=NN  ->  EST_BEN(+0)
7  IMPACT=LIT  ->  EST_BEN(+2)
8  IMPACT=SM  ->  EST_BEN(+6)
9  IMPACT=MAJ  ->  EST_BEN(+10)
10 MAJ_IMPR=NN  ->  EST_BEN(+0)
11 MAJ_IMPR=LIT  ->  EST_BEN(+2)
12 MAJ_IMPR=SM  ->  EST_BEN(+6)
13 MAJ_IMPR=LOT  ->  EST_BEN(+10)
14 COMPL=HG  ->  EST_BEN(+10)
15 COMPL=MED  ->  EST_BEN(+6)
16 COMPL=LOW  ->  EST_BEN(+2)
17 COMPL=NN  ->  EST_BEN(+0)
18 TRNG_TOOL=LOT  ->  EST_BEN(+10)
19 TRNG_TOOL=SM  ->  EST_BEN(+5)
21 END_USER=CL  ->  EST_BEN(+4)
23 END_USER=ADM  ->  EST_BEN(+10)
20 TRNG_TOOL=NNVLIT  ->  EST_BEN(+0)
22 END_USER=TECHVPRO  ->  EST_BEN(+8)
25 YRLY_ANAL_EXEC_T>1000  ->  EST_BEN(+10)(0)
56 RATIO>1  ->  go_ahd,LK_GOOD
54 EST_RSK>40,EXP_EST_RSK  ->  exp_sz
55 RATIO<1,EXP_LOWRATIO  ->  exp_low
30 ADV=Y  ->  ORG_RSK_F(+0)(+10)
49 MGMT_LV=3  ->  ORG_RSK_F(+2)
60 ORG_RSK_F<DISP_ORG_RSK+1000  ->  conc_org_rsk
50 MGMT_LV>3  ->  ORG_RSK_F(+5)
32 MGMT_LV=1V2  ->  ORG_RSK_F(+0)
31 MGMT_SUP=Y  ->  ORG_RSK_F(+0)(+10)
33 MGMT_EXP=Y  ->  ORG_RSK_F(+0)(+10)
34 MDL=Y_ALG  ->  DOM_RSK_F=1
61 DOM_RSK_F<DISP_DOM_R_F+1000  ->  conc_dom_rsk
35 MDL=Y_MAN  ->  DOM_RSK_F=5
36 MDL=N  ->  DOM_RSK_F=10
51 ORG_RSK_F,USER_RSK_F,DOM_RSK_F  ->  EST_RSK(ORG_RSK_F,USER_RSK_F,DOM_RSK_F)
37 REQ_PERF=50  ->  DOM_RSK_F(+4)
38 REQ_PERF=80  ->  DOM_RSK_F(+6)
39 REQ_PERF=100  ->  DOM_RSK_F(+10)
40 INTER=Y  ->  DOM_RSK_F(+10)(+0)
41 USER_ENTH=NN  ->  USER_RSK_F=10
62 USER_RSK_F<DISP_USER_R_F+1000  ->  conc_user_rsk
42 USER_ENTH=LTL  ->  USER_RSK_F=6
43 USER_ENTH=SM  ->  USER_RSK_F=2
44 USER_ENTH=LOT  ->  USER_RSK_F=0
```

```
     45  COMP_PROF=NN  ->  USER_RSK_F(+10)
     46  COMP_PROF=LTL  ->  USER_RSK_F(+6)
     47  COMP_PROF=SM  ->  USER_RSK_F(+2)
     48  COMP_PROF=LOT  ->  USER_RSK_F(+0)
Number Stable Patterns in Group = 94
  PattNo  Description
       3  END_USER
       5  MDL
       7  REQ_PERF
      13  INTER
      14  ADV
      15  MGMT_SUP
      16  MGMT_EXP
      17  MGMT_LV
      18  USER_ENTH
      19  COMP_PROF
      29  IMPACT
      36  MAJ_IMPR
      37  TRNG_TOOL
      54  EST_RSK
      55  ORG_RSK_F
      56  USER_RSK_F
      57  DOM_RSK_F
      58  RATIO
      59  EXP_WHY_BEN
      60  EXP_LOWRATIO
      61  EXP_EST_RSK
      62  DISP_BEN
      63  DISP_ORG_RSK
      64  DISP_DOM_R_F
      65  DISP_USER_R_F
      66  LK GOOD
      71  conc_ben
      72  conc_org_risk
      73  conc_dom_risk
      74  conc_user_risk
      75  Go ahd
      77  Exp_gap
      78  Exp_low
      79  Exp_sz
     103  none(M_NEED)
     104  a little (M_NEED)
     105  some(M_NEED)
     108  clerical(END_USER)
     109  technicians(END_USER)
```

```
110  professional(END_USER)
111  administrative(END_USER)
117  yes, an algorithmic system(MDL)
118  yes, a manual system(MDL)
119  no(MDL)
125  50% as good as senior experts(REQ_PERF)
126  80% as good as senior experts(REQ_PERF)
127  as good as senior experts(REQ_PERF)
140  medium(COMPL)
142  low(COMPL)
143  none(COMPL)
145  yes(INTER)
147  yes(ADV)
149  yes(MGMT_SUP)
151  yes(MGMT_EXP)
153  one(MGMT_LV)
154  two(MGMT_LV)
155  three(MGMT_LV)
156  more than three(MGMT_LV)
157  none(USER_ENTH)
158  a little(USER_ENTH)
159  some(USER_ENTH)
160  a lot(USER_ENTH)
161  none(COMP_PROF)
162  a little(COMP_PROF)
163  some(COMP_PROF)
164  a lot(COMP_PROF)
182  major improvement(IMPACT)
183  some improvement(IMPACT)
184  a little improvement(IMPACT)
185  no improvement(IMPACT)
201  none(MAJ_IMPROV)
202  a little(MAJ_IMPROV)
203  some(MAJ_IMPROV)
204  a lot(MAJ_IMPROV)
205  none(TRNG_TOOL)
206  a little(TRNG_TOOL)
207  some(TRNG_TOOL)
208  a lot(TRNG_TOOL)
221  > 1
223  + 0
224  + 10
225  > 1000
226  + 2
227  + 6
```

```
228   + 5
229   + 4
230   + 8
232   10
234   <> ""
235   < 1
236   <40
237   > 40
238   > 1
241   + 1000
```

At this point, the groups become too large and mining for any additional useful information now becomes a tedious process. Qualifiers that stablize later on in the clustering process (at cluster 1, 2, and 3) are ID_NEED, M_NEED, COMPL, EXP_AVL, RES, END_USERS, REG, EXPED_COMPL, and EXP_LVL. More insight into their dependency information is obtained by clustering the knowledge base through the antecedent and consequent metric.

Antecedent Metric

The antecedent metric clusters rules based on common information across antecedents of the rules. Patterns that stabilize in the last stages of mergings through the data-flow metric have a chance to stabilize earlier through this metric to reveal important *dependency information.*

This metric also reveals information on *incompleteness in qualifier value specifications,* because every time a qualifier stabilizes in a cluster, it can easily be seen if all the associated values with it have also stabilized. If not, an incompleteness condition can be flagged, signifying that if the unspecified value were to appear in the data base, this knowledge base would have no rule to handle this.

Conflicting and redundant conditions are also detected easily through this metric when rules with structural similarity are juxtaposed through clustering.

The three objectives will be discussed in the order listed above.

Dependency Information Among Qualifiers and Variables

COMPL stabilizes in Cluster 29 and REG and EXPED_COMPL also stabilizes with it. This reveals that the last two qualifiers feed into the concept of setting the value for COMPL.

```
Cluster No.: 29
Properties of newly merged group:
Number Rules in Group = 5
  RuleNo  Description
      14  COMPL=HG  ->  EST_BEN(+10)
      15  COMPL=MED  ->  EST_BEN(+6)
      16  COMPL=LOW  ->  EST_BEN(+2)
      17  COMPL=NN  ->  EST_BEN(+0)
      65  COMPL=U,REG=N,EXPED_COMPL=Y  ->  COMPL=H
Number Stable Patterns in Group = 10
  PattNo  Description
      12  COMPL
      32  REG
      33  EXPED_COMPL
     139  high(COMPL)
     140  medium(COMPL)
     142  low(COMPL)
     143  none(COMPL)
     144  unknown(COMPL)
     194  no(REG)
     195  yes(EXPED_COMPL)
```

ID_NEED stabilizes in Cluster 27 and reveals its dependent qualifiers in the following group. Since PRBLM, PROF$\tilde{A}$VL, and EXP_STF also stabilizes with ID_NEED for the first time, it is obvious that these qualifiers are used in setting the value of ID_NEED.

```
Cluster No.: 27
Properties of newly merged group:
Number Rules in Group = 5
  RuleNo  Description
       1  ID_NEED=Y  ->  EST_BEN=10
      63  ID_NEED=N  ->  EST_BEN=0
      66  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y,EXP_STF=Y  ->  ID_NEED=Y
      67  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y  ->  ID_NEED=Y
      68  ID_NEED=U,PRBLM=CMPLX,EXP_STF=Y  ->  ID_NEED=Y
Number Stable Patterns in Group = 11
  PattNo  Description
       1  ID_NEED
      22  PRBLM
      23  PROF_~AVL
      24  EXP_STF
     100  yes(ID_NEED)
     101  no(ID_NEED)
```

```
    102  unknown(ID_NEED)
    168  complex(PRBLM)
    170  yes(PROF_~AVL)
    172  yes(EXP_STF)
    242  0
```

Next, the qualifier M_NEED stabilizes in Cluster 25 with EXP_AVL, RES, END_USERS, and EXP_LVL; the latter four are used in setting M_NEED.

```
Cluster No.: 25
Properties of newly merged group:
Number Rules in Group = 5
  RuleNo  Description
       2  M_NEED=NN -> EST_BEN(+0)
       3  M_NEED=LIT -> EST_BEN(+2)
       4  M_NEED=SM -> EST_BEN(+6)
       5  M_NEED=LOT -> EST_BEN(+10)
      64  M_NEED=U,EXP_AVL,EXP_LVL,RES,END_USERS -> M_NEED=LOT
Number Stable Patterns in Group = 14
  PattNo  Description
       2  M_NEED
      25  EXP_AVL
      27  RES
      28  END_USERS
      35  EXP_LVL
     103  none(M_NEED)
     104  a little (M_NEED)
     105  some(M_NEED)
     106  a lot(M_NEED)
     107  unknown(M_NEED)
     174  available(EXP_AVL)
     178  yes(RES)
     180  receptive(END_USERS)
     199  yes(EXP_LVL)
```

We can now complete the entire tree for ESAA, displaying the relationships of the various qualifiers to key variables, as shown in figure 2.

Incompleteness in Qualifier-Value Specifications

This section will address the issue of incompleteness in qualifier-value specifications. It will cover the important merges that have helped reveal this type of information.

In Cluster 66, the qualifier PROF_ÃVL stabilizes. The Qualifier section declares two associated values - "yes" and "no" - for PROF_ÃVL. As can be seen, the rule base addresses only the former and has no action for the latter value.

```
Cluster No.: 66
Properties of newly merged group:
Number Rules in Group = 2
  RuleNo  Description
      66  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y,EXP_STF=Y -> ID_NEED=Y
      67  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y -> ID_NEED=Y
Number Stable Patterns in Group = 2
  PattNo  Description
      23  PROF_~AVL
     170  yes(PROF_~AVL)
```

The next qualifier to stabilize in Cluster 65 is PRBLM and EXP_STF. Again, only the "complex" value for PRBLM and "yes" for EXP_STF are addressed. However, EXP_STF can take on the value of "no" and PRBLM can have the value "not complex" as declared in the Qualifier section.

```
Cluster No.: 65
Properties of newly merged group:
Number Rules in Group = 3
  RuleNo  Description
      66  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y,EXP_STF=Y -> ID_NEED=Y
      67  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y -> ID_NEED=Y
      68  ID_NEED=U,PRBLM=CMPLX,EXP_STF=Y -> ID_NEED=Y
Number Stable Patterns in Group = 7
  PattNo  Description
      22  PRBLM
      23  PROF_~AVL
      24  EXP_STF
     102  unknown(ID_NEED)
     168  complex(PRBLM)
     170  yes(PROF_~AVL)
     172  yes(EXP_STF)
```

IMPACT stabilizes in Cluster 58 and all values for this qualifier have been addressed in the rule base.

```
Cluster No.: 58
Properties of newly merged group:
Number Rules in Group = 4
  RuleNo  Description
       6  IMPACT=NN -> EST_BEN(+0)
       7  IMPACT=LIT -> EST_BEN(+2)
       8  IMPACT=SM -> EST_BEN(+6)
       9  IMPACT=MAJ -> EST_BEN(+10)
Number Stable Patterns in Group = 5
  PattNo  Description
      29  IMPACT
     182  major improvement(IMPACT)
     183  some improvement(IMPACT)
     184  a little improvement(IMPACT)
     185  no improvement(IMPACT)
```

MAJ_IMPR stabilizes in Cluster 55 and is completely specified.

```
Cluster No.: 55
Properties of newly merged group:
Number Rules in Group = 4
  RuleNo  Description
      10  MAJ_IMPR=NN -> EST_BEN(+0)
      11  MAJ_IMPR=LIT -> EST_BEN(+2)
      12  MAJ_IMPR=SM -> EST_BEN(+6)
      13  MAJ_IMPR=LOT -> EST_BEN(+10)
Number Stable Patterns in Group = 5
  PattNo  Description
      36  MAJ_IMPR
     201  none(MAJ_IMPROV)
     202  a little(MAJ_IMPROV)
     203  some(MAJ_IMPROV)
     204  a lot(MAJ_IMPROV)
```

MDL stabilizes in Cluster 47 and addresses all its specified values, too.

```
Cluster No.: 47
Properties of newly merged group:
Number Rules in Group = 3
  RuleNo  Description
      34  MDL=Y_ALG -> DOM_RSK_F=1
      35  MDL=Y_MAN -> DOM_RSK_F=5
```

```
      36  MDL=N -> DOM_RSK_F=10
Number Stable Patterns in Group = 4
  PattNo  Description
       5  MDL
     117  yes, an algorithmic system(MDL)
     118  yes, a manual system(MDL)
     119  no(MDL)
```

REQ_PERF also passes the completeness inspection in Cluster 45:

```
Cluster No.: 45
Properties of newly merged group:
Number Rules in Group = 3
  RuleNo  Description
      37  REQ_PERF=50 -> DOM_RSK_F(+4)
      38  REQ_PERF=80 -> DOM_RSK_F(+6)
      39  REQ_PERF=100 -> DOM_RSK_F(+10)
Number Stable Patterns in Group = 4
  PattNo  Description
       7  REQ_PERF
     125  50% as good as senior experts(REQ_PERF)
     126  80% as good as senior experts(REQ_PERF)
     127  as good as senior experts(REQ_PERF)
```

USER_ENTH is next to stabilize in Cluster 42 and is also found to be complete:

```
Cluster No.: 42
Properties of newly merged group:
Number Rules in Group = 4
  RuleNo  Description
      41  USER_ENTH=NN -> USER_RSK_F=10
      42  USER_ENTH=LTL -> USER_RSK_F=6
      43  USER_ENTH=SM -> USER_RSK_F=2
      44  USER_ENTH=LOT -> USER_RSK_F=0
Number Stable Patterns in Group = 6
  PattNo  Description
      18  USER_ENTH
     157  none(USER_ENTH)
     158  a little(USER_ENTH)
     159  some(USER_ENTH)
     160  a lot(USER_ENTH)
     232  10
```

COMP_PROF is also found to be complete in Cluster 39:

```
Cluster No.: 39
Properties of newly merged group:
Number Rules in Group = 4
  RuleNo  Description
      45  COMP_PROF=NN -> USER_RSK_F(+10)
      46  COMP_PROF=LTL -> USER_RSK_F(+6)
      47  COMP_PROF=SM -> USER_RSK_F(+2)
      48  COMP_PROF=LOT -> USER_RSK_F(+0)
Number Stable Patterns in Group = 5
  PattNo  Description
      19  COMP_PROF
     161  none(COMP_PROF)
     162  a little(COMP_PROF)
     163  some(COMP_PROF)
     164  a lot(COMP_PROF)
```

The three merges, from Cluster 37 down to Cluster 35, show three qualifiers to be complete — TRNG_TOOL, END_USER, and MGMT_LV.

```
Cluster No.: 37
Properties of newly merged group:
Number Rules in Group = 3
  RuleNo  Description
      18  TRNG_TOOL=LOT -> EST_BEN(+10)
      19  TRNG_TOOL=SM -> EST_BEN(+5)
      20  TRNG_TOOL=NNVLIT -> EST_BEN(+0)
Number Stable Patterns in Group = 5
  PattNo  Description
      37  TRNG_TOOL
     205  none(TRNG_TOOL)
     206  a little(TRNG_TOOL)
     207  some(TRNG_TOOL)
     208  a lot(TRNG_TOOL)
```

```
Cluster No.: 36
Properties of newly merged group:
Number Rules in Group = 3
  RuleNo  Description
      21  END_USER=CL -> EST_BEN(+4)
```

```
      23  END_USER=ADM -> EST_BEN(+10)
      22  END_USER=TECHVPRO -> EST_BEN(+8)
Number Stable Patterns in Group = 6
  PattNo  Description
       3  END_USER
     108  clerical(END_USER)
     109  technicians(END_USER)
     110  professional(END_USER)
     111  administrative(END_USER)
     230  + 8


Cluster No.: 35
Properties of newly merged group:
Number Rules in Group = 3
  RuleNo  Description
      32  MGMT_LV=1V2 -> ORG_RSK_F(+0)
      49  MGMT_LV=3 -> ORG_RSK_F(+2)
      50  MGMT_LV>3 -> ORG_RSK_F(+5)
Number Stable Patterns in Group = 5
  PattNo  Description
      17  MGMT_LV
     153  one(MGMT_LV)
     154  two(MGMT_LV)
     155  three(MGMT_LV)
     156  more than three(MGMT_LV)
```

All the display variables stabilize in Cluster 30:

```
Cluster No.: 30
Properties of newly merged group:
Number Rules in Group = 4
  RuleNo  Description
      59  DISP_BEN+1000>EST_BEN -> conc_ben
      60  ORG_RSK_F<DISP_ORG_RSK+1000 -> conc_org_rsk
      61  DOM_RSK_F<DISP_DOM_R_F+1000 -> conc_dom_rsk
      62  USER_RSK_F<DISP_USER_R_F+1000 -> conc_user_rsk
Number Stable Patterns in Group = 9
  PattNo  Description
      62  DISP_BEN
      63  DISP_ORG_RSK
      64  DISP_DOM_R_F
      65  DISP_USER_R_F
```

```
     71  conc_ben
     72  conc_org_risk
     73  conc_dom_risk
     74  conc_user_risk
    241  + 1000
```

The next qualifiers to stabilize are COMPL, REG, and EXPED_COMPL. COMPL is found to be complete, but the "yes" value for REG and the "no" value for EXPED_COMPL have not been addressed.

```
Cluster No.: 29
Properties of newly merged group:
Number Rules in Group = 5
  RuleNo  Description
      14  COMPL=HG -> EST_BEN(+10)
      15  COMPL=MED -> EST_BEN(+6)
      16  COMPL=LOW -> EST_BEN(+2)
      17  COMPL=NN -> EST_BEN(+0)
      65  COMPL=U,REG=N,EXPED_COMPL=Y -> COMPL=H
Number Stable Patterns in Group = 10
  PattNo  Description
      12  COMPL
      32  REG
      33  EXPED_COMPL
     139  high(COMPL)
     140  medium(COMPL)
     142  low(COMPL)
     143  none(COMPL)
     144  unknown(COMPL)
     194  no(REG)
     195  yes(EXPED_COMPL)
```

All the values for ID_NEED have been addressed as shown through Cluster 27:

```
Cluster No.: 27
Properties of newly merged group:
Number Rules in Group = 5
  RuleNo  Description
       1  ID_NEED=Y -> EST_BEN=10
      63  ID_NEED=N -> EST_BEN=0
      66  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y,EXP_STF=Y -> ID_NEED=Y
```

```
     67  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y -> ID_NEED=Y
     68  ID_NEED=U,PRBLM=CMPLX,EXP_STF=Y -> ID_NEED=Y
Number Stable Patterns in Group = 11
  PattNo  Description
       1  ID_NEED
      22  PRBLM
      23  PROF_~AVL
      24  EXP_STF
     100  yes(ID_NEED)
     101  no(ID_NEED)
     102  unknown(ID_NEED)
     168  complex(PRBLM)
     170  yes(PROF_~AVL)
     172  yes(EXP_STF)
     242  0
```

M_NEED, EXP_AVL, RES, END_USERS, and EXP_LVL all stabilize in Cluster 25. Except for M_NEED, all the other qualifiers are found to be incomplete. For EXP_AVL, "unavailable" is not addressed; for RES, "no" is not addressed; for END_USERS, "unreceptive" is not addressed; and for EXP_LVL, "no" is not addressed.

```
Cluster No.: 25
Properties of newly merged group:
Number Rules in Group = 5
  RuleNo  Description
       2  M_NEED=NN -> EST_BEN(+0)
       3  M_NEED=LIT -> EST_BEN(+2)
       4  M_NEED=SM -> EST_BEN(+6)
       5  M_NEED=LOT -> EST_BEN(+10)
      64  M_NEED=U,EXP_AVL,EXP_LVL,RES,END_USERS -> M_NEED=LOT
Number Stable Patterns in Group = 14
  PattNo  Description
       2  M_NEED
      25  EXP_AVL
      27  RES
      28  END_USERS
      35  EXP_LVL
     103  none(M_NEED)
     104  a little (M_NEED)
     105  some(M_NEED)
     106  a lot(M_NEED)
     107  unknown(M_NEED)
     174  available(EXP_AVL)
```

```
    178  yes(RES)
    180  receptive(END_USERS)
    199  yes(EXP_LVL)
```

ADV is also completely specified. Note that the consequent has an *else* part for addressing values other than "yes."

```
Cluster No.: 12
Properties of newly merged group:
Number Rules in Group = 36
  RuleNo  Description
       1  ID_NEED=Y -> EST_BEN=10
      63  ID_NEED=N -> EST_BEN=0
      66  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y,EXP_STF=Y -> ID_NEED=Y
      67  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y -> ID_NEED=Y
      68  ID_NEED=U,PRBLM=CMPLX,EXP_STF=Y -> ID_NEED=Y
       2  M_NEED=NN -> EST_BEN(+0)
       3  M_NEED=LIT -> EST_BEN(+2)
       4  M_NEED=SM -> EST_BEN(+6)
       5  M_NEED=LOT -> EST_BEN(+10)
      64  M_NEED=U,EXP_AVL,EXP_LVL,RES,END_USERS -> M_NEED=LOT
       6  IMPACT=NN -> EST_BEN(+0)
       7  IMPACT=LIT -> EST_BEN(+2)
       8  IMPACT=SM -> EST_BEN(+6)
       9  IMPACT=MAJ -> EST_BEN(+10)
      10  MAJ_IMPR=NN -> EST_BEN(+0)
      11  MAJ_IMPR=LIT -> EST_BEN(+2)
      12  MAJ_IMPR=SM -> EST_BEN(+6)
      13  MAJ_IMPR=LOT -> EST_BEN(+10)
      14  COMPL=HG -> EST_BEN(+10)
      15  COMPL=MED -> EST_BEN(+6)
      16  COMPL=LOW -> EST_BEN(+2)
      17  COMPL=NN -> EST_BEN(+0)
      65  COMPL=U,REG=N,EXPED_COMPL=Y -> COMPL=H
      18  TRNG_TOOL=LOT -> EST_BEN(+10)
      19  TRNG_TOOL=SM -> EST_BEN(+5)
      20  TRNG_TOOL=NNVLIT -> EST_BEN(+0)
      21  END_USER=CL -> EST_BEN(+4)
      23  END_USER=ADM -> EST_BEN(+10)
      22  END_USER=TECHVPRO -> EST_BEN(+8)
      24  C_AN_T<>0 -> AN_T_SAV(C_AN_T, ES_AN_T)
      27  C_AN_T<>0 -> RAW_AN_EXEC(AN_T_SAV,EXEC_T_SAV)
      26  C_EXEC_T<>0 -> EXEC_T_SAV(C_EXEC_T,ES_EXEC_T)
```

```
     25  YRLY_ANAL_EXEC_T>1000 -> EST_BEN(+10)(0)
     28  NUM_REDO=ES_NUM_REDO -> REDO_SAV(RAW_AN_EXEC,NUM_REDO,ES_NUM_REDO)
     29  YRLY_FREQ>0 -> YRLY_AN_EXEC_T(YRLY_FREQ,REDO_SAV)
     30  ADV=Y -> ORG_RSK_F(+0)(+10)
Number Stable Patterns in Group = 73
  PattNo  Description
       1  ID_NEED
       2  M_NEED
       3  END_USER
      12  COMPL
      14  ADV
      22  PRBLM
      23  PROF_~AVL
      24  EXP_STF
      25  EXP_AVL
      27  RES
      28  END_USERS
      29  IMPACT
      32  REG
      33  EXPED_COMPL
      35  EXP_LVL
      36  MAJ_IMPR
      37  TRNG_TOOL
      40  C_AN_T
      41  ES_AN_T
      42  AN_T_SAV
      43  C_EXEC_T
      44  EXEC_T_SAVINGS
      45  ES_EXEC_T
      46  RAW_AN_EXEC
      47   NUM_REDO
      48  REDO_SAV
      49  ES_NUM_REDO
      51  YRLY_FREQ
      52  YRLY_AN_EXEC_T
     100  yes(ID_NEED)
     101  no(ID_NEED)
     102  unknown(ID_NEED)
     103  none(M_NEED)
     104  a little (M_NEED)
     105  some(M_NEED)
     106  a lot(M_NEED)
     107  unknown(M_NEED)
     108  clerical(END_USER)
     109  technicians(END_USER)
```

```
110  professional(END_USER)
111  administrative(END_USER)
139  high(COMPL)
140  medium(COMPL)
142  low(COMPL)
143  none(COMPL)
144  unknown(COMPL)
147  yes(ADV)
168  complex(PRBLM)
170  yes(PROF_~AVL)
172  yes(EXP_STF)
174  available(EXP_AVL)
178  yes(RES)
180  receptive(END_USERS)
182  major improvement(IMPACT)
183  some improvement(IMPACT)
184  a little improvement(IMPACT)
185  no improvement(IMPACT)
194  no(REG)
195  yes(EXPED_COMPL)
199  yes(EXP_LVL)
201  none(MAJ_IMPROV)
202  a little(MAJ_IMPROV)
203  some(MAJ_IMPROV)
204  a lot(MAJ_IMPROV)
205  none(TRNG_TOOL)
206  a little(TRNG_TOOL)
207  some(TRNG_TOOL)
208  a lot(TRNG_TOOL)
220  <> 0
222  > 0
225  > 1000
230  + 8
242  0
```

Again, MGMT_SUP is similarly found to be complete.

```
Cluster No.: 10
Properties of newly merged group:
Number Rules in Group = 40
  RuleNo  Description
       1  ID_NEED=Y -> EST_BEN=10
      63  ID_NEED=N -> EST_BEN=0
```

```
      66  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y,EXP_STF=Y -> ID_NEED=Y
      67  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y -> ID_NEED=Y
      68  ID_NEED=U,PRBLM=CMPLX,EXP_STF=Y -> ID_NEED=Y
       2  M_NEED=NN -> EST_BEN(+0)
       3  M_NEED=LIT -> EST_BEN(+2)
       4  M_NEED=SM -> EST_BEN(+6)
       5  M_NEED=LOT -> EST_BEN(+10)
      64  M_NEED=U,EXP_AVL,EXP_LVL,RES,END_USERS -> M_NEED=LOT
       6  IMPACT=NN -> EST_BEN(+0)
       7  IMPACT=LIT -> EST_BEN(+2)
       8  IMPACT=SM -> EST_BEN(+6)
       9  IMPACT=MAJ -> EST_BEN(+10)
      10  MAJ_IMPR=NN -> EST_BEN(+0)
      11  MAJ_IMPR=LIT -> EST_BEN(+2)
      12  MAJ_IMPR=SM -> EST_BEN(+6)
      13  MAJ_IMPR=LOT -> EST_BEN(+10)
      14  COMPL=HG -> EST_BEN(+10)
      15  COMPL=MED -> EST_BEN(+6)
      16  COMPL=LOW -> EST_BEN(+2)
      17  COMPL=NN -> EST_BEN(+0)
      65  COMPL=U,REG=N,EXPED_COMPL=Y -> COMPL=H
      18  TRNG_TOOL=LOT -> EST_BEN(+10)
      19  TRNG_TOOL=SM -> EST_BEN(+5)
      20  TRNG_TOOL=NNVLIT -> EST_BEN(+0)
      21  END_USER=CL -> EST_BEN(+4)
      23  END_USER=ADM -> EST_BEN(+10)
      22  END_USER=TECHVPRO -> EST_BEN(+8)
      24  C_AN_T<>0 -> AN_T_SAV(C_AN_T, ES_AN_T)
      27  C_AN_T<>0 -> RAW_AN_EXEC(AN_T_SAV,EXEC_T_SAV)
      26  C_EXEC_T<>0 -> EXEC_T_SAV(C_EXEC_T,ES_EXEC_T)
      25  YRLY_ANAL_EXEC_T>1000 -> EST_BEN(+10)(0)
      28  NUM_REDO=ES_NUM_REDO -> REDO_SAV(RAW_AN_EXEC,NUM_REDO,ES_NUM_REDO)
      29  YRLY_FREQ>0 -> YRLY_AN_EXEC_T(YRLY_FREQ,REDO_SAV)
      30  ADV=Y -> ORG_RSK_F(+0)(+10)
      31  MGMT_SUP=Y -> ORG_RSK_F(+0)(+10)
      32  MGMT_LV=1V2 -> ORG_RSK_F(+0)
      49  MGMT_LV=3 -> ORG_RSK_F(+2)
      50  MGMT_LV>3 -> ORG_RSK_F(+5)
Number Stable Patterns in Group = 80
  PattNo  Description
       1  ID_NEED
       2  M_NEED
       3  END_USER
      12  COMPL
      14  ADV
```

15	MGMT_SUP
17	MGMT_LV
22	PRBLM
23	PROF_~AVL
24	EXP_STF
25	EXP_AVL
27	RES
28	END_USERS
29	IMPACT
32	REG
33	EXPED_COMPL
35	EXP_LVL
36	MAJ_IMPR
37	TRNG_TOOL
40	C_AN_T
41	ES_AN_T
42	AN_T_SAV
43	C_EXEC_T
44	EXEC_T_SAVINGS
45	ES_EXEC_T
46	RAW_AN_EXEC
47	NUM_REDO
48	REDO_SAV
49	ES_NUM_REDO
51	YRLY_FREQ
52	YRLY_AN_EXEC_T
100	yes(ID_NEED)
101	no(ID_NEED)
102	unknown(ID_NEED)
103	none(M_NEED)
104	a little (M_NEED)
105	some(M_NEED)
106	a lot(M_NEED)
107	unknown(M_NEED)
108	clerical(END_USER)
109	technicians(END_USER)
110	professional(END_USER)
111	administrative(END_USER)
139	high(COMPL)
140	medium(COMPL)
142	low(COMPL)
143	none(COMPL)
144	unknown(COMPL)
147	yes(ADV)
149	yes(MGMT_SUP)

```
153  one(MGMT_LV)
154  two(MGMT_LV)
155  three(MGMT_LV)
156  more than three(MGMT_LV)
168  complex(PRBLM)
170  yes(PROF_~AVL)
172  yes(EXP_STF)
174  available(EXP_AVL)
178  yes(RES)
180  receptive(END_USERS)
182  major improvement(IMPACT)
183  some improvement(IMPACT)
184  a little improvement(IMPACT)
185  no improvement(IMPACT)
194  no(REG)
195  yes(EXPED_COMPL)
199  yes(EXP_LVL)
201  none(MAJ_IMPROV)
202  a little(MAJ_IMPROV)
203  some(MAJ_IMPROV)
204  a lot(MAJ_IMPROV)
205  none(TRNG_TOOL)
206  a little(TRNG_TOOL)
207  some(TRNG_TOOL)
208  a lot(TRNG_TOOL)
220  <> 0
222  > 0
225  > 1000
230  + 8
242  0
```

MGMT_EXP is also found to be complete in Cluster 9 through the if-then-else construct.

```
Cluster No.: 9
Properties of newly merged group:
Number Rules in Group = 41
  RuleNo  Description
       1  ID_NEED=Y -> EST_BEN=10
      63  ID_NEED=N -> EST_BEN=0
      66  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y,EXP_STF=Y -> ID_NEED=Y
      67  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y -> ID_NEED=Y
      68  ID_NEED=U,PRBLM=CMPLX,EXP_STF=Y -> ID_NEED=Y
       2  M_NEED=NN -> EST_BEN(+0)
```

```
      3  M_NEED=LIT -> EST_BEN(+2)
      4  M_NEED=SM -> EST_BEN(+6)
      5  M_NEED=LOT -> EST_BEN(+10)
     64  M_NEED=U,EXP_AVL,EXP_LVL,RES,END_USERS -> M_NEED=LOT
      6  IMPACT=NN -> EST_BEN(+0)
      7  IMPACT=LIT -> EST_BEN(+2)
      8  IMPACT=SM -> EST_BEN(+6)
      9  IMPACT=MAJ -> EST_BEN(+10)
     10  MAJ_IMPR=NN -> EST_BEN(+0)
     11  MAJ_IMPR=LIT -> EST_BEN(+2)
     12  MAJ_IMPR=SM -> EST_BEN(+6)
     13  MAJ_IMPR=LOT -> EST_BEN(+10)
     14  COMPL=HG -> EST_BEN(+10)
     15  COMPL=MED -> EST_BEN(+6)
     16  COMPL=LOW -> EST_BEN(+2)
     17  COMPL=NN -> EST_BEN(+0)
     65  COMPL=U,REG=N,EXPED_COMPL=Y -> COMPL=H
     18  TRNG_TOOL=LOT -> EST_BEN(+10)
     19  TRNG_TOOL=SM -> EST_BEN(+5)
     20  TRNG_TOOL=NNVLIT -> EST_BEN(+0)
     21  END_USER=CL -> EST_BEN(+4)
     23  END_USER=ADM -> EST_BEN(+10)
     22  END_USER=TECHVPRO -> EST_BEN(+8)
     24  C_AN_T<>0 -> AN_T_SAV(C_AN_T, ES_AN_T)
     27  C_AN_T<>0 -> RAW_AN_EXEC(AN_T_SAV,EXEC_T_SAV)
     26  C_EXEC_T<>0 -> EXEC_T_SAV(C_EXEC_T,ES_EXEC_T)
     25  YRLY_ANAL_EXEC_T>1000 -> EST_BEN(+10)(0)
     28  NUM_REDO=ES_NUM_REDO -> REDO_SAV(RAW_AN_EXEC,NUM_REDO,ES_NUM_REDO)
     29  YRLY_FREQ>0 -> YRLY_AN_EXEC_T(YRLY_FREQ,REDO_SAV)
     30  ADV=Y -> ORG_RSK_F(+0)(+10)
     31  MGMT_SUP=Y -> ORG_RSK_F(+0)(+10)
     32  MGMT_LV=1V2 -> ORG_RSK_F(+0)
     49  MGMT_LV=3 -> ORG_RSK_F(+2)
     50  MGMT_LV>3 -> ORG_RSK_F(+5)
     33  MGMT_EXP=Y -> ORG_RSK_F(+0)(+10)
Number Stable Patterns in Group = 82
  PattNo  Description
       1  ID_NEED
       2  M_NEED
       3  END_USER
      12  COMPL
      14  ADV
      15  MGMT_SUP
      16  MGMT_EXP
      17  MGMT_LV
```

22	PRBLM
23	PROF_~AVL
24	EXP_STF
25	EXP_AVL
27	RES
28	END_USERS
29	IMPACT
32	REG
33	EXPED_COMPL
35	EXP_LVL
36	MAJ_IMPR
37	TRNG_TOOL
40	C_AN_T
41	ES_AN_T
42	AN_T_SAV
43	C_EXEC_T
44	EXEC_T_SAVINGS
45	ES_EXEC_T
46	RAW_AN_EXEC
47	NUM_REDO
48	REDO_SAV
49	ES_NUM_REDO
51	YRLY_FREQ
52	YRLY_AN_EXEC_T
100	yes(ID_NEED)
101	no(ID_NEED)
102	unknown(ID_NEED)
103	none(M_NEED)
104	a little (M_NEED)
105	some(M_NEED)
106	a lot(M_NEED)
107	unknown(M_NEED)
108	clerical(END_USER)
109	technicians(END_USER)
110	professional(END_USER)
111	administrative(END_USER)
139	high(COMPL)
140	medium(COMPL)
142	low(COMPL)
143	none(COMPL)
144	unknown(COMPL)
147	yes(ADV)
149	yes(MGMT_SUP)
151	yes(MGMT_EXP)
153	one(MGMT_LV)

```
154  two(MGMT_LV)
155  three(MGMT_LV)
156  more than three(MGMT_LV)
168  complex(PRBLM)
170  yes(PROF_~AVL)
172  yes(EXP_STF)
174  available(EXP_AVL)
178  yes(RES)
180  receptive(END_USERS)
182  major improvement(IMPACT)
183  some improvement(IMPACT)
184  a little improvement(IMPACT)
185  no improvement(IMPACT)
194  no(REG)
195  yes(EXPED_COMPL)
199  yes(EXP_LVL)
201  none(MAJ_IMPROV)
202  a little(MAJ_IMPROV)
203  some(MAJ_IMPROV)
204  a lot(MAJ_IMPROV)
205  none(TRNG_TOOL)
206  a little(TRNG_TOOL)
207  some(TRNG_TOOL)
208  a lot(TRNG_TOOL)
220  <> 0
222  > 0
225  > 1000
230  + 8
242  0
```

INTER stabilizes in Cluster 6 and is completely specified:

```
Cluster No.: 6
Properties of newly merged group:
Number Rules in Group = 48
  RuleNo  Description
       1  ID_NEED=Y -> EST_BEN=10
      63  ID_NEED=N -> EST_BEN=0
      66  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y,EXP_STF=Y -> ID_NEED=Y
      67  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y -> ID_NEED=Y
      68  ID_NEED=U,PRBLM=CMPLX,EXP_STF=Y -> ID_NEED=Y
       2  M_NEED=NN -> EST_BEN(+0)
       3  M_NEED=LIT -> EST_BEN(+2)
```

```
     4  M_NEED=SM -> EST_BEN(+6)
     5  M_NEED=LOT -> EST_BEN(+10)
    64  M_NEED=U,EXP_AVL,EXP_LVL,RES,END_USERS -> M_NEED=LOT
     6  IMPACT=NN -> EST_BEN(+0)
     7  IMPACT=LIT -> EST_BEN(+2)
     8  IMPACT=SM -> EST_BEN(+6)
     9  IMPACT=MAJ -> EST_BEN(+10)
    10  MAJ_IMPR=NN -> EST_BEN(+0)
    11  MAJ_IMPR=LIT -> EST_BEN(+2)
    12  MAJ_IMPR=SM -> EST_BEN(+6)
    13  MAJ_IMPR=LOT -> EST_BEN(+10)
    14  COMPL=HG -> EST_BEN(+10)
    15  COMPL=MED -> EST_BEN(+6)
    16  COMPL=LOW -> EST_BEN(+2)
    17  COMPL=NN -> EST_BEN(+0)
    65  COMPL=U,REG=N,EXPED_COMPL=Y -> COMPL=H
    18  TRNG_TOOL=LOT -> EST_BEN(+10)
    19  TRNG_TOOL=SM -> EST_BEN(+5)
    20  TRNG_TOOL=NNVLIT -> EST_BEN(+0)
    21  END_USER=CL -> EST_BEN(+4)
    23  END_USER=ADM -> EST_BEN(+10)
    22  END_USER=TECHVPRO -> EST_BEN(+8)
    24  C_AN_T<>0 -> AN_T_SAV(C_AN_T, ES_AN_T)
    27  C_AN_T<>0 -> RAW_AN_EXEC(AN_T_SAV,EXEC_T_SAV)
    26  C_EXEC_T<>0 -> EXEC_T_SAV(C_EXEC_T,ES_EXEC_T)
    25  YRLY_ANAL_EXEC_T>1000 -> EST_BEN(+10)(0)
    28  NUM_REDO=ES_NUM_REDO -> REDO_SAV(RAW_AN_EXEC,NUM_REDO,ES_NUM_REDO)
    29  YRLY_FREQ>0 -> YRLY_AN_EXEC_T(YRLY_FREQ,REDO_SAV)
    30  ADV=Y -> ORG_RSK_F(+0)(+10)
    31  MGMT_SUP=Y -> ORG_RSK_F(+0)(+10)
    32  MGMT_LV=1V2 -> ORG_RSK_F(+0)
    49  MGMT_LV=3 -> ORG_RSK_F(+2)
    50  MGMT_LV>3 -> ORG_RSK_F(+5)
    33  MGMT_EXP=Y -> ORG_RSK_F(+0)(+10)
    34  MDL=Y_ALG -> DOM_RSK_F=1
    35  MDL=Y_MAN -> DOM_RSK_F=5
    36  MDL=N -> DOM_RSK_F=10
    37  REQ_PERF=50 -> DOM_RSK_F(+4)
    38  REQ_PERF=80 -> DOM_RSK_F(+6)
    39  REQ_PERF=100 -> DOM_RSK_F(+10)
    40  INTER=Y -> DOM_RSK_F(+10)(+0)
Number Stable Patterns in Group = 95
  PattNo  Description
     1  ID_NEED
     2  M_NEED
```

3	END_USER
5	MDL
7	REQ_PERF
12	COMPL
13	INTER
14	ADV
15	MGMT_SUP
16	MGMT_EXP
17	MGMT_LV
22	PRBLM
23	PROF_~AVL
24	EXP_STF
25	EXP_AVL
27	RES
28	END_USERS
29	IMPACT
32	REG
33	EXPED_COMPL
35	EXP_LVL
36	MAJ_IMPR
37	TRNG_TOOL
40	C_AN_T
41	ES_AN_T
42	AN_T_SAV
43	C_EXEC_T
44	EXEC_T_SAVINGS
45	ES_EXEC_T
46	RAW_AN_EXEC
47	NUM_REDO
48	REDO_SAV
49	ES_NUM_REDO
51	YRLY_FREQ
52	YRLY_AN_EXEC_T
100	yes(ID_NEED)
101	no(ID_NEED)
102	unknown(ID_NEED)
103	none(M_NEED)
104	a little(M_NEED)
105	some(M_NEED)
106	a lot(M_NEED)
107	unknown(M_NEED)
108	clerical(END_USER)
109	technicians(END_USER)
110	professional(END_USER)
111	administrative(END_USER)

```
117 yes, an algorithmic system(MDL)
118 yes, a manual system(MDL)
119 no(MDL)
125 50% as good as senior experts(REQ_PERF)
126 80% as good as senior experts(REQ_PERF)
127 as good as senior experts(REQ_PERF)
139 high(COMPL)
140 medium(COMPL)
142 low(COMPL)
143 none(COMPL)
144 unknown(COMPL)
145 yes(INTER)
147 yes(ADV)
149 yes(MGMT_SUP)
151 yes(MGMT_EXP)
153 one(MGMT_LV)
154 two(MGMT_LV)
155 three(MGMT_LV)
156 more than three(MGMT_LV)
168 complex(PRBLM)
170 yes(PROF_~AVL)
172 yes(EXP_STF)
174 available(EXP_AVL)
178 yes(RES)
180 receptive(END_USERS)
182 major improvement(IMPACT)
183 some improvement(IMPACT)
184 a little improvement(IMPACT)
185 no improvement(IMPACT)
194 no(REG)
195 yes(EXPED_COMPL)
199 yes(EXP_LVL)
201 none(MAJ_IMPROV)
202 a little(MAJ_IMPROV)
203 some(MAJ_IMPROV)
204 a lot(MAJ_IMPROV)
205 none(TRNG_TOOL)
206 a little(TRNG_TOOL)
207 some(TRNG_TOOL)
208 a lot(TRNG_TOOL)
220 <> 0
222 > 0
225 > 1000
228 + 5
229 + 4
```

```
230  + 8
231  + 1
242  0
```

END and BEGIN BEN stabilize during the last merge. The clusters become too large at this point to provide any information easily. Hence, other clustering strategies where these variables stabilize in a smaller group, reveal information on them more easily. This ends the completeness specification check on the values of the qualifiers.

Incompleteness in Variable-Value Specifications

This subsection focuses on the variables that have been incompletely specified. In Cluster 26, there is incomplete variable-value specification due to the following: EST_RSK>40 is addressed; however, there are no rules to address EST_RSK≤40.

```
Cluster No.: 26
Properties of newly merged group:
Number Rules in Group = 4
  RuleNo  Description
      51  ORG_RSK_F,USER_RSK_F,DOM_RSK_F -> EST_RSK(ORG_RSK_F,USER_RSK_F,DOM_RSK_F)
      52  EST_BEN>-1,EST_RSK>-1 -> RATIO(EST_BEN/EST_RSK)
      53  EST_BEN<40,EXP_WHY_BEN -> exp_gap
      54  EST_RSK>40,EXP_EST_RSK -> exp_sz
Number Stable Patterns in Group = 8
  PattNo  Description
      54  EST_RSK
      59  EXP_WHY_BEN
      61  EXP_EST_RSK
      77  Exp_gap
      79  Exp_sz
     221  > 1
     236  < 40
     237  > 40
```

Similarly, in Cluster 23, no action is specified for the variable RATIO=1. A close inspection of Rule 52 in this cluster reveals that RATIO is a ratio of EST_BEN and EST_RSK when both are greater than -1. This condition can lead to a situation where both numerator and denominator are equal, in which case, RATIO will be equal to one! There is no action specified for this situation in the rule base.

```
Cluster No.: 23
Properties of newly merged group:
Number Rules in Group = 10
  RuleNo  Description
      51  ORG_RSK_F,USER_RSK_F,DOM_RSK_F -> EST_RSK(ORG_RSK_F,USER_RSK_F,DOM_RSK_F)
      52  EST_BEN>-1,EST_RSK>-1 -> RATIO(EST_BEN/EST_RSK)
      53  EST_BEN<40,EXP_WHY_BEN -> exp_gap
      54  EST_RSK>40,EXP_EST_RSK -> exp_sz
      59  DISP_BEN+1000>EST_BEN -> conc_ben
      60  ORG_RSK_F<DISP_ORG_RSK+1000 -> conc_org_rsk
      61  DOM_RSK_F<DISP_DOM_R_F+1000 -> conc_dom_rsk
      62  USER_RSK_F<DISP_USER_R_F+1000 -> conc_user_rsk
      55  RATIO<1,EXP_LOWRATIO -> exp_low
      56  RATIO>1 -> go_ahd,LK_GOOD
Number Stable Patterns in Group = 25
  PattNo  Description
      54  EST_RSK
      58  RATIO
      59  EXP_WHY_BEN
      60  EXP_LOWRATIO
      61  EXP_EST_RSK
      62  DISP_BEN
      63  DISP_ORG_RSK
      64  DISP_DOM_R_F
      65  DISP_USER_R_F
      66  LK GOOD
      71  conc_ben
      72  conc_org_risk
      73  conc_dom_risk
      74  conc_user_risk
      75  Go ahd
      77  Exp_gap
      78  Exp_low
      79  Exp_sz
     221  > 1
     234  <> ""
```

Certain qualifiers, such as MGMT_SUP, MGMT_EXP, and INTER, that stabilized much later in the clustering process through the antecedent metric, could have been studied much more easily through the data-flow metric where they had stabilized in Clusters 20, 17, and 21, respectively. A noteworthy point is that multiple clusterings are sometimes needed to complete one aspect of study of the knowledge base.

Detection of Anomalous Conditions

There are a number of anomalous conditions that surface quite early in the clustering process with the antecedent metric.

The first merge at Cluster 67 flags C_AN_T as stable. Examining the group closely, it can be seen that two rules have the same premise — C_AN_T <> 0 — but different conclusions. This is an anomalous condition in the rule base as one of the rules (probably Rule 27) will never get a chance to fire (if the expert system shell uses rule ordering as its conflict resolution strategy). To correct the problem, one of these rules needs to be made more specific.

```
Cluster No.: 67
Properties of newly merged group:
Number Rules in Group = 2
  RuleNo  Description
      24  C_AN_T<>0 -> AN_T_SAV(C_AN_T, ES_AN_T)
      27  C_AN_T<>0 -> RAW_AN_EXEC(AN_T_SAV,EXEC_T_SAV)
Number Stable Patterns in Group = 3
  PattNo  Description
      40  C_AN_T
      41  ES_AN_T
      42  AN_T_SAV
```

A redundant condition occurs in Cluster 27 as ID_NEED seems to be overspecified. Rule 66 is asserting that both PROF_ÃVL and EXP_STF have to be affirmative in order to set the value of ID_NEED to "yes." However, Rules 67 and 68 are creating an OR condition for these two qualifiers to set ID_NEED. This conflict needs to be resolved with the help of the domain experts to bring the knowledge base to a consistent state.

```
Cluster No.: 27
Properties of newly merged group:
Number Rules in Group = 5
  RuleNo  Description
       1  ID_NEED=Y -> EST_BEN=10
      63  ID_NEED=N -> EST_BEN=0
      66  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y,EXP_STF=Y -> ID_NEED=Y
      67  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y -> ID_NEED=Y
      68  ID_NEED=U,PRBLM=CMPLX,EXP_STF=Y -> ID_NEED=Y
```

```
Number Stable Patterns in Group = 11
  PattNo  Description
       1  ID_NEED
      22  PRBLM
      23  PROF_~AVL
      24  EXP_STF
     100  yes(ID_NEED)
     101  no(ID_NEED)
     102  unknown(ID_NEED)
     168  complex(PRBLM)
     170  yes(PROF_~AVL)
     172  yes(EXP_STF)
     242  0
```

Consequent Metric

Clusters obtained from the consequent metric can be used to gain a better understanding of the conditions under which various variable settings are being affected. In other words, this metric is useful in exposing all the prerequisites for setting or incrementing the variables. In doing so, it also provides a second check on the dependency relationship of various qualifiers and variables, which has also been studied earlier through the data-flow metric. It is important in this metric to look for various stable pattern combinations instead of isolated stable patterns. This feature will be automated in the next version of the tool. However, for this study, the process has been accomplished manually through the editor. In presenting results on the consequent metric, there is a break from the tradition of presenting clusters in their chronological order. Instead, related stable variable-value pattern combinations are tracked and show how a deeper understanding of the rule base can be brought about through this focused approach.

First, all stable pattern combinations related to EST_BEN are tracked. In Cluster 23, when all possible settings of EST_BEN stabilize, it is realized that ID_NEED is critical in setting the value of EST_BEN to 10 or 0, depending on a "yes" or a "no" answer. A quick check on the various settings for EST_BEN can be performed at this point.

```
Cluster No.: 23
Properties of newly merged group:
Number Rules in Group = 26
  RuleNo  Description
       1  ID_NEED=Y -> EST_BEN=10
       5  M_NEED=LOT -> EST_BEN(+10)
       9  IMPACT=MAJ -> EST_BEN(+10)
      13  MAJ_IMPR=LOT -> EST_BEN(+10)
```

```
     14  COMPL=HG -> EST_BEN(+10)
     18  TRNG_TOOL=LOT -> EST_BEN(+10)
     23  END_USER=ADM -> EST_BEN(+10)
     25  YRLY_ANAL_EXEC_T>1000 -> EST_BEN(+10)(0)
      2  M_NEED=NN -> EST_BEN(+0)
      6  IMPACT=NN -> EST_BEN(+0)
     10  MAJ_IMPR=NN -> EST_BEN(+0)
     17  COMPL=NN -> EST_BEN(+0)
     20  TRNG_TOOL=NNVLIT -> EST_BEN(+0)
      3  M_NEED=LIT -> EST_BEN(+2)
      7  IMPACT=LIT -> EST_BEN(+2)
     11  MAJ_IMPR=LIT -> EST_BEN(+2)
     16  COMPL=LOW -> EST_BEN(+2)
      4  M_NEED=SM -> EST_BEN(+6)
      8  IMPACT=SM -> EST_BEN(+6)
     12  MAJ_IMPR=SM -> EST_BEN(+6)
     15  COMPL=MED -> EST_BEN(+6)
     19  TRNG_TOOL=SM -> EST_BEN(+5)
     21  END_USER=CL -> EST_BEN(+4)
     22  END_USER=TECHVPRO -> EST_BEN(+8)
     52  EST_BEN>-1,EST_RSK>-1 -> RATIO(EST_BEN/EST_RSK)
     63  ID_NEED=N -> EST_BEN=0
Number Stable Patterns in Group = 30
  PattNo  Description
      3  END_USER
     29  IMPACT
     36  MAJ_IMPR
     37  TRNG_TOOL
    101  no(ID_NEED)
    103  none(M_NEED)
    104  a little(M_NEED)
    105  some(M_NEED)
    108  clerical(END_USER)
    109  technicians(END_USER)
    110  professional(END_USER)
    111  administrative(END_USER)
    140  medium(COMPL)
    142  low(COMPL)
    143  none(COMPL)
    182  major improvement(IMPACT)
    183  some improvement(IMPACT)
    184  a little improvement(IMPACT)
    185  no improvement(IMPACT)
    201  none(MAJ_IMPROV)
    202  a little(MAJ_IMPROV)
```

```
203  some(MAJ_IMPROV)
204  a lot(MAJ_IMPROV)
205  none(TRNG_TOOL)
206  a little(TRNG_TOOL)
207  some(TRNG_TOOL)
208  a lot(TRNG_TOOL)
225  > 1000
230  + 8
242  0
```

With regards to ORG_RSK_F, Cluster 52 brings together a pair of very similar rules that can possibly be combined because management support (MGMT_SUP) and management expectations (MGMT_EXP) are very closely related semantically.

```
Cluster No.: 52
Properties of newly merged group:
Number Rules in Group = 2
  RuleNo  Description
      31  MGMT_SUP=Y -> ORG_RSK_F(+0)(+10)
      33  MGMT_EXP=Y -> ORG_RSK_F(+0)(+10)
Number Stable Patterns in Group = 4
  PattNo  Description
      15  MGMT_SUP
      16  MGMT_EXP
     149  yes(MGMT_SUP)
     151  yes(MGMT_EXP)
```

However, Cluster 32 captures all possible settings of ORG_RSK_F. It is evident from this cluster that none of the rules set the value of ORG_RSK_F. All the qualifiers only increment it. This is a flaw in the rule base and needs to be corrected with a domain expert. Since this metric is clustering on consequent similarity alone, it is incapable of stabilizing ORG_RSK_F as this pattern occurs in both the antecedent and the consequent of the rules. However, a facility in the tool will be provided where, if a number of pattern combinations (such as all settings of ORG_RSK_F) are specified, the appropriate cluster(s) can be automatically presented.

```
Cluster No.: 32
Properties of newly merged group:
Number Rules in Group = 6
  RuleNo  Description
      30  ADV=Y -> ORG_RSK_F(+0)(+10)
      32  MGMT_LV=1V2 -> ORG_RSK_F(+0)
```

```
     31  MGMT_SUP=Y -> ORG_RSK_F(+0)(+10)
     33  MGMT_EXP=Y -> ORG_RSK_F(+0)(+10)
     49  MGMT_LV=3 -> ORG_RSK_F(+2)
     50  MGMT_LV>3 -> ORG_RSK_F(+5)
Number Stable Patterns in Group = 11
  PattNo  Description
     14  ADV
     15  MGMT_SUP
     16  MGMT_EXP
     17  MGMT_LV
    147  yes(ADV)
    149  yes(MGMT_SUP)
    151  yes(MGMT_EXP)
    153  one(MGMT_LV)
    154  two(MGMT_LV)
    155  three(MGMT_LV)
    156  more than three(MGMT_LV)
```

Cluster 24 provides all the settings of DOM_RSK_F. MDL is critical in setting this variable, while all the other qualifiers can increment it.

```
Cluster No.: 24
Properties of newly merged group:
Number Rules in Group = 7
  RuleNo  Description
     34  MDL=Y_ALG -> DOM_RSK_F=1
     35  MDL=Y_MAN -> DOM_RSK_F=5
     36  MDL=N -> DOM_RSK_F=10
     39  REQ_PERF=100 -> DOM_RSK_F(+10)
     40  INTER=Y -> DOM_RSK_F(+10)(+0)
     37  REQ_PERF=50 -> DOM_RSK_F(+4)
     38  REQ_PERF=80 -> DOM_RSK_F(+6)
Number Stable Patterns in Group = 10
  PattNo  Description
      5  MDL
      7  REQ_PERF
     13  INTER
    117  yes, an algorithmic system(MDL)
    118  yes, a manual system(MDL)
    119  no(MDL)
    125  50% as good as senior experts(REQ_PERF)
    126  80% as good as senior experts(REQ_PERF)
    127  as good as senior experts(REQ_PERF)
    145  yes(INTER)
```

The next qualifier to be tracked is USER_RSK_F and all settings of this variable are achieved in Cluster 26. As is evident here, USER_RSK_F is initially set depending on the value of USER_ENTH.

```
Cluster No.: 26
Properties of newly merged group:
Number Rules in Group = 8
  RuleNo  Description
      41  USER_ENTH=NN -> USER_RSK_F=10
      42  USER_ENTH=LTL -> USER_RSK_F=6
      46  COMP_PROF=LTL -> USER_RSK_F(+6)
      45  COMP_PROF=NN -> USER_RSK_F(+10)
      43  USER_ENTH=SM -> USER_RSK_F=2
      47  COMP_PROF=SM -> USER_RSK_F(+2)
      44  USER_ENTH=LOT -> USER_RSK_F=0
      48  COMP_PROF=LOT -> USER_RSK_F(+0)
Number Stable Patterns in Group = 11
  PattNo  Description
      18  USER_ENTH
      19  COMP_PROF
     157  none(USER_ENTH)
     158  a little(USER_ENTH)
     159  some(USER_ENTH)
     160  a lot(USER_ENTH)
     161  none(COMP_PROF)
     162  a little(COMP_PROF)
     163  some(COMP_PROF)
     164  a lot(COMP_PROF)
     232  10
```

Some other anomalies that surfaced through this clustering had also become evident through clustering with other metrics. However, they are worth mentioning again in this context. Cluster 50 exposed the redundancy condition in ID_NEED setting.

```
Cluster No.: 50
Properties of newly merged group:
Number Rules in Group = 3
  RuleNo  Description
      66  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y,EXP_STF=Y -> ID_NEED=Y
      67  ID_NEED=U,PRBLM=CMPLX,PROF_~AVL=Y -> ID_NEED=Y
      68  ID_NEED=U,PRBLM=CMPLX,EXP_STF=Y -> ID_NEED=Y
```

```
Number Stable Patterns in Group = 7
  PattNo  Description
      22  PRBLM
      23  PROF_~AVL
      24  EXP_STF
     102  unknown(ID_NEED)
     168  complex(PRBLM)
     170  yes(PROF_~AVL)
     172  yes(EXP_STF)
```

Cluster 22 exposes the anomalous condition of C_AN_T. It also shows that C_AN_T has to be set externally.

```
Cluster No.: 22
Properties of newly merged group:
Number Rules in Group = 2
  RuleNo  Description
      24  C_AN_T<>0 -> AN_T_SAV(C_AN_T,ES_AN_T)
      27  C_AN_T<>0 -> RAW_AN_EXEC(AN_T_SAV,EXEC_T_SAV)
Number Stable Patterns in Group = 3
  PattNo  Description
      40  C_AN_T
      41  ES_AN_T
      42  AN_T_SAV
```

Total Metric

ESAA is a rule base where the bulk of the closeness between rules is due to the similarity in the antecedents. In this manner, it resembles a monitoring system. However, the presence of a few variables do chain some rules from right to left also. When the total metric was run on ESAA, information obtained from the clusters was very similar to the ones in the antecedent metric. Almost the same clusters were generated; only the ordering of their generation differed. That is, some qualifiers stabilized earlier with the total metric than with the antecedent metric. Since no new information was generated through this metric, the contractor opted not to present the clusters from it.

CONCLUSIONS

The value of using the MVP-CA tool for analyzing a poorly structured rule base such as ESAA has been shown. In particular, the MVP-CA methodology is capable of exposing the current underlying software architecture of the knowledge base. This is especially useful when a knowledge base is in an evolving state or if there are multiple experts updating the

Table 1: Incompleteness in qualifier value specifications.

Qualifier Name	Specified Value	Unspecified Value
PROF_ AVL	yes	no
PRBLM	complex	not complex
EXP_STF	yes	no
REG	no	yes
EXPED_COMPL	yes	no
EXP_AVL	available	unavailable
RES	yes	no
END_USERS	receptive	unreceptive
EXP_LVL	yes	no

Table 2: Incompleteness in variable value specifications.

Variable Name	Specified Value	Unspecified Value
EST_RSK	> 40	≤ 40
RATIO	$> 1, < 1$	$= 1$
ORG_RSK_F	always incremented	never set

knowledge base with their own expertise. Having a **(semi-)automated means of exposing the current underlying structure**, as shown in figure 2, can pave the way towards showing alternate methods of restructuring the system while the system is evolving. Also, inconsistencies generated by multiple expert opinions need to be detected soon enough in the software lifecycle. MVP-CA technology juxtaposes rules with similar content and structure so that **inconsistencies and anomalies become apparent easily.** As was shown in ESAA, **incomplete specifications on various variable and qualifier values also surfaced quite easily** through the grouping of rules from multiple perspectives.

To recapitulate the results, we present our findings in a tabular form in tables 1, 2, and 3.

Table 3: Anomalies across ESAA rules.

Rule Numbers	Type of Anomaly
24,27	Conflict Condition
1,63,66,67,68	Redundancy

It should be noted that there are different types and levels of information to be obtained through the MVP-CA analysis that depend on the size of the rule base, as well as on how thoroughly issues have been thought out during the development of the knowledge base.

In rule bases such as PAMEX, where there is quantification of domain knowledge in terms of qualifiers and variables having well-demarcated numeric values, MVP-CA technology has been able to expose the relevant combinatorial regions for qualifiers and variables that need to be tested. In addition, since there are no incompleteness or inconsistency issues across the substructures exposed in PAMEX by the MVP-CA tool, restructuring possibilities for such an expert system become apparent when inter-cluster analysis is done on this knowledge base.

Until the errors in ESAA are corrected and the rule base is brought to a semantically complete and consistent state, MVP-CA cannot reveal information that provides insight into, for example, restructuring the knowledge base or cutting down on the testing aspects of the knowledge base.

Another difference between PAMEX and ESAA is that the former has a lot of numeric ranges attached to the variable values. Therefore, partitioning this rule base through the MVP-CA tool gives valuable information on the pertinent combinations for testing this rule base with a reduced number of Hoffman regions.[5] ESAA, on the other hand, has qualifiers and variables with mostly literal settings, such as little, small, high, etc. These can be highly subjective values and hence the rule base can be especially difficult to verify or validate. However, figure 2 does expose the qualifiers that can be separately tested. For example, EXPED_COMPL or PROF_ÃVL never impacts EST_BEN or RATIO directly. Hence, qualifiers at that level can be tested out separately.

Some amount of restructuring information can be obtained through the consequent metric results on ESAA. However, these are mainly in terms of combining rules that are addressing similar types of information, such as management support and management expectations setting the organizational risk factor in the same manner. However, since both these qualifiers are incompletely specified, one cannot yet advocate combining the two rules together.

Thus, it is believed that the MVP-CA technology is capable of providing useful information to both developers and maintainers of software systems regardless of which stage the software is in. In the earlier stages when there are likely to be anomalous and conflicting conditions, MVP-CA technology can aid in the process of exposing and removing the inconsistencies. When software evolves to a more mature state, MVP-CA technology can help in restructuring the knowledge base so that maintenance and testing can be realized more effectively and efficiently.

FUTURE WORK

- The contractor is currently designing a stylized interface language for allowing users to interact with the tool.
- The contractor will provide user control on concatenating patterns together before the pattern numbers are generated so that domain-specific information can be incorporated into the clustering process.
- The contractor is planning on providing software for the tool to be able to expose combinations of stable patterns. This information can feed directly into testing efforts for the rule base by showing which subdomains combine legitimately with other subdomains in the knowledge base.
- A documentation phase is also planned where cluster-based information (such as the dominant pattern of a cluster, outlier rules in the cluster, parent and child of the cluster, etc.) can be captured and stored.
- An infrastructure to generate a diagrammatic representation of the software architecture of the system will also be provided in the tool.
- Generalization of the MVP-CA technology on other types of testbeds is planned. In a rule base, rules are the basic entities to be clustered, and the patterns in the rules form the concepts around which clustering takes place. In any other system, once the set of entities to be clustered and a basis for clustering the entities are defined, we can use this technology to obtain similar types of information from other systems that are used to represent information.

APPENDIX A. ESAA RULES

RULE NUMBER: 1 ID_NEED=Y⟶ESTBEN=10

IF:

Is there a clearly identified need that can best be addressed by an expert system? Note: Existing conditions must be known and described to establish a need. Yes.

THEN:

[EST BEN] IS GIVEN THE VALUE + 10

RULE NUMBER: 2 M_NEED=NN⟶ESTBEN(+0)

IF:

What is the probability that the expert system will meet the identified need? Note: If a conventional system exists that meets the identified need, then use it. None.

THEN:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 0

RULE NUMBER: 3 M_NEED=LIT⟶ESTBEN(+2)

IF:

What is the probability that the expert system will meet the identified need? Note: If a conventional system exists that meets the identified need, then use it. A little.

THEN:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 2

RULE NUMBER: 4 M_NEED=SM⟶ESTBEN(+6)

IF:

What is the probability that the expert system will meet the identified need? Note: If a conventional system exists that meets the identified need, then use it. Some.

THEN:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 6

RULE NUMBER: 5   M_NEED=LOT⟶ESTBEN(+10)

IF:

What is the probability that the expert system will meet the identified need? Note: If a conventional system exists that meets the identified need, then use it. A lot.

THEN:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 10

RULE NUMBER: 6   IMPACT=NN⟶ESTBEN(+0)

IF:

What will the impact of the expert system be on ANY of the following categories? * cost reduction * efficiency * transfer of data and/or results * legal implications? No improvement.

THEN:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 0

RULE NUMBER: 7   IMPACT=LIT⟶ESTBEN(+2)

IF:

What will the impact of the expert system be on ANY of the following categories? * cost reduction * efficiency * transfer of data and/or results * legal implications? A little improvement.

THEN:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 2

RULE NUMBER: 8 IMPACT=SM⟶ESTBEN(+6)

IF:

What will the impact of the expert system be on ANY of the following categories? * cost reduction * efficiency * transfer of data and/or results * legal implications? Some improvement.

THEN:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 6

RULE NUMBER: 9 IMPACT=MAJ⟶ESTBEN(+10)

IF:

What will the impact of the expert system be on ANY of the following categories? * cost reduction * efficiency * transfer of data and/or results * legal implications? Major improvement.

THEN:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 10

RULE NUMBER: 10 MAJ_IMPR=NN⟶ESTBEN(+0)

IF:

To what extent will use of the expert system result in major improvement of conditions because of * better solutions * more complete solutions * identification of a better set of alternatives? None.

THEN:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 0

RULE NUMBER: 11 MAJ_IMPR=LIT⟶ESTBEN(+2)

IF:

To what extent will use of the expert system result in major improvement of conditions because of * better solutions * more complete solutions * identification of a better set of alternatives? A little.

THEN:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 2

RULE NUMBER: 12 MAJ_IMPR=SM⟶ESTBEN(+6)

IF:

To what extent will use of the expert system result in major improvement of conditions because of * better solutions * more complete solutions * identification of a better set of alternatives? Some.

THEN:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 6

RULE NUMBER: 13 MAJ_IMPR=LOT⟶ESTBEN(+10)

IF:

To what extent will use of the expert system result in major improvement of conditions because of * better solutions * more complete solutions * identification of a better set of alternatives? A lot.

THEN:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 10

RULE NUMBER: 14 COMPL=HG⟶ESTBEN(+10)

IF:

What is the extent to which the expert system will improve compliance with internal or external requirements (regulations, procedures, guidelines, etc.)? High.

THEN:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 10

RULE NUMBER: 15 COMPL=MED⟶ESTBEN(+6)

IF:

What is the extent to which the expert system will improve compliance with internal or external requirements (regulations, procedures, guidelines, etc.)? Medium.

THEN:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 6

RULE NUMBER: 16 COMPL=LOW⟶ESTBEN(+2)

IF:

What is the extent to which the expert system will improve compliance with internal or external requirements (regulations, procedures, guidelines, etc.)? Low.

THEN:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 2

RULE NUMBER: 17 COMPL=NN⟶ESTBEN(+0)

IF:

What is the extent to which the expert system will improve compliance with internal or external requirements (regulations, procedures, guidelines, etc.)? None.

THEN:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 0

RULE NUMBER: 18 TRNG_TOOL=LOT—→ESTBEN(+10)

IF:

To what extent will the expert system * be used as a training tool * increase users' proficiency and understanding so they can function better independently of the expert system? A lot.

THEN:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 10

RULE NUMBER: 19 TRNG_TOOL=SM—→ESTBEN(+5)

IF:

To what extent will the expert system * be used as a training tool * increase users' proficiency and understanding so they can function better independently of the expert system? Some.

THEN:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 5

RULE NUMBER: 20 TRNG_TOOL=ESTBEN(+0)

IF:

To what extent will the expert system * be used as a training tool * increase users' proficiency and understanding so they can function better independently of the expert system? None OR a little.

THEN:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 0

RULE NUMBER: 21 END_USER=CL—→ESTBEN(+4)

IF:

Who will be the end user of the expert system? Clerical.

THEN:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 4

RULE NUMBER: 22 END_USER=TECHVPRO⟶ESTBEN(+8)

IF:

Who will be the end user of the expert system? Technicians OR professional.

THEN:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 8

RULE NUMBER: 23 END_USER=ADM⟶ESTBEN(+10)

IF:

Who will be the end user of the expert system? Administrative.

THEN:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 10

RULE NUMBER: 24 C_AN_TM⟶AN_TIME_SAV(C_AN_T)(ES_AN_T)

IF:

[C AN TIME] <> 0

THEN:

[AN TIME SAVINGS] IS GIVEN THE VALUE [C AN TIME] - [ES AN TIME]

RULE NUMBER: 25 YRLY_ANAL_EXEC_T>1000⟶ESTBEN(+10)(0)

IF:

[YRLY ANAL AND EXEC T] > 1000

THEN:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 10

ELSE:

[EST BEN] IS GIVEN THE VALUE [EST BEN] + 0

RULE NUMBER: 26 C_EXEC_T<>0⟶EXEC_T_SAV(C_EXEC_T)(ES_EXEC_T)

IF:

[C EXEC TIME] <> 0

THEN:

[EXEC TIME SAVINGS] IS GIVEN THE VALUE [C EXEC TIME] - [ES EXEC TIME]

RULE NUMBER: 27 C_AN_T<>0⟶RAW_AN_EXEC(AN_T_SAV)(EXEC_T_SAV)

IF:

[C AN TIME] <> 0

THEN:

[RAW ANALYSIS AND EXE] IS GIVEN THE VALUE [AN TIME SAVINGS] + [EXEC TIME SAVINGS]

RULE NUMBER: 28
NUM_REDO=ES_NUM_REDO⟶REDO_SAV(RAW_AN_EXEC)(NUM_REDO)(ES_NUM_REDO)

IF:

[NUM REDO] = [ES NUM REDO]

THEN:

[REDO SAVINGS] IS GIVEN THE VALUE [RAW ANALYSIS AND EXE]

ELSE:

[REDO SAVINGS] IS GIVEN THE VALUE ([NUM REDO] - [ES NUM REDO] + 1) * [RAW ANALYSIS AND EXE]

RULE NUMBER: 29 YRLY_FREQ>0⟶YRLY_AN_EXEC_T(YRLY_FREQ)(REDO_SAV)

IF:

[YEARLY FREQ] > 0

THEN:

[YRLY ANAL AND EXEC T] IS GIVEN THE VALUE [YEARLY FREQ] * [REDO SAVINGS]

RULE NUMBER: 30 ADV=Y⟶ORG_RSK_F(+0)(+10)

IF:

Is there a dedicated advocate who wants the system to be a success? Yes.

THEN:

[ORG RISK FACTORS] IS GIVEN THE VALUE + 0

ELSE:

[ORG RISK FACTORS] IS GIVEN THE VALUE + 10

RULE NUMBER: 31 MGMT_SUP=Y⟶ORG_RSK_F(+0)(+10)

IF:

Is there management support? Yes.

THEN:

[ORG RISK FACTORS] IS GIVEN THE VALUE [ORG RISK FACTORS] + 0

ELSE:

[ORG RISK FACTORS] IS GIVEN THE VALUE [ORG RISK FACTORS] + 10

RULE NUMBER: 32 MGMT_LV=1V2⟶ORG_RSK_F(+0)

IF:

How many levels (management) have to approve the results of the expert system before they are applied? One OR two.

THEN:

[ORG RISK FACTORS] IS GIVEN THE VALUE [ORG RISK FACTORS] + 0

RULE NUMBER: 33 MGMT_EXP=Y⟶ORG_RSK_F(+0)(+10)

IF:

Does management have realistic expectations regarding the performance of the developed system? Yes.

THEN:

[ORG RISK FACTORS] IS GIVEN THE VALUE [ORG RISK FACTORS] + 0

ELSE:

[ORG RISK FACTORS] IS GIVEN THE VALUE [ORG RISK FACTORS] + 10

RULE NUMBER: 34 MDL=Y_ALG⟶DOM_RSK_F=+1

IF:

Is there a procedure to be used as a model for the expert system? Yes, an algorithmic system.

THEN:

[DOMAIN RISK FACTORS] IS GIVEN THE VALUE + 1

RULE NUMBER: 35 MDL=Y_MAN⟶DOM_RSK_F=+5

IF:

Is there a procedure to be used as a model for the expert system? Yes, a manual system.

THEN:

[DOMAIN RISK FACTORS] IS GIVEN THE VALUE + 5

RULE NUMBER: 36 MDL=N⟶DOM_RSK_F=+10

IF:

Is there a procedure to be used as a model for the expert system? No.

THEN:

[DOMAIN RISK FACTORS] IS GIVEN THE VALUE + 10

RULE NUMBER: 37 REQ_PERF=50⟶DOM_RSK_F(+4)

IF:

Required performance (in terms of finding the best solution) as compared to senior experts. 50 percent as good as senior experts.

THEN:

[DOMAIN RISK FACTORS] IS GIVEN THE VALUE [DOMAIN RISK FACTORS] + 4

RULE NUMBER: 38 REQ_PERF=80⟶DOM_RSK_F(+6)

IF:

Required performance (in terms of finding the best solution) as compared to senior experts. 80 percent as good as senior experts.

THEN:

[DOMAIN RISK FACTORS] IS GIVEN THE VALUE [DOMAIN RISK FACTORS] + 6

RULE NUMBER: 39 REQ_PERF=100⟶DOM_RSK_F(+10)

IF:

Required performance (in terms of finding the best solution) as compared to senior experts. As good as senior experts.

THEN:

[DOMAIN RISK FACTORS] IS GIVEN THE VALUE [DOMAIN RISK FACTORS] + 10

RULE NUMBER: 40 INTER=Y⟶DOM_RSK_F(10)(+0)

IF:

Is the interaction with external programs required? Yes.

THEN:

[DOMAIN RISK FACTORS] IS GIVEN THE VALUE [DOMAIN RISK FACTORS] + 10

ELSE:

[DOMAIN RISK FACTORS] IS GIVEN THE VALUE [DOMAIN RISK FACTORS] + 0

RULE NUMBER: 41 USER_ENTH=NN⟶USER_RSK_F=10

IF:

What is the level of enthusiasm for the expert system by the intended users (i.e., how much do the intended users want the expert system)? None.

THEN:

[USER RISK FACTORS] IS GIVEN THE VALUE 10

RULE NUMBER: 42 USER_ENTH=LTL⟶USER_RSK_F=6

IF:

What is the level of enthusiasm for the expert system by the intended users (i.e., how much do the intended users want the expert system)? A little.

THEN:

[USER RISK FACTORS] IS GIVEN THE VALUE + 6

RULE NUMBER: 43 USER_ENTH=SM⟶USER_RSK_F=2

IF:

What is the level of enthusiasm for the expert system by the intended users (i.e., how much do the intended users want the expert system)? Some.

THEN:

[USER RISK FACTORS] IS GIVEN THE VALUE + 2

RULE NUMBER: 44 USER_ENTH=LOT⟶USER_RSK_F=0

IF:

What is the level of enthusiasm for the expert system by the intended users (i.e., how much do the intended users want the expert system)? A lot.

THEN:

[USER RISK FACTORS] IS GIVEN THE VALUE + 0

RULE NUMBER: 45 COMP_PROF=NN⟶USER_RSK_F(10)

IF:

The level of computer proficiency of the expert system users. None.

THEN:

[USER RISK FACTORS] IS GIVEN THE VALUE [USER RISK FACTORS] + 10

RULE NUMBER: 46 COMP_PROF=LTL⟶USER_RSK_F(6)

IF:

The level of computer proficiency of the expert system users. A little.

THEN:

[USER RISK FACTORS] IS GIVEN THE VALUE [USER RISK FACTORS] + 6

RULE NUMBER: 47 COMP_PROF=SM⟶USER_RSK_F(2)

IF:

The level of computer proficiency of the expert system users. Some.

THEN:

[USER RISK FACTORS] IS GIVEN THE VALUE [USER RISK FACTORS] + 2

RULE NUMBER: 48   COMP_PROF=LOT⟶USER_RSK_F(0)

IF:

The level of computer proficiency of the expert system users. A lot.

THEN:

[USER RISK FACTORS] IS GIVEN THE VALUE [USER RISK FACTORS] + 0

RULE NUMBER: 49   MGMT_LV=3⟶ORG_RSK_F(+2)

IF:

How many levels (management) have to approve the results of the expert system before they are applied? Three.

THEN:

[ORG RISK FACTORS] IS GIVEN THE VALUE [ORG RISK FACTORS] + 2

RULE NUMBER: 50   MGMT_LV>3⟶ORG_RSK_F(+5)

IF:

How many levels (management) have to approve the results of the expert system before they are applied? More than three.

THEN:

[ORG RISK FACTORS] IS GIVEN THE VALUE [ORG RISK FACTORS] + 5

RULE NUMBER: 51   ORG_RSK_F>-1,USER_RSK_F>-1,DOM_RSK_F>-1⟶EST_RSK(ORG_RSK_F)(USER_RSK_F)(DOM_RSK_F)

IF:

[ORG RISK FACTORS] > -1 and [USER RISK FACTORS] > -1 and [DOMAIN RISK

FACTORS] > -1

THEN:

[EST RISK] IS GIVEN THE VALUE [ORG RISK FACTORS] + [USER RISK FACTORS] + [DOMAIN RISK FACTORS]

RULE NUMBER: 52 EST_BEN>-1,EST_RSK>-1⟶RATIO(EST_BEN,EST_RSK)

IF:

[EST BEN] > - 1 and [EST RISK] > -1

THEN:

[EST BEN EST RISK RAT] IS GIVEN THE VALUE [EST BEN] / [EST RISK]

RULE NUMBER: 53 EST_BEN<40,EXP_WHY_BEN⟶EXP_GAP

IF:

[EST BEN] < 40 and [EXPLAIN WHY BEN] <> ""

THEN:

Explanation for gap between system's view and advocate's view re benefit has been given - Confidence=1.

RULE NUMBER: 54 EST_RSK>40,EXP_EST_RSK⟶EXP_SZ

IF:

[EST RISK] > 40 and [EXPLAIN EST RISK] <> ""

THEN:

Explanation of why the proposed expert system should be built despite the size of the estimated risk - Confidence=1.

RULE NUMBER: 55 RATIO<1,EXP_LOWRATIO—→EXP_LOW

IF:

[EST BEN EST RISK RAT] < 1 and [EXPLAIN LOW BEN RISK RATE] <> ""

THEN:

Explanation of why the proposed expert system should be built despite the low benefit/risk ratio - Confidence=1.

RULE NUMBER: 56 RATIO>1,GO_AHD—→ALTERN

IF:

[EST BEN EST RISK RAT] > 1

THEN:

Go ahead with the proposed expert system since the estimated benefit is greater than the estimated risk - Confidence=1 and consider alternatives to building the expert system.

RULE NUMBER: 57 END=Y—→INP_COMP

IF:

All questions have been asked? Yes.

THEN:

Input complete - Confidence=1.

RULE NUMBER: 58 (BEGIN)BEG_BEN=Y—→BEG_BEN

IF:

Begin benefits? Yes.

THEN:

Begin benefits - Confidence=1.

RULE NUMBER: 59 DISP_BEN+1000>EST_BEN⟶CONC_BEN

IF:

[DISPLAY BEN] + 1000 > [EST BEN]

THEN:

Estimation of benefits concluded - Confidence=1.

RULE NUMBER: 60 ORG_RSK_F<DISP_ORG_RSK+1000⟶CONC_ORG_RSK

IF:

[ORG RISK FACTORS] < [DISPLAY ORG RISK] + 1000

THEN:

Estimation of organizational risk factors concluded - Confidence=1.

RULE NUMBER: 61 DOM_RSK_F<DISP_DOM_R_F+1000⟶CONC_DOM_RSK

IF:

[DOMAIN RISK FACTORS] < [DISPLAY DOMAIN R F] + 1000

THEN:

Estimation of domain risk factors concluded - Confidence=1.

RULE NUMBER: 62 USER_RSK_F<DISP_USER_R_F+1000⟶CONC_USER_RSK

IF:

[USER RISK FACTORS] < [DISPLAY USER R F] + 1000

THEN:

Estimation of user risk factors concluded - Confidence=1.

RULE NUMBER: 63 ID_NEED=N⟶ESTBEN=0

IF:

Is there a clearly identified need that can best be addressed by an expert system? Note: Existing conditions must be known and described to establish a need. No.

THEN:

[EST BEN] IS GIVEN THE VALUE 0

RULE NUMBER: 64
PROB_ID_NEED,EXP_AVAIL,EXP_LVL,RES,END_USERS⟶MT_NEED=LOT

IF:

What is the probability that the expert system will meet the identified need? Note: If a conventional system exists that meets the identified need, then use it. Unknown.

and Experts are available?

and Experts' level of performance is sufficiently better than users' level of performance to make the expert system worthwhile? Yes.

and Sufficient resources (time and money) are available to build and test the expert system? Yes.

and End users of the expert system will be receptive?

THEN:

clear(Q “meet need”)

and What is the probability that the expert system will meet the identified need? Note: If a conventional system exists that meets the identified need, then use it. A lot.

RULE NUMBER: 65 COMPL=U,REG=N,EXPED_COMPL$\longrightarrow$COMPL=H

IF:

What is the extent to which the expert system will improve compliance with internal or external requirements (regulations, procedures, guidelines, etc.)? Unknown.

and Current regulations are uniformly understood and followed? No.

and The expert system will expedite uniform compliance? Yes.

THEN:

clear(Q “compliance”)

and What is the extent to which the expert system will improve compliance with internal or external requirements (regulations, procedures, guidelines, etc.)? High.

RULE NUMBER: 66
ID_NEED=U,PRBLM=CMPLX,PROF_ AVL=Y,EXP_STF=Y$\longrightarrow$ID_NEED=Y

IF:

Is there a clearly identified need that can best be addressed by an expert system? Note: Existing conditions must be known and described to establish a need. Unknown.

and The problem is complex?

and Professionals are currently needed to solve the problem and these professionals are not always available? Yes.

and Less experienced staff need suport and/or advice in solving the problem? Yes.

THEN:

clear(Q "identified need")

and Is there a clearly identified need that can best be addressed by an expert system? Note: Existing conditions must be known and described to establish a need. Yes.

RULE NUMBER: 67 ID_NEED=U,PRBLM=CMPLX,PROF_ AVL=Y⟶ID_NEED=Y

IF:

Is there a clearly identified need that can best be addressed by an expert system? Note: Existing conditions must be known and described to establish a need. Unknown.

and The problem is complex?

and Professionals are currently needed to solve the problem and these professionals are not always available? Yes.

THEN:

clear(Q "identified need")

and Is there a clearly identified need that can best be addressed by an expert system? Note: Existing conditions must be known and described to establish a need. Yes.

RULE NUMBER: 68 ID_NEED=U,PRBLM=CMPLX,EXP_STF=Y⟶ID_NEED=Y

IF:

Is there a clearly identified need that can best be addressed by an expert system? Note: Existing conditions must be known and described to establish a need. Unknown.

and The problem is complex?

and Less experienced staff need suport and/or advice in solving the problem? Yes.

THEN:

clear(Q "identified need")

and Is there a clearly identified need that can best be addressed by an expert system? Note: Existing conditions must be known and described to establish a need. Yes.

REFERENCES

1. K. Bellman and C. Landauer. Designing testable, heterogeneous software environments. *Journal of Systems and Software*, March 1996.

2. K. L. Bellman and D. O. Walter. Analyzing and correcting knowledge-based systems requires explicit models. In *AAAI-88 Workshop on Verification, Validation and Testing of Knowledge-Based Systems*, July 1988.

3. C. Landauer. Correctness principles for rule-based expert systems. *Journal of Expert Systems With Applications*, 1:291–316, 1990.

4. C. Landauer and K. L. Bellman. Constructed complex systems: Architectures, mathematics, and semiotics. In *Proceedings, 1995 Semiotics Conference*, May 1995.

5. M. Mehrotra. *Application of Multi-Viewpoint Clustering Analysis to a Highway Maintenance System.* Report No. FHWA-RD-95-087, Federal Highway Administration, McLean, VA, August 1995 (can be obtained from the National Technical Information Service (NTIS)).

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

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