경영정보학연구
제4권 1호
1994 년 6 월

# 지식기반 임금예측시스템 설계와 구축 사례

조 재 희[1)]

## Design and Implementation of a Knowledge-Based Wage Rate Prediction System

*Potential employers considering locations for production or service facilities typically equire detailed advance knowledge of the wages they will be expected to offer for workers in various occupational categories. The State of Missouri's Department of Labor and Industrial Relations is often contacted by organizations requesting such information. The current wage rate survey approach, initiated in 1988, allows the Department to predict an appropriate wage rate for a given occupation in certain counties, adjusted for changes in the Consumer Price Index (CPI). However, both Department employees and firms have indicated that improved prediction responsiveness and accuracy are desirable. A major deficiency of the current approach is its inability to predict wages for unsurveyed counties. This paper describes a knowledge-based system (KBS), currently in the prototype testing stage, that is expected to supplement the wage rate survey in the near future.*

1) 광운대학교 경영정보학과

# Ⅰ. INTRODUCTION

Wages often constitute a substantial portion of an organization's costs. Given the significant variation in wage rates among locations, the siting decisions made by a firm can have a major impact on profitability or survival. Information detailing wage rates by location allows a potential employer to select the most cost-effective site among several choices. Such information may also be utilized by established employers to compare wage rates among their existing sites, for purposes of shifting production or equalizing disparities among occupational categories by location.

Location-dependent wage rate prediction offers several interesting problem areas, including data collection, similarity determination, and prediction. Data collection is problematic due to scale, scope, and temporal considerations : The number of locations is often very large, with an even larger count of occupation categories, complicated by continual changes in wage rates to compensate for inflation and other factors. Similarity determination enters the situation when the decision is made to approximate some location/occupation/industry/time combinations by analogy to known combinations, instead of attempting to measure every possible combination. Prediction is of relevance when a potential employer will have a major impact on the location, as well as when estimates of wage rates are desired for purposes of location selection, budgeting, or staffing decisions.

The State of Missouri's Department of Labor and Industrial Relations utilizes a wage rate survey to furnish estimates of wage rates by occupation and location. The state of Missouri consists of 114 counties, which constitute the locational units examined by both the wage rate survey and the knowledge-based system introduced in this paper. Occupational Employment Statistics (OES) codes are used to represent over 400 job types, and Standard Industry Classification (SIC) codes represent over 500 primary employer activities [OMB, 1987].

The 1989 and 1990 Wage Rate Surveys produced databases on 61 counties, surveying 33 counties and 28 counties respectively. The sizes of the 61 county wage rate databases vary widely in terms of the number of jobs and industry. However, this department receives constant inquiries from potential investors or established employers about wage rates not only on surveyed counties but also on unsurveyed counties. Even with the database of a surveyed county, the department cannot provide information on all potential wage rates because certain jobs simply do not exist in certain counties. Thus, the department wanted to build a computer-based system to predict wage rates by using the existing 61 wage rate databases.

## Ⅱ. METHODOLOGY

Tihree approaches were considered as prediction-enabling extensions to the survey system : regression analysis, artificial neural systems, and knowledge-based systems. Regression analysis, relying heavily on the use of continuous variables, was deemed unsuitable for this problem due to the existence of two categorical variables (OES and SIC) with large numbers of values. Were a regression equation for wage rate prediction to be devised, it would require at least several hundred (if not several thousand) binary independent variables. For example, it is not possible toassign relative or rank-ordered values to job type ; "chef" and "financial manager" have only transient relationships. The same holds for industry type. The result would be little more than anextremely cumbersome look-up table.

Artificial neural systems (ANS), or neural nets, are a novel and rather intriguing approach whichholds promise in resolving the wage rate prediction problem. The possible application of ANS inthis domain is currently explored, and it is believed that the two primary hurdles experienced thus far may be discrimination among the several hundred categories characteristic of this problem. The secondrelated concern involves the ability of an ANS to adequately distinguish between categorical andcontinuous variables. The lack of an explanation mechanism would hinder implementation of acurrent black-box style ANS-based system in this instance, as potential and current employers typically require justification before accepting the predicted wage rate.

The current wage rate survey essentially produces an unenhanced database without even simplequery capabilities. It is incapable of performing any decision support tasks of interest to its users.Further, not all of the users are equally skilled at manipulating the data ; some have proventhemselves true experts, while others can perform only lower-level tasks or provide predictionswith a lesser degree of confidence. Examination of the situation found numerous characteristics indicating that a knowledge-based system would be appropriate, justified, and feasible : expertiseexists, experts were available and willing to help, there was insufficient expertise for all needs,an algorithmic solution would be less efficient than a heuristic, and the size of the problem was not overwhelming[Waterman, 1986]. Thus, with the concurrence and willing participation of domain specialists in the Missouri Division of Employment Security, Research and Analysis Section, this project was initiated to supplement the wage rate survey system with aknowledge-based system for wage rate prediction.

## Ⅲ. FOUNDATIONS FOR THE KBS

The evolution of the Missouri KBS followed the list of tasks presented by Buchanan et al. [1983] : identification, conceptualization, formalization, implementation, and testing. The time frame fordevelopment of the Missouri wage rate KBS is shown in ⟨Figure 1⟩. An initial feasibility analysis established to our satisfaction that a knowledge-based system was indeed appropriate in thissituation, that a system could be designed and developed with reasonable expenditures of timeand effort, and that domain expertise was sufficiently available.

Over the course of five months of extensive interviews and observations, the cognitive behavior ofthe principal expert in the targeted organization was studied in detail. Unstructured interviewswere followed by written questions and structured interviews. Manuals detailing the procedures to be followed, reports prepared inresponse to several actual and hypothetical queries, and published data summaries were carefully examined [Missouri, 1989]. The resulting mental process model of the domain expert is presentedin⟨Figure 2⟩. Two primary tasks constitute the contribution of the experts, and hence also that ofour KBS : matching

- ▣ Identification (9/89-12/89)
  - Situation analysis
  - Feasibility study
- ▣ Conceptualization (1/90-4/90)
  - Knowledge acquisition techniques
  - Mental process model of expert
- ▣ Formalization (5/90-8/90)
  - AI formalism (production system)
  - Development tool selection (Prolog)
- ▣ Implementation (9/90-12/90)
  - Knowledge base construction
  - Prototyping of KBS
- ▣ Testing (1/91-3/91)
  - Test run in real environment
  - Validation and evaluation
- ▣ Maintenance

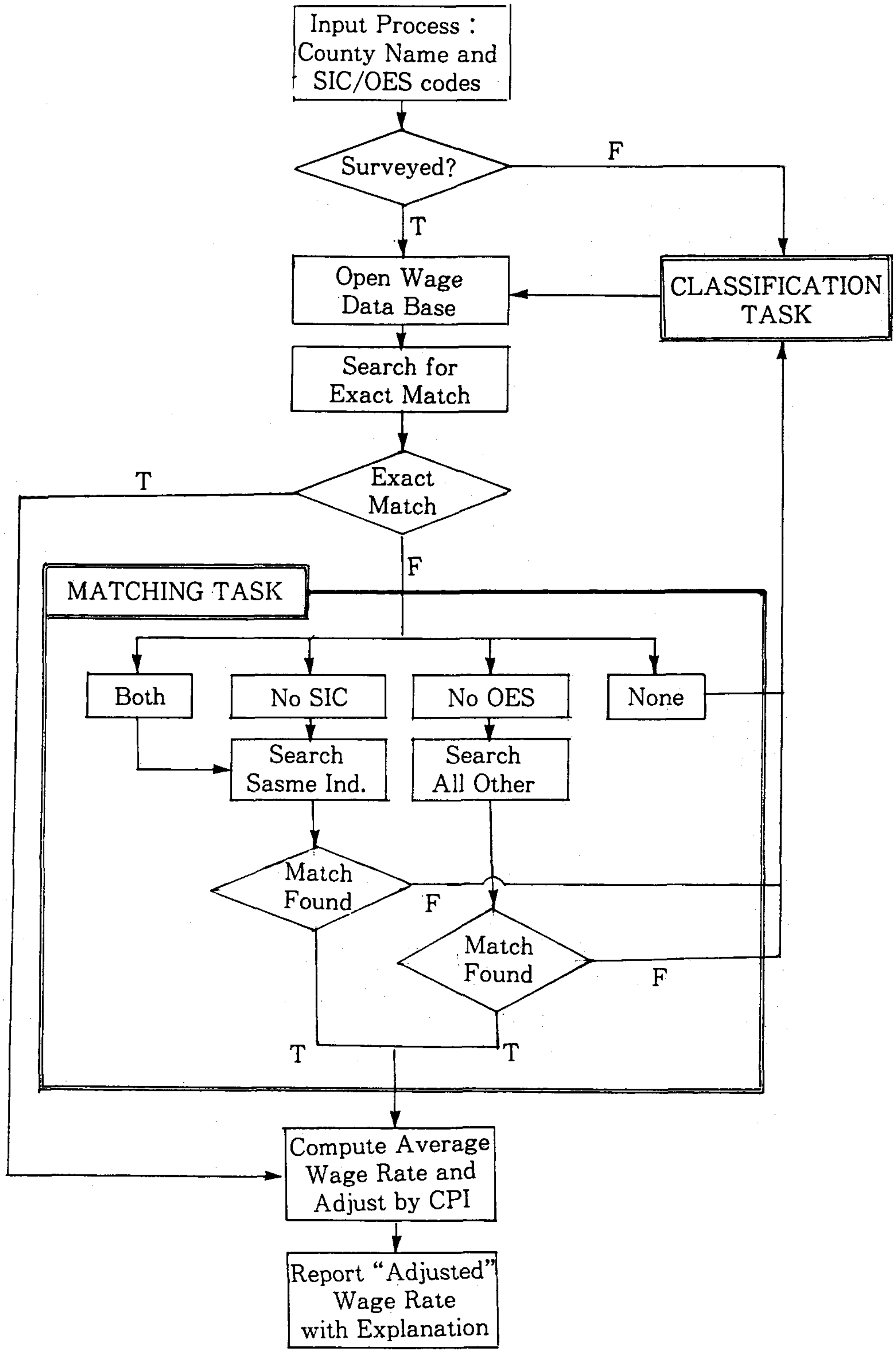
⟨Figure 1⟩ Time Frame for Development of Missouri Wage Rate KBS and similarity determination. Matching requires relatively little expertise,while similarity determination may involve considerable use thereof. A breakdown of the matching task is shown in⟨Figure 2⟩.

Similarity determination is required whenever a match cannot be found for an input county/OES/SIC combination. This may result from a request concerning a county which has notyet been surveyed, a job category which does not exist in a surveyed county, or an industrywhich does not yet exist in the county. Each of the 114 counties in Missouri is described by ten attributes (selected by the domain experts), including average weekly wage, six industry mix attributes, and three demographic attributes. (See Appendix A.)

Selecting a proper similarity measure is crucial and closely related to the type of data involved.Sneath and Sokal [1973] divided similarity measures into four groups : distance measures, association coefficients, correlation coefficients, and probabilistic similarity coefficients. Among these, association coefficients and correlation coefficients were dropped from consideration for this county similarity determination problem because the data type is not appropriate to those two measures. In addition, the Missouri wage specialist (the domain expert) suggested that the similarity measure should not only categorize similar counties but also show the degree of similarity among counties. With these two guidelines in mind, several statistical similarity measures were tested before settling on the Weighted Manhattan Distance Similarity Measure(WMDSM).

There are many measures for estimation of distance or similarity between two objects. Different measures weigh data characteristics differently. The choice among the measures should be basedon which differences or similarities in the data are important for a particular application. Adistance measure called City-block or Manhattan is selected since large differences are not weighted heavily. It is based on the absolute values of differences. However, weights were assigned to attributes according on their relative importance to reflect the expert's guidelines.

In order to avoid unduly emphasizing attributes with inherently large values (such as population, when compared with percentage change in population), attribute values were scaled before computing distances. The Z-score was used for the data transformation. Based on the transformed data, the weighted Manhattan distance similarity measure was calculated. The Manhattan distance relies on pairwise difference between the attribute values of instances (Figure 3).




〈Figure 2〉 Mental Process Model of the Domain Expert

| | 001 | 003 | 005 | 007 | 009 | 011 | 013 | 015 | 017 | 019 | 021 | 023 | | 025 | 027 | 029 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 001 | — | .89 | .82 | .76 | .88 | .88 | .83 | .81 | .72 | .70 | .68 | .96 | ·· | .88 | .78 | .84 |
| 003 | | — | .87 | .72 | .83 | .82 | .90 | .86 | .79 | .66 | .64 | .89 | ·· | .88 | .87 | .88 |
| 005 | | | | .63 | .73 | .75 | .90 | .89 | .84 | .57 | .55 | .81 | ·· | .78 | .88 | .85 |
| 007 | | | | — | .80 | .82 | .69 | .66 | .60 | .85 | .88 | .78 | ·· | .78 | .64 | .72 |
| 009 | | | | | — | .95 | .78 | .77 | .74 | .67 | .68 | .89 | ·· | .90 | .75 | .87 |
| 011 | | | | | | — | .79 | .79 | .72 | .70 | .70 | .89 | ·· | .91 | .78 | .86 |
| 013 | | | | | | | — | .94 | .82 | .62 | .62 | .86 | ·· | .86 | .90 | .91 |
| 015 | | | | | | | | — | .83 | .59 | .58 | .84 | ·· | .85 | .92 | .90 |
| 017 | | | | | | | | | — | .50 | .48 | .75 | ·· | .79 | .80 | .85 |
| 109 | | | | | | | | | | — | .88 | .70 | ·· | .66 | .58 | .59 |
| 021 | | | | | | | | | | | — | .70 | ·· | .66 | .55 | .61 |
| 023 | | | | | | | | | | | | — | ·· | .91 | .79 | .85 |
| ⋮ | | | | | | | | | | | | | ·· | ⋮ | ⋮ | ⋮ |
| 225 | | | | | | | | | | | | | | — | .81 | .89 |
| 227 | | | | | | | | | | | | | | | — | .87 |
| 229 | | | | | | | | | | | | | | | | — |

Note : The three-digit integer numbers indicate the county codes and the two-digit decimal numbers indicate the degree of similarities among counties. The largest possible degree of similarity is "1.00" and it means the two counties are identical.

〈Figure 3〉 A Sample from the Similarity Matrix Created by the Weighted Manhattan Distance Similarity Measure (WMDSM)

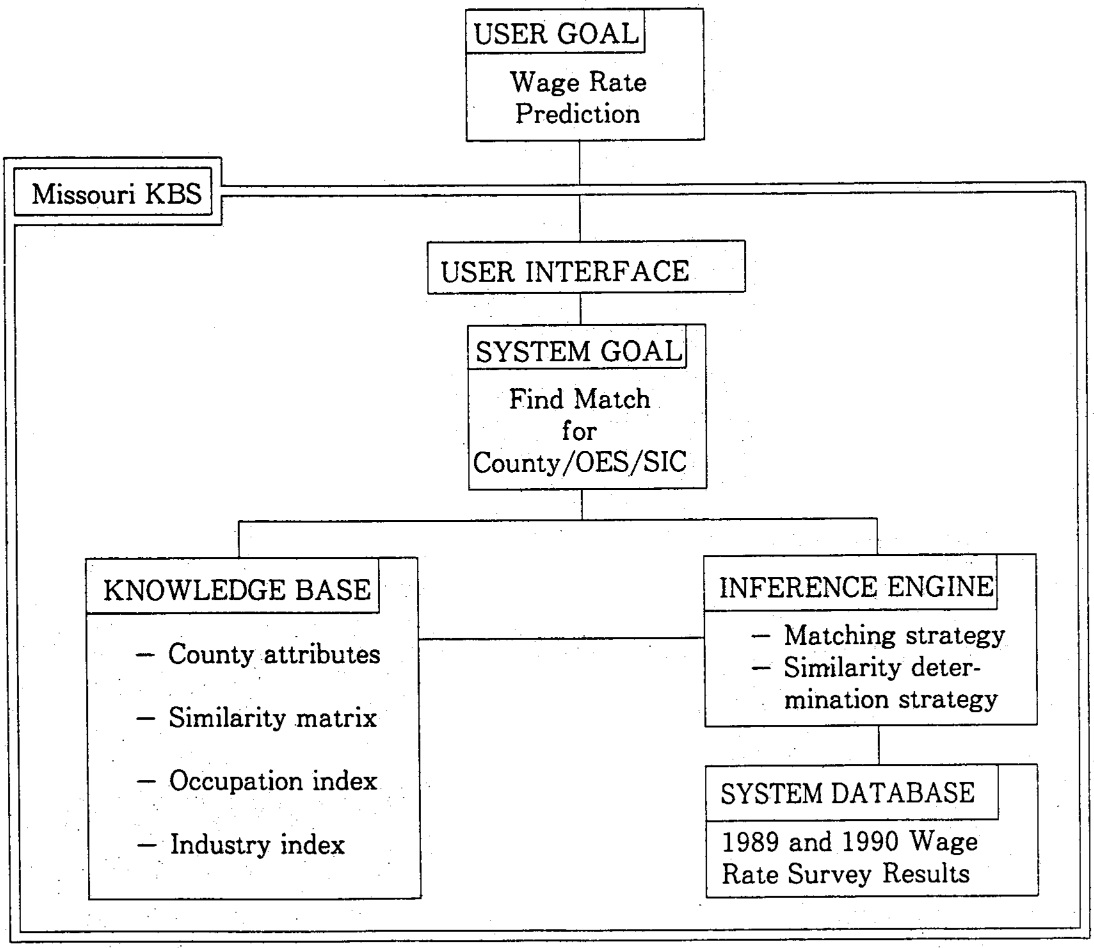
Relative weights assigned by domain expert were attached to attributes, as they tended to enhance performance.

# Ⅳ. KBS DESIGN AND DEVELOPMENT

In order to forestall the appearance of

lengthy delays, to maintain expert and user interest, and to enhance the utility of the delivered system, prototyping approach was used as the system development method. Prototyping has been shown to improve organizational acceptance, induce meaningful user feedback, and highlight errors and inadequacies, especially for systems which depend on heuristics and expertise [Naumann & Jenkins, 1982 ; Bratko, 1989]. The use of prototyping has been treated in the literature as if it were both indispensable and unavoidable in expert system development [Buchanan et al., 1983 ; Hayes-Roth et al., 1983].

The general architecture of the system is shown in⟨Figure 4⟩. The traditional distinction among knowledge base, inference engine, and user interface were retained. User access to the system is menu driven, with input prompts and screen output (see ⟨Figure 5⟩ for a sample query response).




⟨Figure 4⟩ Architecture of the Missouri KBS

```
Missouri Wage Rate Prediction System

---------- Menu ----------
1. Wage rate prediction
2. List county names
3. List occupations & OES codes
4. List industries & SIC codes
5. Exit
Enter a choice : 1
```

```
Missouri Wage Rate Prediction System

---------- Database Access Option ----------
1. County Average

2. State Average

Enter an option : 1
```

```
Missouri Wage Rate Prediction System

---------- Input Process ----------
Enter the county name : Buchanan
Enter the job type(by OES code)
  : 25105
Enter the industry type(by SIC
  code) : 3400
```

```
Missouri Wage Rate Prediction System

The job you requested is
COMPUTER PROGRAMMERS
(25105) in 3400 industry.

*** Buchanan(021) ***

No exact match found, however 25105
  exists.
Searching for the match within the Manu-
  facturing industry.

25105 2013   $15.40   1
25105 3111   $11.10   1
25105 2099   $ 6.10   1
25105 2099   $18.00   1
25105 3841   $14.30   1
25105 3841   $16.40   3
25105 3841   $20.80   1
25105 3841   $19.90   1

Average wage rate= $15.50
Number of people=10

Press any key :
```

〈Figure 5〉 Input and Output Screens : Wage Rate Prediction Task

The knowledge representation formalism adopted is the rule-based or production system. Control and search strategies are implemented as if-then-else rules. The key search strategy in this domain is illustrated in the following pseudocode rules :

| | |
|---|---|
| IF OES | code is found in wage database |
| AND SIC | code is not found in wage database |
| THEN | search for OES code within the same industry and save the matches in a list |

| | |
|---|---|
| IF OES | code is not found in wage database |
| AND SIC | code is found in wage database |
| THEN | search for SIC code with "all other" |
| OES | type and save the matches in a list |

| | |
|---|---|
| IF OES | code is not found in wage database |
| AND SIC | code is not found in wage database |
| THEN | find the surveyed county which is most similar to the county in question |

| | |
|---|---|
| IF | the list of matches is empty |
| THEN | find the surveyed county which is most similar to the county in question |
| ELSE | (compute the average of the contents of the list) |

The domain knowledge is composed of both facts and rules. Internal and external databases store the characteristics of each county (an internal database ; ⟨Figure 6⟩), OES definitions (internal ; ⟨Figure 7⟩), the

```
county("001", 1990,2, "Adair")
county("003", 1989,1, "Andrew")
county("005", 9999,1, "Atchison")
county("007", 1990,2, "Audrain")
county("009", 1990,4, "Barry")
county("011", 9999,4, "Barton")
county("013", 9999,4, "Bates")
county("015", 1989,3, "Benton")
county("017", 1989,5, "Bollinger")
county("019", 1989,3, "Boone")
   ⋮      ⋮     ⋮      ⋮
county("225", 9999,4, "Webster")
county("227", 9999,1, "Worth")
county("229", 1989,4, "Wright")
```

⟨Figure 6⟩ Contents of the County Characteristic Knowledge Base (County code, year of survey, location indicator, and county name ; File name : COUNTY.KB2)

```
job("21114", "ACCOUNTANTS AND AUDITORS")
job("53123", "ADJUSTMENT CLERKS")
job("13014", "ADMINISTRATIVE SERVICES MANAGERS")
job("85323", "AIRCRAFT MECHANICS")
job("97702", "AIRCRAFT PILOTS AND FLIGHT ENGINEERS")
job("79017", "ANIMAL CARETAKERS, EXCEPT FARM")
job("34021", "ANNOUNCERS, EXCEPT RADIO AND TELEVISION")
job("34017", "ANNOUNCERS, RADIO AND TELEVISION")
job("43011", "APPRAISERS, REAL ESTATE")
job("22302", "ARCHITECTS, EXCEPT LANDSCAPE AND MARINE")
job("34035", "ARTISTS AND RELATED WORKERS")
 :    :        :          :
job("91705", "WELDING MACHINE OPERATORS AND TENDERS")
job("89308", "WOOD MACHINISTS")
job("34002", "WRITERS AND EDITORS")
```

〈Figure 7〉 Contents of Occupational Employment Statistics(OES) Definition Knowledge Base (File name : OES-DEF.KB)

calculated similarity table (internal ; 〈Figure 8〉), and instances of each wage (external ; 〈Figure 9〉). 〈Figure 10〉 depicts the conversion process of the Missouri data into the system knowledge base and database. The filename extensions "PAS" and "PRO" denote Pascal programs and Prolog programs, respectively.

Prolog was chosen as the system development tool because it proved far more flexible than an expert system shell. Among the aspects where Prolog excels are data-handling (internal and external databases), searching (symbolic match), computational power (numeric processing), and ease of rule modification (declarative expressions) [Lazarev, 1989 ; Bratko, 1990 ; Borland, 1988 ; Genesereth & Ginsberg, 1985]. Prolog has its own inference (backward-chaining) and search mechanisms [Sterling & Shapiro, 1986], like a shell, but unlike shells it also has the computational power of first-order logic [Lloyd, 1984]. However, Prolog is no panacea ; it does not have some fea-

```
sim("001", ["023", ".96", "187", ".93", "195", ".92", …, "015", ".81"])
sim("003", ["083", ".93", "109", ".93", "069", ".92", …, "041", ".84"])
sim("005", ["185", ".94", "041", ".93", "181", ".91", …, "069", ".84"])
sim("007", ["071", ".94", "186", ".94", "099", ".93", …, "055", ".72"])
sim("009", ["195", ".93", "157", ".92", "105", ".92", …, "186", ".82"])
sim("011", ["009", ".95", "157", ".95", "065", ".94", …, "003", ".82"])
sim("013", ["033", ".96", "015", ".94", "087", ".94", …, "023", ".86"])
sim("015", ["033", ".96", "055", ".93", "059", ".93", …, "043", ".86"])
sim("017", ["223", ".95", "153", ".92", "123", ".91", …, "083", ".80"])
sim("019", ["175", ".90", "099", ".89", "077", ".89", …, "087", ".62"])
sim("021", ["175", ".91", "077", ".89", "007", ".88", …, "125", ".62"])
sim("023", ["001", ".96", "187", ".95", "195", ".95", …, "063", ".84"])
 ⋮     ⋮         ⋮           ⋮            ⋮             ⋮
sim("225", ["105", ".95", "043", ".94", "069", ".93", …, "015", ".85"])
sim("227", ["015", ".92", "033", ".92", "211", ".92", …, "203", ".84"])
sim("229", ["055", ".97", "215", ".97", "125", ".96", …, "211", ".86"])
```

⟨Figure 8⟩ Contents of the County Similarity Knowledge Base(File name : SIMIL.KB2)

```
wg("7261", "13002" 15.50,1)
wg("7311", "13002", 11.50,1)
wg("6022", "13002", 11.50,1)
wg("6022", "13002", 23.08,1)
wg("6022", "13002", 36.00,1)
wg("6062", "13002", 10.50,1)
wg("6062", "13002", 15.50,1)
wg("6411", "13002", 14.50,1)
wg("8331", "13002", 15.50,1)
wg("6036", "13002",  9.50,1)
wg("2013", "13005", 26.00,1)
 ⋮      ⋮         ⋮       ⋮
wg("4210", "98999",  4.75,4)
wg("4210", "98999",  5.25,5)
wg("4210", "98999",  5.75,1)
```

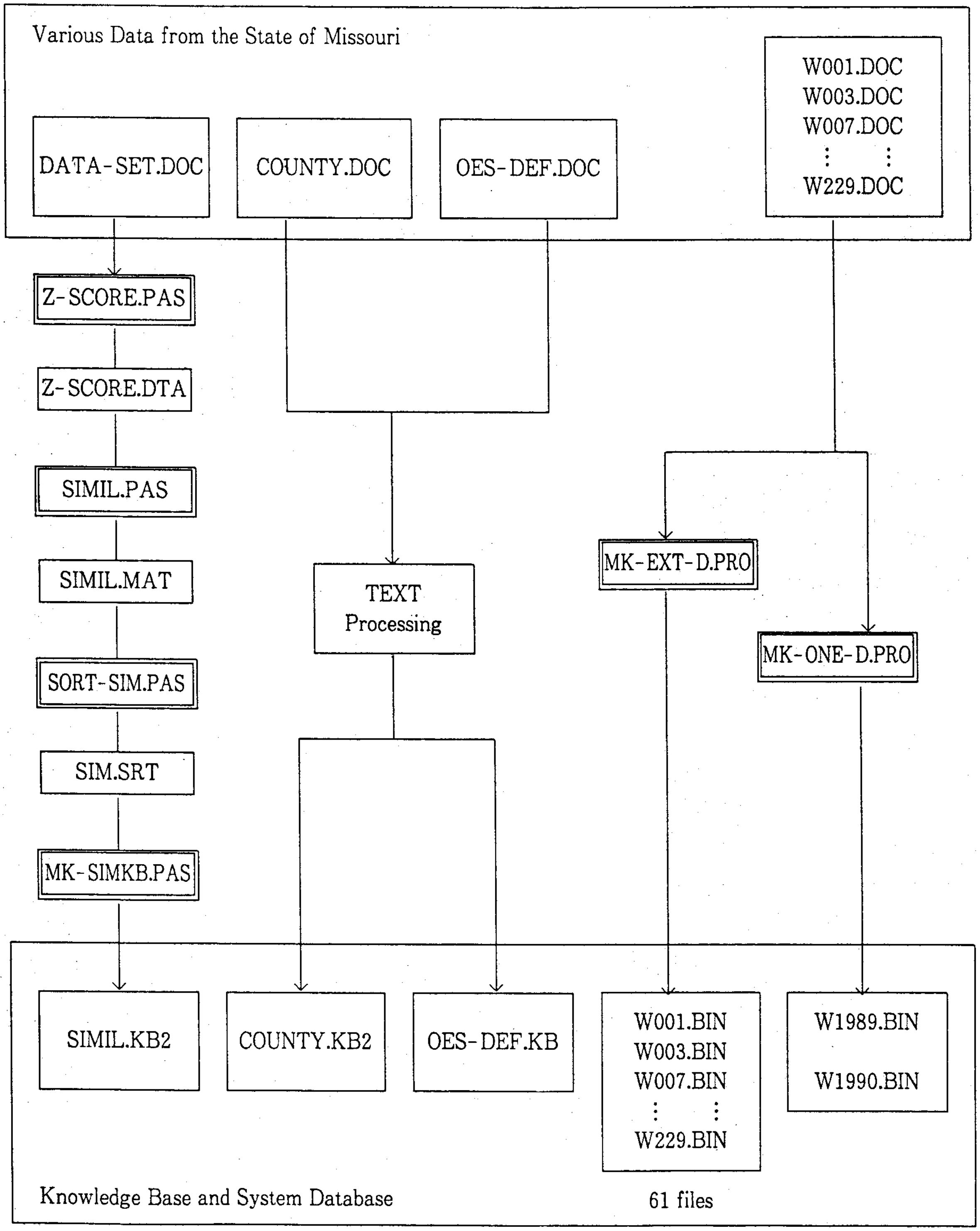
⟨Figure 9⟩ Contents of a County's Wage Database

Database name : WAGES

Chain name : chain019

County Name : Boone (Pop. : 77,702)

Database size : 7,098 instances

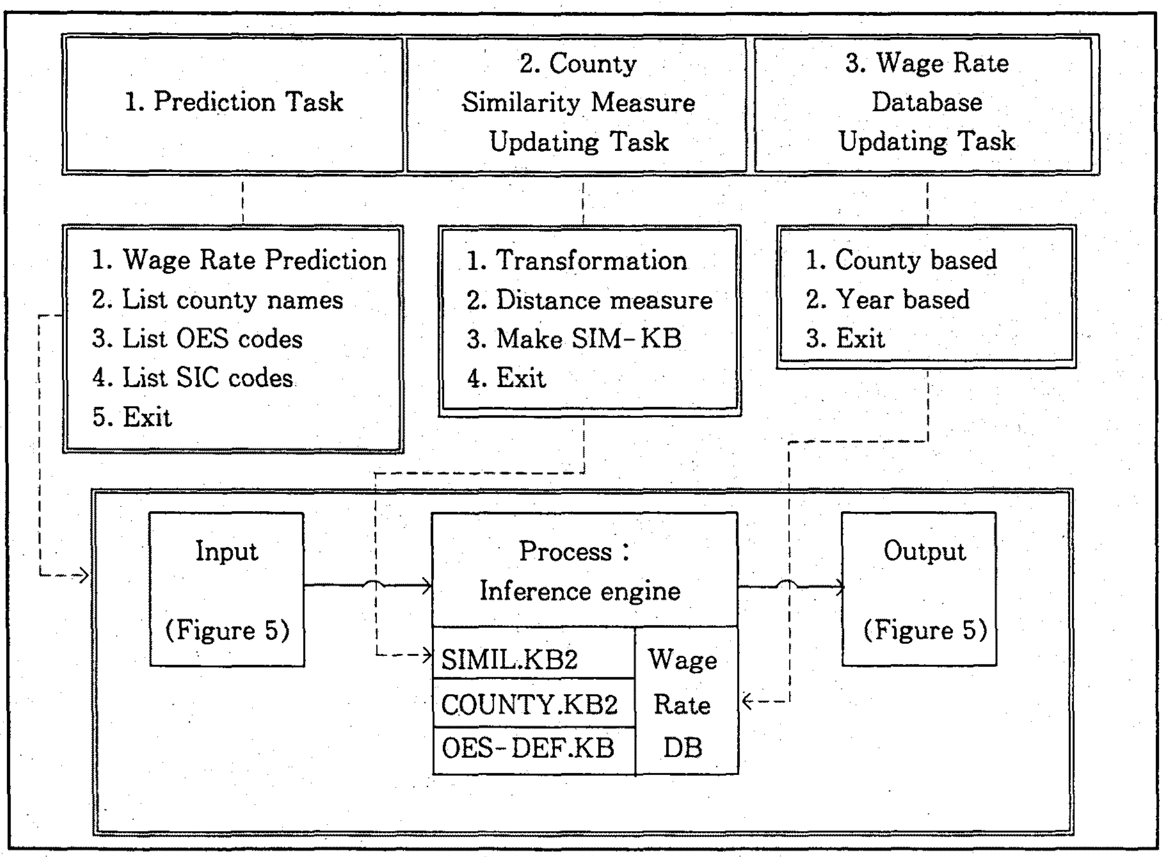



⟨⟨Figure 10⟩ Conversion Process of the Missouri Data into the System Knowledge Base and Database tures that a shell might have, such as the ability to handle uncertain information or generate explanations [Jackson, 1986]. Various authors have argued that Prolog should be used more widely for expert systems [Sterling & Shapiro, 1986; Smith, 1988]. They cite some of the aforementioned advantages and suggest that expert systems written in Prolog require less code and execute more quickly than systems written in other languages.

The system consists of three modules: a prediction task module, a county similarity measure updating task module, and a wage rate database updating task module. The main module is the prediction task module, while the remaining two modules are added to facilitate the system maintenance process. ⟨Figure 11⟩ provides an overview of the system.

## V. VALIDATION OF THE SYSTEM

Validations of knowledge-based systems are rarely published. There are many




⟨Figure 11⟩ Overview of the Missouri KBS dimensions by which a knowledge-based system might be judged. Barr et al.[1989, p. 177] considered computational, psychological "look and feel," and performance as the three most important dimensions. Feigenbaum [1989] claimed that the "competence/performance" dimension is receiving the most attention among the various dimensions of machine intelligence. Laswell[1990] mentioned that the best standard for KBS testing is the past performance of human decision makers. Liang [1986] accumulated 19 factors affecting system implementation ; his statistical analysis showed that accuracy was rated as the most important.

These evaluation criteria guided the formal validation process of this research. Among several system performance measures, the validation study here focused strictly on measuring system accuracy. This is not only because accuracy was considered by Liang as the most important factor, but also because the other factors (such as motivation and user attitude) would require a long-term study of the system in an organization, which was beyond the scope of this study. The validation process involves two initial tests and two main system performance tests.

## Initial Test:Database Conversion Accuracy

The purpose of this test is simple : to examine whether the wage rate database (in ASCII format) from Missouri was correctly transformed to the KBS database (in a chained external database in Prolog predicate format ; as in ⟨Figure 9⟩) because data conversion on such a scale might involve errors. This preliminary test is an essential prerequisite for the subsequent tests.

Fifteen jobs were tested and the result was outstanding (see Table 1). The minimal deviation between reported values and system-produced values is due to rounding errors. The system can handle more specific industry code input than was employed in this test. The system would constitute a better decision aid in this respect. Based on the results of this initial test, it was determined that the database was properly converted and no data was lost or altered during the conversion process.

## County Similarity Measure Validation

This second initial test concerns the county similarity determination task, which is one of the key tasks of the system.

(Table 1) Result of the Initial Test

| Selected jobs | | | Wage rate from reports | Wage rate found by the KBS |
|---|---|---|---|---|
| County | Job title | Industry | | |
| Franklin and Jefferson (1990) | Electricians (87202) | Manuf. | $ 10.81 (53) | $ 10.70 (43)<br>$ 13.10 (10) |
| | Secretaries (55100) | Manuf. | $ 7.78 (78) | $ 7.58 (55)<br>$ 8.27 (23) |
| | Marketing Mgr (13011) | Trade | $ 31.56 (4) | –<br>$ 31.60 (4) |
| Jasper and Newton (1990) | Cashiers (49023) | Trade | $ 5.23 (432) | $ 4.66(202)<br>$ 5.73 (78) |
| | Tellers (53102) | Finance | $ 5.62 (190) | $ 5.38(112)<br>$ 5.94 (78) |
| | Housekeepers (61008) | Services | $ 4.95 (63) | $ 4.24 (24)<br>$ 5.38 (39) |
| Carroll and Saline (1990) | First Line Sup (81008) | Manuf. | $ 12.79 (85) | $ 10.80 (40)<br>$ 14.56 (45) |
| | Welders/Cutters (93914) | Manuf. | $ 7.99 (168) | $ 7.60(132)<br>$ 9.43 (36) |
| | Pharmacists (32517) | Trade | $ 17.42 (3) | –<br>$ 17.42 (3) |
| Barry and Lawrence (1990) | Stock Clerks (49021) | Trade | $ 6.08 (108) | $ 4.89 (46)<br>$ 6.96 (36) |
| | Data Entry Keyr (56017) | Finance | $ 5.61 (14) | $ 4.43 (6)<br>$ 6.50 (8) |
| | Pharmacists (32517) | Trade | $ 17.15 (10) | $ 14.50 (4)<br>$ 18.92 (6) |
| Christian and Greene (1989) | Cost Estimators (21902) | Manuf. | $ 12.07 (15) | $ 9.70 (7)<br>$ 14.14 (8) |
| | System Analysts (25102) | Manuf. | $ 13.68 (7) | $ 11.34 (2)<br>$ 14.61 (5) |
| | Purchasing Mgrs (13008) | Trade | $ 14.10 (22) | $ 12.50 (10)<br>$ 15.43 (12) |

Note : Values in (  ) in the last two columns represent number of occurrences.

Although the overall performance of the system will determine the accuracy of the similarity measure (WMDSM), this test was performed to demonstrate the reliability of the measure before its implementation within the system.

⟨Figure 8⟩ shows an extract of the knowledge base of similarity relationships among counties. This knowledge base contains not only a listing of the most similar surveyed counties for each county but also the degrees of similarity. This extra information was deemed useful to increase user acceptance of the predicted values.

The similarity relationships determined by the WMDSM were reviewed by the Missouri experts, who confirmed their general validity. Diagrams were used instead of tables to facilitate the evaluators' understanding of the material. ⟨Figure 12⟩ shows a sample of the diagrams which were sent to the Missouri department. Each diagram contained several pieces of information : (1) identification of a surveyed county, (2) its attribute values, (3) a list of similar counties and their degrees of similarity, and (4) similar counties' attribute values. With the Missouri department's confirmation, the new similarity measure was determined to be valid for the domain of county similarity determination.

## Mental Process Model Validation

The purpose of this test is to validate the mental process model of the domain expert, which is shown in⟨Figure 2⟩. Also, if the test result is positive, one can conclude that not only the domain knowledge is valid but it is also correctly represented within the system. The majority of the domain knowledge consists of rules, priorities, and weights concerning the matching task and the county similarity determination task.

Recent survey result (1991 Wage Survey : Prepared by Research and Analysis Section, Missouri Division of Employment of Security) were utilized as a hold-out sample to validate the system performance. These hold-out sample involve four counties including Mississippi, New Madrid, Scott, and Stoddard.

For testing purposes, fifty samples from the survey reports were randomly selected. In selecting the samples, an effort was made to include jobs from all five industries (Manufacturing, Wholesale/Retail Trade, Services Finance/Insurance/Real Estate, an Transportation/Communications). A part of the results of the comparison test between the hold-out samples' published wage rates and the system's predicted values are listed in(Table 2). The result was sent to the Mis-

| Similarity | County | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 001 | 275.53 | 12 | 4 | 30 | 3 | 28 | 23 | 15772 | 67 | 3.7 |
| .96 | 023 | 277.17 | 15 | 4 | 29 | 3 | 21 | 28 | 24418 | 63 | 8.6 |
| .93 | 187 | 273.44 | 21 | 3 | 24 | 3 | 20 | 28 | 29069 | 68 | 8.2 |
| .89 | 069 | 262.71 | 25 | 4 | 24 | 4 | 19 | 24 | 20394 | 59 | 7.8 |
| .89 | 083 | 261.20 | 24 | 5 | 27 | 3 | 16 | 25 | 11627 | 59 | 8.4 |
| .89 | 57 | 287.23 | 31 | 5 | 26 | 4 | 12 | 22 | 10129 | 60 | 5.4 |
| .89 | 003 | 259.93 | 1 | 6 | 25 | 3 | 19 | 46 | 9614 | 65 | 5.5 |
| .88 | 043 | 262.57 | 36 | 2 | 24 | 4 | 10 | 24 | 19430 | 67 | 5.3 |
| .88 | 105 | 268.61 | 45 | 3 | 25 | 4 | 10 | 13 | 17213 | 64 | 7.9 |
| .87 | 091 | 254.56 | 32 | 4 | 24 | 3 | 19 | 18 | 18658 | 61 | 6.6 |
| .87 | 029 | 247.12 | 10 | 3 | 30 | 6 | 32 | 19 | 17708 | 67 | 6.1 |
| .86 | 049 | 256.24 | 5 | 3 | 25 | 8 | 24 | 35 | 10520 | 63 | 6.3 |
| .84 | 087 | 251.84 | 0 | 7 | 34 | 5 | 17 | 37 | 3603 | 56 | 5.8 |
| .84 | 155 | 245.11 | 21 | 4 | 23 | 3 | 17 | 32 | 13677 | 58 | 7.9 |

〈Figure 12〉 Counties, Attributes, and Their Similarity Relationships

(Table 2) Result of Comparison Study between Hold-out Samples'(from 1991 Wage Rate Survey) Published Value and the System's Predicted Value

| Selected Job | | | 1991 Data | Predicted Wage Rate by the system | | |
|---|---|---|---|---|---|---|
| County | Job | Ind | | Wages | Similar County | WMDSM |
| Mississippi | 21902 | 2000 | $ 15.32 (14) | $ 6.44 (1) | Crawford | 0.95 |
| New Madrid | | | | 15.00 (6) | Buchanan | 0,90 |
| Scott | | | | 10.80 (1) | Phelps | 0.93 |
| Stoddard | | | | 13.90 (2) | Saline | 0.96 |

| Selected Job | | | 1991 Data | Predicted Wage Rate by the system | | |
|---|---|---|---|---|---|---|
| County | Job | Ind | | Wages | Similar County | WMDSM |
| Mississippi | 13011 | 2000 | 12.68 (8) | 13.20 (5) | Pemiscot | 0.95 |
| New Madrid | | | | 22.40 (1) | Buchanan | 0.90 |
| Scott | | | | 19.00 (1) | Phelps | 0.93 |
| Stoddard | | | | 5.71 (8) | Saline | 0.96 |
| Mississippi | 81008 | 2000 | 11.01 (153) | 10.10 (23) | Crawford | 0.95 |
| New Madrid | | | | 13.80(130) | Buchanan | 0.90 |
| Scott | | | | 12.00 (89) | Pettis | 0.95 |
| Stoddard | | | | 13.40 (80) | Saline | 0.96 |
| Mississippi | 34017 | 4000 | 4.89 (15) | 9.02 (2) | Pemiscot | 0.95 |
| New Madrid | | | | 9.37 (9) | Buchanan | 0.90 |
| Scott | | | | 6.34 (6) | Pettis | 0.95 |
| Stoddard | | | | 4.06 (4) | Laclede | 0.96 |
| Mississippi | 97102 | 4000 | 11.25 (359) | 14.90 (1) | Crawford | 0.95 |
| New Madrid | | | | 12.50(123) | Buchanan | 0.90 |
| Scott | | | | 10.60 (15) | Pettis | 0.95 |
| Stoddard | | | | 14.50 (54) | Saline | 0.96 |
| Mississippi | Q3008 | 5000 | 8.50 (10) | 6.63 (1) | Pemiscot | 0.95 |
| New Madrid | | | | 10.70 (9) | Buchanan | 0.90 |
| Scott | | | | 12.90 (1) | Pettis | 0.95 |
| Stoddard | | | | 10.20 (2) | Barry | 0.94 |
| Mississippi | 85321 | 5000 | 7.75 (14) | 7.16 (4) | Pemiscot | 0.95 |
| New Madrid | | | | 8.55 (9) | Buchanan | 0.90 |
| Scott | | | | 8.56 (9) | Perry | 0.92 |
| Stoddard | | | | 8.75 (3) | Saline | 0.96 |
| Mississippi | 13002 | 6000 | 19.74 (19) | 13.20 (5) | Permiscot | 0.95 |
| New Madrid | | | | 22.40 (13) | Buchanan | 0.90 |
| Scott | | | | 19.00 (1) | Phelps | 0.93 |
| Stoddard | | | | 12.71 (2) | Saline | 0.96 |
| Mississippi | 53102 | 6000 | 5.78 (136) | 3.98 (45) | Crawford | 0.95 |
| New Madrid | | | | 7.50 (36) | Buchanan | 0.90 |
| Scott | | | | 7.40 (24) | Pettis | 0.95 |
| Stoddard | | | | 6.10 (37) | Saline | 0.96 |
| Mississippi | 27302 | 7000 | 11.73 (13) | 10.02 (1) | Pemiscot | 0.95 |
| New Madrid | | | | 13.32 (8) | Buchanan | 0.90 |
| Scott | | | | 12.33 (7) | Pettis | 0.95 |
| Stoddard | | | | 9.02 (5) | Laclede | 0.96 |
| Mississippi | 85305 | 7000 | 10.33 (6) | 7.63 (1) | Pemiscot | 0.95 |
| New Madrid | | | | 10.70 (4) | Buchanan | 0.90 |
| Scott | | | | 12.90 (5) | Pettis | 0.95 |
| Stoddard | | | | 10.20 (2) | Barry | 0.94 |

Note : Source of the 1991 data is 1991 Wage Rate in Selected Occupations for Mississippi, New Madrid, Scott, and Stoddard Counties.

souri department for evaluation ; they were very satisfied with the system's performance. It was thus determined that not only was the mental process model of the Missouri expert appropriate, but also that it was correctly represented in the system.

## Performance Comparison : Systems with Different Database Size

The purpose of this portion of the study is to test the hypothesis forming the foundation of this-project. At the beginning of the project, only thirty three counties had been surveyed. It was assumed that as the wage database grew, the values predicted by the system would become more accurate and reliable because the matches would be retrieved from the databases of more similar surveyed county. Currently, the system contains wage rate databases from sixty one counties.

An experiment was conducted to test the above hypothesis. The experiment involved two systems : one with 33 wage rate databases (1990 version) and the other with 61 wage rate databases (1991 version). The same input (county/job code/industry code combination) was used on both systems and the results were accumulated. (Table 3) shows these results, which are decidedly indicatory of enhanced system performance as the database expands. The two sets of predicted values were compared with the hold-out sample counties' reported values. As shown in Table 3, the predicted values from the system with the larger database were found to be closer to the reported values. The table also shows the degree of similarity (WMDSM) of the county from which the predicted-values were retrieved. Generally, the values of the degree of similarity from the larger database system are higher. Thus, it can be concluded that as the database expands the accuracy of the-predicted values will be enhanced.

## Psychological and Computational Issues

Psychological issues include ease of use, understandability, and explanation capability. Although these system attributes were not tested, previous research (such as Bratko, 1989) on the prototyping approach indicates that systems developed using the prototyping approach tend to have higher organizational acceptance.

To date, the system has been sent to the users four times. Feedback from the users led to several important adjustments and corrections to the system, including (1) wage rate database update, (2) county similarity measure update, (3) database search

(Table 3) Performance Comparison between Two Systems with Different Database Size

| Selected jobs | | | Predicted Values | | | |
|---|---|---|---|---|---|---|
| County | Job Code | SIC Code | 1989 System(33 DB) | | 1990 System(61 DB) | |
| | | | County(WMDSM) | Wage/hr | County(WMDSM) | Wage/hr |
| Barry (1990) | 32517 | 5084 | Laclede(0.92) | $ 19.90 | Barry(1.00) | $ 17.30 |
| | 49021 | 5084 | Laclede(0.92) | 9.67 | Barry(1.00) | 5.98 |
| | 56017 | 6023 | Laclede(0.92) | 5.39 | Barry(1.00) | 5.41 |
| Carroll (1990) | 32517 | 5411 | Benton(0.96) | 18.00 | Barry(1.00) | 18.50 |
| | 81008 | 2421 | Benton(0.96) | 11.10 | Barry(1.00) | 9.48 |
| | 93914 | 2352 | Benton(0.96) | 6.10 | Benton(0.96) | 6.10 |
| Franklin (1990) | 13011 | 5084 | Ste. Genev(0.95) | 9.39 | Jasper(0.96) | 12.70 |
| | 55100 | 6023 | Perry(0.86) | 7.80 | Franklin(1.00) | 7.63 |
| | 87202 | 2352 | Ste.Genev(0.95) | 9.39 | Franklin(1.00) | 10.70 |
| Jasper (1990) | 49023 | 5411 | Ste.Genev(0.92) | 7.27 | Jasper(1.00) | 6.60 |
| | 53102 | 6023 | Ste.Genev(0.92) | 6.74 | Jasper(1.00) | 5.81 |
| | 61008 | 7539 | Ste.Genev(0.92) | 4.08 | Jasper(1.00) | 5.22 |
| Jefferson (1990) | 13011 | 5084 | Greene(0.95) | 19.60 | Jefferson(1.00) | 32.50 |
| | 55100 | 6023 | Perry(0.80) | 7.80 | Jefferson(1.00) | 7.84 |
| | 87202 | 2352 | Greene(0.95) | 13.10 | Jefferson(1.00) | 13.10 |
| Lawrence (1990) | 32517 | 5411 | Dunklin(0.94) | 16.40 | Lawrence(1.00) | 20.60 |
| | 49021 | 5084 | Henry(0.95) | 5.03 | Lawrence(1.00) | 8.03 |
| | 56017 | 6023 | Dunklin(0.94) | 6.63 | Lawrence(1.00) | 6.05 |
| Newton (1990) | 49023 | 6023 | Laclede(0.86) | 4.27 | Laclede(0.86) | 4.27 |
| | 53102 | 6023 | Perry(0.91) | 5.77 | Newton(1.00) | 5.73 |
| | 61008 | 7539 | Perry(0.91) | 5.15 | Newton(1.00) | 4.59 |
| Saline (1990) | 32517 | 5411 | Butler(0.95) | 20.10 | Saline(1.00) | 19.30 |
| | 81008 | 2421 | Butler(0.95) | 8.75 | Saline(1.00) | 13.40 |
| | 93914 | 2352 | Butler(0.95) | 10.10 | Saline(1.00) | 9.62 |
| Cass (9999) | 13011 | 5084 | Perry(0.92) | 17.50 | Dent(0.94) | 6.44 |
| | 55100 | 6023 | Perry(0.92) | 7.80 | Dent(0.94) | 6.44 |
| | 87202 | 2352 | Butler(0.93) | 11.10 | Dent(0.94) | 12.40 |
| Clay (9999) | 32517 | 5411 | Buchanan(0.80) | 17.60 | Buchanan(0.80) | 17.60 |
| | 49021 | 5084 | Buchanan(0.80) | 5.44 | Buchanan(0.80) | 5.44 |
| | 56017 | 6023 | Buchanan(0.80) | 5.90 | Buchanan(0.80) | 5.90 |

strategy, (4) detection of errors, (5) adaptation of the input/output format to more closely match user preferences, and (6) addition of updating functionality. It is expected that the high (Table 3) Performance Comparison between Two Systems with Different Database Size degree of meaningful user feedback, and its consideration in system development, will in turn produce a higher user acceptance rate. Among several system features, the users indicate that they liked the output contents most (see ⟨Figure 5⟩). The current report shows simply an average wage rate of a certain job in a certain industry located in a certain county. This information is not enough to serve various types of inquiries and it does not provide accurate wage rates. However, the system provides the user with a list of information (the job code, industry code, hourly wage rate, and number of people receiving the wage rate) for each match. It also identifies which county's database was used to find the match and the degree of similarity between the two counties. This list helps the user to see a broader picture of the labor market, thereby providing more accurate information to their customers.

It is expected that the users' productivity will improve with the help of this system. According to one of Missouri wage specialists, any one request for wage information can take several hours to answer and the amount of time spent to determine wages for all users is often the majority of any week. In case of a county which has not yet been surveyed, the analyst must choose a similar county in order to respond to the inquiry. This process of determining the most similar county was not systematic. The system selects similar counties automatically and logically. Customer inquiries often involve requests for multiple wage quotes. With this computerized system, users can find appropriate wage rates for various jobs, wherever they locate and to whatever industry those jobs belong, within a matter of seconds.

Computational issues, according to Barr et al. [1989], involve memory required, extensibility, and portability. The current system requires 179 KB of compiled memory space and 2.24 MB of database space. The database is expected to grow up to 4 MB when the results of 1991 and 1992 survey are added. In an effort to test the system's extensibility, the system was reviewed by wage specialists of Nebraska Department of Labor. They expressed that the system has potential in the state of Nebraska. The wage rate prediction system for Nebraska can be completed within a week once they accept the idea of using Missouri wage rate database. Until Nebraska come up with their own wage rate survey result (actually Nebraska started the survey in 1991), the Nebraska wage rate prediction system with Missouri wage rate database can be a viable option.

The portability was an important issue in this project. Like most practical system development projects, the portable or stand-alone system seemed appropriate since the Missouri Department of Labor and Industrial Relations (which initiated this project) as well as other Missouri government agencies also expect to have this facility within their office.

# Ⅵ. CONCLUSION

The knowledge-based system developed for this study is intended for use in the State of Missouri's Department of Labor and Industrial Relations. This department is often contacted by organizations requesting information on actual or hypothetical wages for workers in various occupational categories. These requests are related to the geographical location of the position. The current wage rate survey approach has numerous deficiencies, particularly relative to counties which have not been surveyed.

There are three main themes throughout this study. First, it provides conceptual frameworks which are based on AI methodologies. Second, it presents a system architecture which has been designed in order to serve the decision making process effectively. Third, it describes a functional wage rate prediction system and the validation study of the system.

## Contributions of the Study

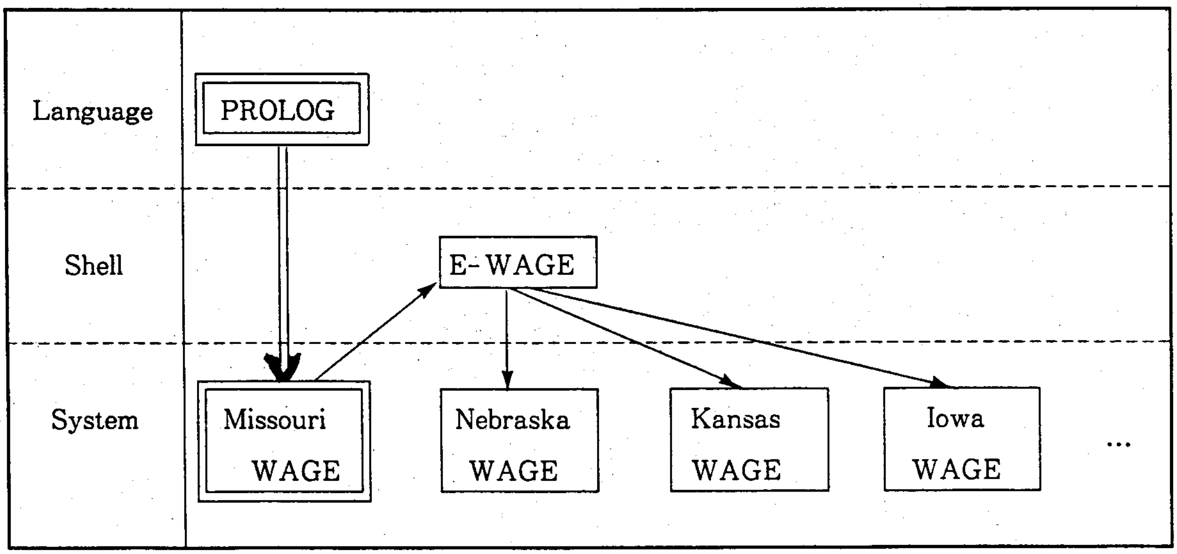
The contributions of this study are a framework for the business application of AI technology, a knowledge-based system architecture to articulate design and implementation issues, a functional system implementation to demonstrate the applicability of the wage rate prediction system. This KBS is the first computerized wage rate prediction system and contains a vast amount of information concerning wage rate. The system includes virtually all occupations and their surveyed wage rates in the state of Missouri. Such wage rate information is vital for those who are considering opening shops, clinics, or manufacturing facilities ; they provide at least a general idea of the labor cost portion of the total investment. Also, potential employees, especially those who are foreign to a location with job openings, can use this system to see the market prices of their skills.

This system has already engendered interest by analysts in the Nebraska Department of Labor; two domain experts in Lincoln served as knowledge sources and assisted in testing the knowledge base. As a result, we anticipate "emptying" the Missouri system of its knowledge, forming an E-Wage shell (on the lines of EMYCIN-see [Barr & Feigenbaum, 1982]). Derivative systems would follow, as shown in ⟨Figure 13⟩.

## Limitations of the Study

To be categorized as an intelligent machine, a system should have various competences (such as performance, learning, generality, and common sense). Among these the performance dimension is currently receiving the most attention [Feigenbaum, 1989]. The Missouri KBS is such a system that performs well. However, other intelligence dimension such as learning could have been added to the system. AI researchers are working on adding such dimensions to system architectures.

Since this Missouri system were developed under PC environment, the programming effort to control memory overflow were most cumbersome. Some difficulties were encountered due to the selection of Turbo Prolog, which is a nonstandard version of Prolog lacking standard predicates. However, Turbo Prolog includes a compiler,




⟨Figure 13⟩ Wage Rate Prediction System Development Tool and the Possible Extensions enabling development of a stand-alone system for increased portability.

The validation study examined only one performance measure : prediction accuracy. As the system is used in the actual working environment, behavioral aspects of system limitations can be discovered.

## Future Research Directions

Based on the results of the current research, two major research directions can be identified. The first direction involves extending the implementation of the KBS. For the short term, the focus will be on developing other functions (such as on-line help for county names, descriptions of occupations, industry classification, and so on). The inclusion of such features has been postponed since the main users of the system are wage rate specialists in the Missouri department, who are already adequately acquainted with their problem domain. These are essential features, however, if the system is to be placed in the public domain, because the general populace is not familiar with the decimal codes used in the system's input process. Since the program is coded in modules, the input process conversion would involve only a small amount of time and effort.

For the longer term the focus will be on the application of the system to other states for use in their wage rate prediction processes. States such as Iowa, Nebraska, Arkansas, and Kansas would be appropriate targets since these states have demographic and economic characteristics similar to those of the state of Missouri. This system's extendibility or generalized applicability is excellent. As far as midwestern states are concerned, the wage rate experts presume, one would not have to perform the time-consuming wage rate survey to build the database portion of the system. They suggest matching each area (county-wise) with similar counties in Missouri using the ten attributes. With some degree of adjustments in the predicted wage rate or even without adjustments, some idea of approximate wage rates for a given occupation can be obtained. The new system development process would first involve "emptying" the Missouri system of its knowledge, forming an E-Wage Shell. The main system architecture and wage rate database will be maintained. Only the state-specific knowledge base of the system will be changed according to the target state's attributes.

This paper has taken a few first steps with regard to the theoretical aspect by providing a framework for organizing the wage rate prediction structure, and the technical aspect by defining an architecture. These initial steps will hopefully motivate more research in this area so that the KBS concept can achieve its potential in solving high-level problems and problems involving complex information processing.

# 참 고 문 헌


Barr, A., and Feigenbaum, E.A. (ed.) *The Handbook of Artificial Intelligence*, Vol. 2, Addison-Wesley, Reading, MA, 1982.

Barr, A., P.R. Cohen, and E.A. Feigenbaum. (ed.) *The Handbook of Artificial Intelligence*, Vol. 4, Addison-Wesley, Reading, MA, 1989.

Borland International. *Turbo Prolog 2.0: Reference Guide, User's Guide, and Toolbox*, Scotts Valley, CA, 1988.

Bratko, I. "Fast Prototyping of Expert Systems Using Prolog," in *Topics in Expert System Design*, G. Guida and C. Tasso (eds.), Elsevier Science, Amsterdam, Holland, 1989, pp. 69-86.

Bratko, I. Prolog *Programming for Artificial Intelligence*, 2nd Ed., Addison-Wesley Publishing Company, Reading, MA, 1990.

Buchanan, B.G. et al. "Constructing an Expert System," in *Building Expert Systems*, F. Hayes-Roth, D. Waterman, and D. Lenat (eds.), Addison-Wesley, New York, 1983, pp. 127-167.

Feigenbaum, E., "What Hath Simon Wrought?," in *Complex Information Processing: The Impact of Herbert A. Simon*, D. Klahr and K. Kotovsky(eds.), Lawrence Erlbaum, Hillsdale, New Jersey, 1989, pp. 165-182.

Genesereth, M.R., and M.L. Ginsberg. "Logic Programming," *Communications of the ACM* (28:9), September 1985, pp. 933-941.

Hayes-Roth, Frederick, Donald A.Waterman, and Douglas B.Lenat *Building Expert Systems.* Reading, MA : Addison-Wesley, 1983.

Jackson, P. *Introduction to Expert Systems (2nd ed.)*, Addison-Wesley, Reading, MA, 1990.

Laswell,L.K. "Expert Systems Testing," *Expert Systems Planning/Implementation/Integration*(2:1), Spring 1990, pp. 19-24.

Lazarev, G. *Why Prolog? Justifying Logic Programming for Practical Applications*, Prentice-Hall, Englewood Cliffs, NJ, 1989.

Liang, T.P. "Critical Success Factors of Decision Support Systems : An Experimental Study," *Data Base*(17:2), Winter 1986, pp. 3-15.

Lloyd,J.W. *Foundations of Logic Programming*, Berlin, Springer-Verlag, New York, 1984.

Missouri Division of Employment Security, Research and Analysis Section. *1989 Wage Rates in Selected Occupations for Andrew, Buchanan, Clinton, De Kalb and Holt Counties*, 1989.

Missouri Division of Employment Security, Research and Analysis Section. *1990 Wage Rates in Selected Occupations for Franklin and Jefferson Counties*, 1990.

Naumann,J.D. and Jenkins,A.M. "Prototyping : The New Paradigm for Systems Development." *MIS Quarterly*, September 1982, 6: 3, pp. 29-43.

Office of Management and the Budget (OMB). *Standard Industrial Classification Manual*, U.S. Government Printing Office, 1987.

Smith, P. *Expert System Development in Prolog and Turbo-Prolog*, Halsted Press, New York, 1988.

Sneath, P.H.A., and R.R. Sokal. *Numerical Taxonomy*, Freeman, San Francisco, 1973.

Sterling, L., and E. Shapiro. *The Art of Prolog*, The MIT Press, Cambridge, MA, 1986.

〈Appendix A〉 List of Ten Attributes and Its Values of the 114 Missouri Counties

| NAME | COD | WKWG | M | U | T | F | S | O | POP | WF | PCR | YR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Adair | 001 | 75.53 | 12 | 4 | 30 | 3 | 28 | 23 | 15772 | 67 | 3.7 | 90 |
| Andrew | 003 | 59.93 | 1 | 6 | 25 | 3 | 19 | 46 | 9614 | 65 | 5.5 | 89 |
| Atchison | 005 | 32.39 | 0 | 6 | 29 | 4 | 25 | 36 | 4666 | 60 | 3.6 | 00 |
| Audrain | 007 | 28.74 | 29 | 7 | 23 | 4 | 13 | 24 | 15588 | 61 | 5.2 | 90 |
| Barry | 009 | 80.52 | 51 | 3 | 18 | 3 | 11 | 16 | 16735 | 61 | 4.7 | 90 |
| Barton | 011 | 83.03 | 41 | 1 | 21 | 4 | 10 | 23 | 6838 | 59 | 3.7 | 00 |
| Bates | 013 | 47.52 | 10 | 7 | 26 | 5 | 18 | 34 | 9484 | 60 | 9.6 | 00 |
| Benton | 015 | 43.50 | 13 | 4 | 35 | 5 | 9 | 34 | 7852 | 60 | 8.3 | 89 |
| Bollinger | 017 | 19.33 | 29 | 3 | 14 | 2 | 13 | 39 | 6953 | 64 | 9.2 | 89 |
| Boone | 019 | 45.46 | 9 | 3 | 22 | 7 | 20 | 39 | 77702 | 72 | 2.7 | 89 |
| Buchanan | 021 | 56.21 | 24 | 8 | 25 | 5 | 19 | 19 | 53618 | 63 | 5.5 | 89 |
| Butler | 023 | 77.17 | 15 | 4 | 29 | 3 | 21 | 28 | 24418 | 63 | 8.6 | 89 |
| Caldwell | 025 | 28.74 | 0 | 3 | 19 | 4 | 8 | 66 | 4844 | 59 | 6.7 | 00 |
| Callaway | 027 | 49.35 | 20 | 14 | 16 | 2 | 13 | 35 | 21351 | 65 | 4.8 | 00 |
| Camden | 029 | 47.12 | 10 | 3 | 30 | 6 | 32 | 19 | 17708 | 67 | 6.1 | 99 |
| Cape Girardeau | 031 | 21.09 | 19 | 5 | 30 | 4 | 25 | 17 | 40834 | 67 | 4.2 | 00 |
| Carroll | 033 | 43.16 | 12 | 3 | 30 | 5 | 19 | 31 | 6446 | 58 | 8.6 | 90 |
| Carter | 035 | 45.32 | 29 | 4 | 16 | 2 | 15 | 34 | 3690 | 64 | 0.0 | 89 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| Wayne | 223 | 18.70 | 43 | 2 | 20 | 2 | 8 | 25 | 7539 | 62 | 9.9 | 89 |
| Webster | 225 | 69.95 | 35 | 5 | 20 | 4 | 6 | 30 | 14857 | 64 | 8.2 | 00 |
| Worth | 227 | 43.82 | 0 | 0 | 27 | 5 | 8 | 60 | 1527 | 57 | 4.3 | 00 |
| Wright | 229 | 48.39 | 38 | 3 | 23 | 3 | 12 | 21 | 10241 | 61 | 7.0 | 99 |

NAME : Name of County
COD : County Code
WKWG : Average Weekly Wage
M : % of Industry Mix (Manufacturing)
U : % of Industry Mix (Utilities)
T : % of Industry Mix (Wholesale/Retail)
F : % of Industry Mix (Finance/Insurance/Real Estate)
S : % of Industry Mix (Services)
O : % of Industry Mix (Others)
POP : Population
WF : % of Work Force
PCR : Population Change Rate
YR : Year of the Wage Rate Survey (89 : 1989 ; 90 : 1990 ; 00 : Unsurveyed County)

# 저 자 소 개

저자 조재희는 현재 광운대학교 경영정보학과 조교수로 재직 중이다. 그는 연세대학교 경영학과에서 학사학위, Miami University 대학원에서 계량경영 전공으로 경영학석사학 위, 그리고 University of Nebraska 대학원에서 경영정보학 박사학위를 받았다. 귀국 후 2년동안 펜타컴퓨터 EIS사업부와 에스.티.엠 기술연구소에 컨설턴트로 재직하며, 국내 여러 기업에 "Multidimensional Model을 기반으로 한 EIS"를 구축하는 현장경험을 하였다. 주요 관심분야는 정보시스템 디자인 및 개발 (EIS, DSS, ES), 다차원 모델/다차원 DB, 객체지향 프로그래밍 (Visual Basic, PowerBuilder, SQLWindows)이다.