Sri Lanka Association for Artificial Intelligence (SLAAI)
Proceedings of second Annual Session
19th February 2005 – Colombo

Using Fuzzy logic for automating knowledge acquisition

D.S. Kalana Mendis
Department of Information Technology,
Advanced Technical Institute,
Labuduwa.
kalanaatil@mail.com

Asoka S. Karunananda
Department of Mathematics & Computer Science,
Open University of Sri Lanka,
Nawala.
asoka@maths.ou.ac.lk

U. Samaratunga
Gampaha Wickramarachi Ayurveda Institute,
University of Kelaniya,
Yakkala.

Abstract

Knowledge acquisition is a key phase in construction of expert systems. This process is very much dependent on the nature of the domain knowledge. In particular, knowledge acquisition is a tedious task when modeling domains with tacit knowledge. Despite fuzzy logic has been used for knowledge acquisition in such domains, a large portion of the process is manually operated. This paper presents an approach to automated knowledge acquisition using fuzzy logic for the domains with tacit knowledge.

The novel approach allows the user or developer to directly interact with the system and enter tacit form of knowledge Acquisition of tacit knowledge is supported through questionnaire emulating the role of an interview of domain expert by the knowledge engineer. Tacit knowledge acquired through this session will be analyzed by statistical technique of principle component analysis to reveal the available dependencies in the knowledge acquired. The principle components generated will be transferred to fuzzy logic module for automatic construction of the fuzzy membership functions. Further, in a usual manner, system can acquired fuzzy rules relevant to manipulation of domain knowledge. Thus, collectively, the approach is consisted of principles component analyzer, fuzzy logic module and a fuzzy expert system.

The approach has been developed to be able to connect with a standard expert system shell. Currently, it has been integrated with FLEX expert system shell. The development has been done using Visual basic and the system runs on Windows platform. The approach has been applied to many domains include Ayurvedic domain of classification of individuals. It has shown 78%. accuracy in using the tacit knowledge for reasoning in the relevant domain. Performances were very close to handling tacit knowledge by the human expert in tacit domain.

1.Introduction

Construction of expert systems is a difficult task and it begins with knowledge acquisition. In particular acquiring of knowledge from tacit domains is still very add-hoc. Among other approaches use of fuzzy logic has shown some potential for knowledge acquisition in domains with tacit knowledge. This is because; fuzzy logic has inherent ability to handle vague situations. However, most approaches using fuzzy logic are largely operated manually, ranging from interviewing experts, analyzing

data, defining fuzzy logic membership function and writing fuzzy rules. In this paper we present our automated approach to use fuzzy logic for acquiring knowledge from domains with tacit knowledge, thereby supporting the development of expert systems for such domains. The rest of the paper is organized as following sections such as Fuzzy logic, Fuzzy logic for knowledge acquisition, Proposed approach, Iintegrating with Expert Systems, How approach works, Evaluation and References.

2.Fuzzy logic

Fuzzy logic deals with finding a truth ness of a concept in a range of values. It is not always considered the measurement for a degree of truth ness as an extreme value. Fuzzy logic underlying approximates, rather than exact, modules of reasoning – is finding applications that range from process control to medical diagnosis [29]. In more specifications, what is central about fuzzy logic is that, unlike classical logic systems, it aims at modeling the imprecise models of reasoning that play an essential role in the remarkable human ability to make rational decisions in an environment of uncertainty and imprecision. This ability depends, in turn, on our ability to infer an approximate answer to a question based on a store of knowledge that is inexact, incomplete, or not totally reliable.

2.1. What is a fuzzy set?

A fuzzy set can be simply defined as a set with fuzzy boundaries. Let X be the universe of discourse and its elements be denoted as *x*. Fuzzy set *A* of universe *X* is defined by function $\mu_A(x)$ called membership function of set *A*.

$$\mu_A(x):\ X \longrightarrow [0.1]$$

This degree, a value between 0 and 1, represents the degree of membership, also called membership value, of element *x* in set *A*.

2.2. Fuzzy rules

A fuzzy rule can be defined as a conditional statement in the form:

IF *X*= A
THEN *Y*= B

Where *X* and *Y* are linguistic variables; and *A* and *B* are linguistic values determined by fuzzy sets on the universe of discourses *X* and *Y*, respectively.

3.Fuzzy logic for Knowledge Acquisition

Tacit knowledge is embedded in skills and cannot be demonstrated and so is very difficult to transfer [30]. Buchanan *et al*. (1983) defines knowledge acquisition as ′the transfer and transformations of potential problem-solving expertise from some knowledge source to a program′. There are a number of reasons why productivity is typically; here are some of them.

- Specialist fields have their own jargon, and it is often difficult for experts to communicate their knowledge in every day language. Analyzing the concepts behind the jargon is rarely straightforward, since these concepts need not admit of precise mathematical or logical definition.
- The facts and principals underling many domains of interest cannot be categorized precisely in term of a mathematical theory of a deterministic model whose properties are well understood.
- Experts need to know more than the mere facts or principals of a domain in order to solve problems.
- Human expertise, even in a relatively narrow domain, is often set in a broader context that involves a good deal of commonsense knowledge about the every day world.

All knowledge can be considered as tacit or rooted in tacit knowledge. It is important to investigate the methods available for tacit knowledge acquisition. Since tacit knowledge is embedded in implicit nature, fuzzy logic gives great interest of handling such kind of nature. Further more, this can be considered as a method for tacit knowledge acquisition using fuzzy logic.

Despite fuzzy logic has been used for knowledge acquisition in such domains, a large portion of the process is manually operated. XpertRule Knowledge Builder extends the graphical knowledge representation paradigm, established since 1988, by its predecessor XpertRule KBS [28], to new levels of scalability and flexibility. Although the knowledge acquisition accompanied with the methods based on fuzzy logic, but it is exploited

the level of transparency and accuracy due to handling manually constructed membership functions. There is a great issue of dealing with constructing membership functions, especially on determinations about intervals of membership functions. Most of the time the knowledge engineer is expected to do this task which leads to arise questions about system validation.

Another knowledge acquisition tool for computer assisted diagnosis of postmenopausal osteoporosis using a fuzzy expert system shell [28] is also seen in a position of a manually operated tool for knowledge acquisition. Although WinProlog LPA [16] gives a toolkit (FUZZYEG) based on fuzzy logic for constructing membership functions effectively, but it appears a manual method for determining the intervals of membership functions.

It is intended that finding a way to a process of automated knowledge acquisition will assure more validations about the tacit knowledge acquisition to a system.

4. Proposed approach

We postulate a new approach enhancing the ability of automated knowledge acquisition using fuzzy logic. It has been exploited the process of the new approach in following steps.

4.1. Acquiring knowledge

The approach begins with by acquiring tacit knowledge. This can be done as an interview between domain experts and the knowledge engineer. Using the interviewing process between expert and knowledge engineer, tacit knowledge has been acquired and mapped in to a questionnaire based on Likert scale technology [14]. WE have chosen to acquire tacit knowledge into a questionnaire since it is more convenient for further analysis. On the other hand, the questionnaire can be automated to interact directly with the domain expert without involving a knowledge engineer. Once tacit knowledge has been qcuired then we should analyses the knowledge for finding dependencies. The questionnaire has been analysed using principal component analysis (PC) [7] to find dependencies.

4.1.1. What is Principle components analysis?

The concept of PCA is based on the derivation of linear combinations of the *p* measured variables X_1, $X_2..X_p$ to produce 'derived variables', that are uncorrelated and are such that explains a different 'dimension' within the data [7]. Such derived variables are referred to as principal components (PCs). As there are *p* response variables within the data set, *p* principal components can be derived. The first PC, denoted PC1, is expressed in the form

$$PC_1 = \alpha_{11} X_1 + \alpha_{12} X_2 + ... + \alpha_{1p} X_p \quad (1)$$

Where the α terms refer to the weights of each variable within this principal component PC_1. The weights of each PC_i represent the eigenvector solution, which maximise the variance of each PC_i, where i is the number of components.

4.1.2. Extracting principal components

The importance of each PC, in terms of level of data variation explained, is specified by its eigenvalue, the λ term, with $\Sigma\lambda$ representing the total of the *p* eigen values. A measure of the proportion of data variation accounted for by each PC, based on the equivalence of eigenvalue and PC variance, is provided by the expression $\lambda/(\Sigma\lambda)$.

Generally, it is required to select those PCs, which account cumulatively for at least 80% to 90% of the data variation. In addition that each PC must exceed eigenvalue more than 1.However, if nearly all the correlations are less than 0.25, then there is probably not much point in carrying out a PCA. But to reduce even that much of interdependency PCs can be computed.

4.1.3. PC for tacit knowledge

Let *S* be the set of all questions in the questionnaire and *P* be the set of all extracted principle components.

Further, $P = \{PC_1, PC_2, ..PC_{n-1}, PC_n\}$

$$S = \{S_1, S_2, ..S_{m-1}, S_m\}$$

$\Rightarrow$

$$PC_i = a_{1i} S_1 + a_{2i} S_2 + .. + a_{m-1,i} S_{m-1} + a_{m,i} S_m$$

Let *M* be the principle components Matrix for filtered tacit knowledge.

$$M = \begin{bmatrix} a_{11}a_{12}..a_{1n} \\ a_{21}a_{22}..a_{2n} \\ . \\ . \\ a_{m1}a_{m2}..a_{mn} \end{bmatrix}$$
(2)

$$\therefore PC_1 = a_{11}S_1 + a_{21}S_2 + .. + a_{m1}S_m$$
(3)

$$PC_2 = a_{12}S_1 + a_{22}S_2 + .. + a_{m2}S_m$$
(4)

$$PC_{n-1} = a_{1n-1}S_1 + a_{2n-1}S_2 + .. + a_{mn-1}S_m$$
(5)

$$PC_n = a_{1n}S_1 + a_{2n}S_2 + .. + a_{mn}S_m$$
(6)

For n number of extracted principal components, following computation is concluded

$$X = \sum_{j=1}^{n} PC_j$$
(7)

$$\therefore X = \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}S_i$$
(8)

4.2. Generating membership function

Let LS be the Likert scale, then

$$LS = [L,..,U]$$
(9)

X_L and X_U values are derived from results of the filtered tacit knowledge. It is computed as given below.

$$\therefore X_L = L\sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}$$
(10)

$$\therefore X_U = \mathrm{U}\sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}$$
(11)

Let A be fuzzy set defined on a fuzzy concept using the interval of $[X_L,..,X_U]$. Then membership function is as follows.

$$A(X) = \begin{cases} 0 & X=<X_L \\ (X-X_L)/(X_U-X_L) & X_L<X<X_U \\ 1 & X>=X_U \end{cases} \quad (12)$$

4.3. Adding fuzzy rules

Fuzzy rules can be constructed as follows,

Rule 1: If $X <= X_L$ then A (X)=0 %

Rule 2: If $X_L<X<X_U$ then $(X-X_L)/(X_U-X_L)$ %

Rule 3: If $X>=X_U$ then A (X)=100 %

Further, fuzzy rule base can be extended by adding dynamically, in order to function the reasoning process for answers given by the fuzzy rules.

5. Integrating with Expert Systems

The approach has been converted for an implementation using the architecture given below (Figure 1). It is consisted of with modules such as principle component analyser, database, knowledge base, and fuzzy logic module and inference engine.

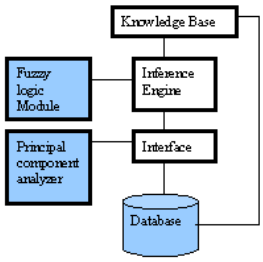

Figure 1: Top level architecture

5.1. Principle components analyser

Tacit knowledge has been extracted from the expert and formulated in a questionnaire. It is evaluated using Likert scale technology. In the first instance of knowledge acquisition, a pilot survey has be done for the purpose of extracting principle components. The SPSS [18] is used for conducting the functions of principle components extracting.

5.2. Fuzzy logic module

The output results of the principle component analyser would be the input for the fuzzy logic module. In the case of generating membership function, finding the interval is considered as an automated process in this module due to instead of using runtime inputs. This module has been implemented using Visual Basic [19] for widening scope of generating membership function. Further, fuzzy rules have been constructed in the fuzzy logic module.

5.3. Database

Extracted principle components have been stored in Ms Access [19] database, which integrated with the principle component analyser through the developer interface that is considered as a sub interface of the user interface. The questionnaire consisted of tacit knowledge also been stored in the database that integrated with the user interface.

5.4. Knowledgebase

Explanations for output generated by the fuzzy logic module has been processed using fuzzy rules in the knowledge base. Further, knowledge engineer is given a facility to add new rules in the runtime. The knowledge base has been implemented using FLEX expert system shell, which embedded in WinProlog [16].

5.5. User interface

The user interface facilitates for both developer and general user. Once knowledge engineer develops a particular framework for required tacit domain with interaction of the expert, and then general user will be given a facility of using the framework for decision-making purposes. So, it has been divided the user interface in developer interface and general user interface. General user will be able to use a developed framework using a questionnaire, which has been implemented as a web page linked to the database.

5.6. Inference engine

The inference engine carries out the reasoning whereby the expert system reaches a solution. This is the inference engine of the FLEX expert system shell. Since this is built in to the system there is no development activities with regard to this component in the system. Note that inference engine has nothing to do with the modelling of tacit knowledge but it runs the expert system.

6. How approach works

We have illustrated our approach using Ayurvedic medicine as a domain with tacit knowledge. In doing so, classification of individuals through clinical examination in Ayurveda has been considered [5]. The clinical examination of Ayurveda is divided into 2 paths, namely: examination through patient and examination through disease.

Prescribing drugs for a disease is depended on both 2 examinations. Classification of individual (human constituents) is included in examination through patient, which defined as a concept called *'prakurti pariksha.'*. Individual can be categorized into *vata* or *pita* or kapha based on the *'prakurti pariksha.'*. It was defined that one type can be dominated but in combination of all 3 types. In the exciting system, the method of analysing constituents is not consistent. Although Ayurvedic practitioners use a questionnaire but leads several problems like dependencies among the questions in the questionnaire and analysis of the constituent type. We addressed these problems to solve using following stages.

6.1. Extracting tacit knowledge in Ayurveda

In the first instance we mapped tacit knowledge regarding to analysis of constituents to a questionnaire with interaction of an Ayurvedic expert. It is consisted of 72 questions to analyse *vata*, *pita* and *kapha*. It is shown as Figure 2.

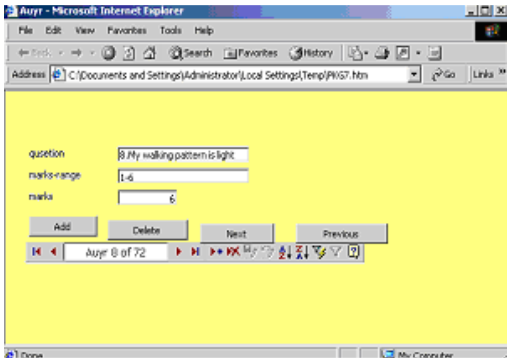

Figure 2: Questionnaire window

6.2. Removing dependencies

We have done a pilot survey using 100 no. of students for statistical modelling. Principal component analyser has been used to remove dependencies. It has been identified 25 principal components as shown in Figure 3.

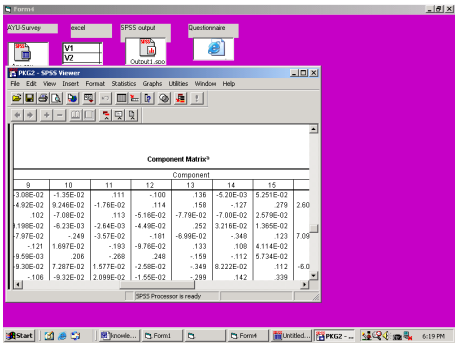

Figure 3: principal component Matrix

6.3. Analysis of human constituents

Human constituents can be computed in to *vata, pita* and *kapha* in percentages as shown in Figure 4. Membership functions for *vata, pita* and *kapha* have been constructed using the out puts of principle component analyser.

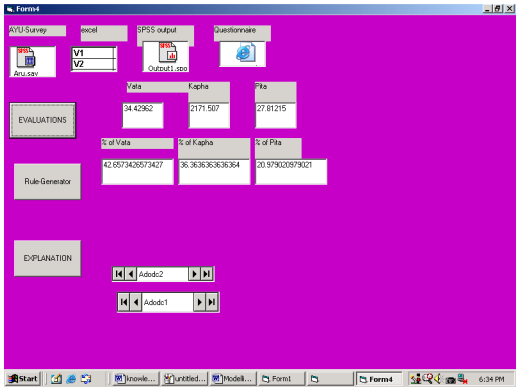

Figure 4: Analysis of human constituents

6.4. Explanations for derived human constituents

Possible diseases can be occurred due to dominated constituent type. It is illustrated as shown in Figure 5.

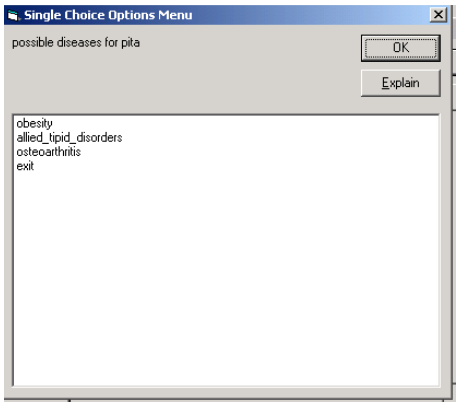

Figure 5: Explanation window

6.5. Adding new rules

The developer is given a privilege to add new rules at the run time. It is shown as Figure 6. given below.

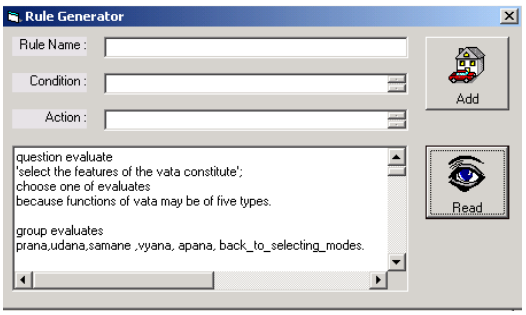

Figure 6: Run time rule addition window

7. Evaluation

The expert system developed using this approach was tested with a group of 35 persons of Ayurvedic experts and Ayurvedic medical students. The evaluation was conducted to see far the answers gerated by the system matches with the identification by Ayurvedic experts and the students. Further, the system's ability to fine-tune the answers were also tested.

It is investigated that 78% of conclusions matches with the system and expert using descriptive statistics.

The system facilitated to derive constituents types in percentages while Ayurvedic experts obtain only the constituent type. As recommendation given by the Ayurvedic experts, determining constituent's types in percentages is an important criterion for prescribing drugs for a disease. Further, our system provide as an option to find out possible diseases.

In generally, the system can be used as a self-assessment for finding constituents. According to Ayurvedic medicine, regiments can be done easily by knowing the constituent type. The human constituents can be computed as a combination. So it would help to find the effectiveness of minimum type in a diagnosis.

8. Conclusion & Further work

In this approach we encourage the domain experts to present their knowledge to construct more successful questionnaire. However, different experts may propose different questionnaire since their emphasis of domain knowledge is different. At present our system is based on one expert view of the domain knowledge. As the further work, we propose to improve the system to capture different experts' tacit knowledge of the given domain and synchronize those for generating most comprehensive set of tacit knowledge. Eventually PC analysis will be done on that knowledge and generate the appropriate fuzzy membership functions. In other words, we emphasise that the system do not operate on a pre-defined set of tacit knowledge, rather the collection of tacit knowledge is evolving in presence of different experts. In the long run the tacit knowledge may lead to become more formalized.

9. REFERENCES

[1] Karunananda A.S (2000), How To Do Research Author Publication.

[2] Jonson L (1988), Expert system Architectures, Kopan Page Limited.

[3] George J (1997), "Fuzzy sets and Fuzzy Logic", Prentice Hall of India.

[4] Noak V(2000), "Discovering the world with Fuzzy logic", A Springer – Verlag Company, PP. 3 – 50.

[5] Dubey G.P , (1978)The Physiological concepts in Indian medicine, Science and Philosophy of Indian medicine.

[6] Tripathi S.N(1978), "Clinical Diagnosis", Science and Philosophy of Indian medicine.

[7] Chatfied C(1996)," Introduction to Multivariate Analysis", Chapman and Hall.

[8] Morrison D.F(1990), "Multivariate Statistical Methods", Mc Graw-Hill, Inc. PP. 312 - 357.

[9] Ross Dawson, (2001), "Developing Knowledge-Based Client Relationship ", Butterworth Heinemann.

[10] Karunananda A.S (2001),Pinki – A motivation for Artificial Intelligence projects.

[11] Mendis D.S.K, Karunananda A.S, Samaratunga U. (2003), "Reasoning with Uncertainty", Health Informatics Conference, Colombo.

[12] Mendis D.S.K, Karunananda A.S, Samaratunga U. (2003), "Tacit knowledge modeling", (2003), Sri Lanka Association for Artificial Intelligence Annual Session and Annual General Meeting, Colombo.

[13] Bellman and Zadeh, (1970). R.E. Bellman and L.A. Zadeh , Decision making in a fuzzy environment. Management Science 17, pp. 141–164.

[14] Richards D. and Bush P., "Measuring, Formalizing and Modeling Tacit Knowledge" IEEE/Web Intelligence Conference (WI-2003) Bejing.

[15] M. Dzbor, 2000. Explication of Design Requirements Through Reflection on Solutions. In Proc. of 4th IEEE Conference on Knowledge-based Intelligent Engineering Systems. Brighton, UK, pages 141-144.

[16] Dave Westwood, Flex reference guide, LPA, U.K

[17] M.Sc.in Architecture, (2003),Field projects,

University of Moratuwa.

[18] Matei Ciobanu Morogan, (1997), SPSS for windows, release 8.0.0, DSV KTH/ SU, Sweden.

[19] Evangelos Petroutsos,(1998), Mastering Visual Basic 6, BPP Publications, New Delhi

[20] George J.K, Yuan B.(1995), Fuzzy sets and Fuzzy logic, prentice hall of India, pp. 280–300.

[21] XpertRule Knowledge Builder, www.attar.com.

[22] Peter J. (1995), Introduction to expert systems, Pearson education Asia, pp. 182–199.

[23] Karl B, Klaus P, A,Yoichi H, Thomas E. Rothenfluh and Harald L (2004), "Knowledge acquisition in the fuzzy knowledge representation framework of a medical consultation system" International Journal of Artificial Intelligence in medicine, pages 1–26.

[24] Rajeev K (1995), "Integrating knowledge bases in expert system shells: an open system approach", International Journal of computer application in technology, pages 78–89.

[25] Coppin G, Skrzyniarz A (2003), "Human–centered processes : Individual and distributed decision support", IEEE Intelligent systems, pages 27–33.

[26] Hayashi Y, Tazaki E, Yoshida K, Dey P (1988), "Medical diagnosis using simplified multi-dimensional fuzzy reasoning", In: Proceedings of IEEE 1988 International Conference on Systems, Man and Cybernetics, Beijing, China, 1988. p. 58–62.

[27] Kolousek G (1997), ″The system architecture of an integrated medical consultation system and its implementation based on fuzzy technology", Doctoral thesis, Technical University of Vienna, Austria.

[28] Binaghi E., De Giorgi O., Maggi G., Motta T., Rampini A (1993), A knowledge acquisition tool for computer assisted diagnosis of postmenopausal osteoporosis using a fuzzy expert system shell, Computers and Biomedical Research, 26,pp.498-516.

[29] Zadeh L. A (1988),Fuzzy Logic, Computer IEEE, pp. 83-93.

[30] Senkar J (1995), ″Network and tacit knowledge in innovation″, Economics et societies, pp. 99-118.

[31] Zimmerman H.J (1985), Zimmerman Fuzzy set theory and its applications, Kluwer–Nijhoff, Boston