Fuzzy Associative Memory-Driven Approach to Knowledge Integration

Myoung-Jong Kim\textsuperscript{1}, Ingoo Han\textsuperscript{1}, Kun Chang Lee\textsuperscript{2}

\textsuperscript{1}Graduate School of Management, Korea Advanced Institute of Science & Technology, mjkim@kgsam.kaist.ac.kr
\textsuperscript{2}School of Business Administration, Sungkyunkwan University, Seoul 110-745, Korea

Abstract

We propose a knowledge integration mechanism that yields a cooperated knowledge by integrating user knowledge, expert knowledge and machine knowledge within the fuzzy logic-driven framework, and then refines it with a fuzzy associative memory (FAM) to enhance the reasoning performance. The proposed knowledge integration mechanism is applied for the prediction of Korea stock price index (KOSPI). Experimental results show that the FAM-driven approach can enhance the reasoning performance by refining the cooperated knowledge of fuzzy logic-driven framework. This result means that the FAM-driven approach can be a robust guidance for knowledge integration.

1. Introduction

The typical types of knowledge which can be used for an intelligent decision support are machine knowledge, expert knowledge, and user knowledge [21]. Both expert and user knowledge means human knowledge. Several researchers in behavioral science compared and investigated the characteristics of expert knowledge and user knowledge [3, 6, 8, 11, 19, 24]. Experts have expertise or expert knowledge which is highly organized and domain-specific enough to encode complex information [19, 24] and result in faster and more accurate performance [3] than user knowledge. Meanwhile, user knowledge is concerned with such knowledge derived from random users who may be novice or experienced. User knowledge is not well organized and especially varies with the user's experienced level. Therefore, it is natural to assume that the performance with expert knowledge outperforms one with user knowledge.

These two types of human knowledge have the advantage in the adaptability to the changing environment. However, it is a quite difficult to extract the consistent knowledge from them because they differ from different experts or users. Meanwhile machine knowledge is consistent because it is derived automatically by applying machine learning techniques such as Inductive learning [27], neural network [28], case-based reasoning [17], genetic algorithms [7, 13], etc. to historical instances that possess regularities useful for interpreting some parts of phenomena. The most important difference between machine knowledge and human knowledge is that the former relies on the objective method and the latter is compiled from hard-to-explain psychological processing of information in human brain.

There have been a variety of efforts to combine multiple sources of knowledge in the fields of forecasting and classification [1, 2, 4, 5, 9, 10, 12, 14, 20, 25]. More recently, [21] propose the fuzzy logic-driven framework that yields the cooperated knowledge (CK) by integrating multiple sources of knowledge. The findings of their mechanism show that it is promising in integrating multiple knowledge. However it suffers from the conflicts between multiple sources of knowledge which are associated with the knowledge pertaining to a case or an object. This may degrade the performance of an intelligent system with multiple sources of knowledge when applied to solving an ill-structured problem [16].

In this sense, this paper proposes the FAM-driven approach to knowledge integration to generate the domain knowledge which can resolve the knowledge conflicts and enhance the reasoning performance. The proposed mechanism will be applied for the prediction of Korea stock price index (KOSPI). Experimental results show that the proposed mechanism can provide an intelligent system with the more adaptable and robust intelligence.

The paper is organized as follows. Section 2 reviews the background of knowledge integration. Fuzzy logic-driven framework for generating the CK is presented in section 3. The knowledge refinement by the FAM-driven approach is introduced in section 4. The empirical test with the prediction of KOSPI is shown in section 5. In section 6, this paper is ended with some concluding remarks.

2. The Background for Multiple Knowledge Integration

The studies about knowledge integration in
the fields of forecasting have shown the better reasoning performance by combining multiple results obtained from different models [4]. The argument underlying the combination of multiple results is that a proper combination provides more accurate results than the single result because it can reduce the magnitude of variance [10]. Most research for combination of multiple knowledge focused on combinations of multiple models or multiple experts where model represents machine knowledge and expert stands for human knowledge [2, 5, 9, 10, 12, 20, 25]. This type of combination means that integration of machine knowledge and human knowledge can reduce uncertainty involved with data and produce more elaborate results [1].

The traditional approaches to knowledge integration in the field of classification are based on linear combination. They are determined and classed by two main elements of the prediction performance and the predictive certainty. The performance of the knowledge is measured by the historical results and the certainty of the knowledge is by the degree of confidence. However, they can not deal with the uncertainty involved in involved in the highly ill-structured problems because they depend on the deterministic linear combination. Another limitation is that they can not applied to n group classification problems where n >2 [14].

More recently, fuzzy logic-driven framework is proposed to generate the CK by integrating multiple sources of knowledge [21]. The findings show that the CK outperforms single sources of knowledge in terms of the reasoning performance. However it suffers from the conflicts between multiple sources of knowledge. The knowledge conflicts may degrade the robustness of an intelligent system with multiple sources of knowledge, thereby the system may fall into the trap of local optima [16]. In this sense, this paper presents the FAM-driven approach to resolve the conflicts between multiple sources of knowledge.

3. Fuzzy Logic-Driven Framework for Generating the Cooperated Knowledge

We use fuzzy logic-driven approach [21] to generate the CK from multiple sources of knowledge. It consists of three phases: (1) machine knowledge-based inference (MKBI) phase (2) expert knowledge-based inference (EKBI) phase, and (3) combining phase.

3.1 MKBI Phase

Variables and Data selection

Machine knowledge (MK) is obtained by applying backpropagation neural network model [28] to technical indicators obtained from KOSPI. We collect 5 technical indicators indicating dynamic trends of stock price index including Moving Average, Relative Strength Index, Stochastic %D, Disparity, and Rate Of Change. Specific values for each technical indicator are 'good', 'not good or not bad', and 'bad', which are determined by users' knowledge.

The output values are the four levels of stock market of next month which are classified into the following categories: Bear, Edged-Down, Edged-Up, and Bull. A criterion that has been used by stock market analysts, (-3%, 0%, +3%), is used for determining category.

The total number of samples available is 649 weekly data from July 1982 to December 1995. The data set is split into three subset according to the time period used for neural net training (493 weeks from July 1985 to December 1992), FAM learning (52 weeks from January 1993 to December 1993), and validation (104 weeks from January 1994 to December 1995).

Generation of Machine Knowledge

We can define a fuzzy prediction vector of MK (FPVMK) as fuzzy membership function derived from a neural network:

\[ FPV^{MK} = \mu_{MK}(\text{Bear}), \mu_{MK}(\text{Edged-Down}), \mu_{MK}(\text{Edged-Up}), \mu_{MK}(\text{Bull}) \]

Suppose that we obtain a particular result of a neural network model as follows:

\[ FPV^{MK} = (0.5390, 0.8520, 0.1222, 0.1012) \]

This result indicates the MK predicts stock market level of next week as Edged-Down with a fuzzy value .8520.

3.2 EKBI Phase

External Factors and Data

This phase is to combine user knowledge (UK) and expert knowledge (EK) expressed as fuzzy membership functions for external factors. We consider only four types of external factors including economy prospects (EP), the amount of stock supply and demand (SSD), the amount of currency ready for buying stocks (AOC), and conditions favorable or unfavorable to stock market trend (CPU).

The data is classified into two data sets used for genetic learning (52 weeks from January 1993 to December 1993) and validation (104 weeks from January 1994 to December 1995).
**EKBI Phase**

We introduce fuzzy membership functions for UK and EK about each external factor to combine UK and EK. We assume a triangular-shaped fuzzy membership function that has a center value $c$ and a width $w$. The center value indicates the most probable value and width means a level of expertise. Fuzzy membership value for the center value is always identical to 1 in the case of triangular-shaped fuzzy membership function. If the width value is large, the expertise level is regarded as low, otherwise high. If he has the width value 0 for a certain external factor, his judgment about the external factor is assumed as extremely reliable.

Let $\mu_{ue}(x)$ and $\mu_{ek}(y)$ denote respectively UK membership function and EK membership function about $ith$ external factor, i=1, 2, 3, 4. Both $x$ and $y$ represent one of five evaluation categories: very bad, bad, not good or not bad, good, and very good. Also let us define a Fuzzy Evaluation Vector for $K$ type knowledge-based evaluation of $ith$ external factor ($FEV^K_i$) as follows:

$$FEV^K_i = (\mu^K_{ub}(very\_bad), \mu^K_{ub}(bad), \mu^K_{ub}(not\_bad),$$

$$\mu^K_{ub}(good), \mu^K_{ub}(very\_good))$$

where $K$ means either UK or EK. Then we can define a $K$ type knowledge-based Fuzzy Evaluation Matrix ($FEM^K_i$) evaluating all the external factors, consisting of row vectors $FEV_i$.

$$FEM^K = [FEV^K_i], i = 1, 2, 3, 4$$

Therefore, the dimension of $FEM^K$ in our case is 4 by 5.

Suppose that UK provides bad for EP, bad for SSD, very bad for AOC, and good for CFU. EK provides very bad for EP, bad for SSD, very bad for AOC, and bad for CFU. In addition to this sort of fuzzy evaluation about external factors, we assume that the width of fuzzy membership function for EK is 2 and the width of fuzzy membership function for UK is 3. Then we can obtain $FEM$s for UK and EK respectively as follows:

$$FEM^UK = \begin{bmatrix}
0.67 & 1.00 & 0.67 & 0.33 & 0.00 \\
0.67 & 1.00 & 0.67 & 0.33 & 0.00 \\
1.00 & 0.67 & 0.33 & 0.00 & 0.00 \\
0.00 & 0.33 & 0.67 & 1.00 & 0.67 \\
1.00 & 0.50 & 0.00 & 0.00 & 0.00 \\
0.50 & 1.00 & 0.50 & 0.00 & 0.00 \\
1.00 & 0.50 & 0.00 & 0.00 & 0.00 \\
0.50 & 1.00 & 0.50 & 0.00 & 0.00 
\end{bmatrix}$$

$$FEM^EK = \begin{bmatrix}
0.67 & 1.00 & 0.67 & 0.33 & 0.00 \\
0.67 & 1.00 & 0.67 & 0.33 & 0.00 \\
1.00 & 0.67 & 0.33 & 0.00 & 0.00 \\
0.00 & 0.33 & 0.67 & 1.00 & 0.67 \\
1.00 & 0.50 & 0.00 & 0.00 & 0.00 \\
0.50 & 1.00 & 0.50 & 0.00 & 0.00 \\
1.00 & 0.50 & 0.00 & 0.00 & 0.00 \\
0.50 & 1.00 & 0.50 & 0.00 & 0.00 
\end{bmatrix}$$

Using $FEM$s for UK and EK, combined fuzzy evaluation matrix of human knowledge ($CFEM^{UK}$), is calculated as follows:

$$CFEM^{UK} = FEM^UK \land FEM^EK = \begin{bmatrix}
0.67 & 0.50 & 0.00 & 0.00 & 0.00 \\
0.50 & 1.00 & 0.50 & 0.00 & 0.00 \\
1.00 & 0.50 & 0.00 & 0.00 & 0.00 \\
0.00 & 0.33 & 0.50 & 0.00 & 0.00 
\end{bmatrix}$$

Let us define the weight vector to consider the effect of each factor on five evaluation categories as $W=(W_1, ..., W_m), 0<\nu<1$ for $m$ factors where the sum of weights should be equal to 1. In our case, we assumed that $W$ is (.25 .25 .25 .25), each element of which means the amount of influences that each factor has on five evaluation categories. By multiplying $W$ with the $CFEM^{UK}$, HK-based combined fuzzy evaluation vector ($CFEV^{HK}$) can be calculated as follows:

$$CFEV^{HK} = W \times CFEM^{UK} = (.5425 .5825 .25 .0 .0) .$$

Finally, consider the following 5 by 4 conversion matrix ($CM$) to transform five evaluation categories of combined fuzzy evaluation vector ($CFEV^{HK}$) into four levels of stock market represented in fuzzy predict vector

$$\begin{bmatrix}
0.5 & 0.0 & 0.0 & 0.0 \\
0.5 & 0.5 & 0.0 & 0.0 \\
0.0 & 0.5 & 0.5 & 0.0 \\
0.0 & 0.0 & 0.5 & 0.5 \\
\end{bmatrix}$$

$$(FPV^{HK}) \ CM = \begin{bmatrix}
0.5625 & 0.4163 & 0.1250 & 0.00 
\end{bmatrix}$$

Therefore $FPV^{HK} = CFEV^{HK} \times CM = (.5625 .4163 .1250 .0)$.

#### 3.3 Combining Phase

We obtained $FPV^{HK}$ and $FPV^{MK}$ from MKBI phase and EKBI phase, respectively. Then to create a cooperated knowledge ($CK$), we use min operator for combining $FPV^{HK}$ with $FPV^{MK}$ and generating a $FPV$ of the cooperated knowledge ($FPV^{CK}$) as follows:

$$FPV^{CK} = (FPV^{HK} \land FPV^{MK})$$

$$= (.5390 .8520 .1222 .1012) \land (.5625 .4163 .1250 .0)$$

$$= (.5390 .4163 .1222 .0).$$

From this FPV of the cooperated knowledge, the stock market level of next month is predicted as bear with a fuzzy value 0.5390.

#### 4. The FAM-Driven Approach
Fuzzy logic-driven framework seems to be more better at generating case-specific knowledge integration than domain knowledge, as it generates a CK from MK and HK pertaining to a case or an object as described in previous section. This may degrade the robustness of an intelligent system with multiple sources of knowledge so that the system may not provide the high quality intelligence [16]. We introduce a notion of FAM to extract the domain knowledge, which is capable of explaining the regularities generally observed in numerous instances, from the results of fuzzy logic-driven framework. The domain knowledge can provide the generality and robustness for an intelligent system and also resolve the conflicts between case-specific CK.

The FAM is to map and associate a $p$-dimensional fuzzy set and another $n$-dimensional where $p$ and $n$ are linguistic values of each fuzzy set, and represents the FAM rules as $p$ by $n$ matrix. Since we use four linguistic values such as Bear, Edgedown, Edged-up and Bull for all types of knowledge, the FAM consists of 4 by 4-dimensional matrix [18].

To build the FAM, we can transform the above results from MKBI, EKBI, and combining phase into the form of a fuzzy rule. In a fuzzy rule, MK and HK are corresponding to the attributes in IF part of the rule, and CK is a set of class in THEN part of the rule. One way to transform a particular result from fuzzy logic-driven framework into a fuzzy rule is to make a crisp-cut for membership function (herein, 0.5). For instance, if membership function of a linguistic term is greater than 0.5 then the linguistic value is converted to 1, otherwise to 0. The fuzzy rule consists of the linguistic term and the corresponding attributes and classes converted to 1. For instance, the above case, $u_1$, can be transformed into a fuzzy rule $r_1$ as following:

$$\begin{align*}
u_1 &= (\text{MK}; \text{HK}; \text{CK}) \\
&= (0.5390, 0.8520, 0.1222, 0.1012, 0.5625, 0.4163, 0.1250, 0.0) \\
r_1 &= (1, 0, 0, 0, 0, 1, 0, 0, 0)
\end{align*}$$

The rule $r_1$ can be interpreted as following: 
"IF MK is Bear or Edged-Down and HK is Bear THEN CK is Bear". The rule $r_1$ could be classified into two rules for mapping the fuzzy rule to the FAM as following: $r_{11} = (1, 0, 0, 0; 1, 0, 0, 0, 1, 0, 0, 0)$ and $r_{12} = (0, 1, 0, 0; 1, 0, 0, 0, 1, 0, 0, 0)$. Suppose that we obtain another different fuzzy rule from the different case as following: $r_2$ ("IF MK is Edged-Down and HK is Bear THEN CK is Edged-Down"). Two rules, $r_{12}$ and $r_2$, have the same IF parts, but different THEN parts. Two fuzzy rules can be applied to an object or a case at the same time, therefore the object may be predicted as two classes. One way to resolve such a conflict between CkS is using the domain knowledge proven to be robust to select only one class. We consider the domain knowledge as a set of fuzzy rules with the highest accuracy for training data between the competitive rules and then represent it in the FAM. Therefore, the FAM consists of a set of fuzzy rules used to resolve the conflicts between knowledge. We can compute the accuracy of the fuzzy rule by means of the degree of matching between an object $u$ and a rule $r$ in the respects of attribute, condition, and conclusion to calculate the accuracy of the fuzzy rule [30].

1) The degree of attribute match between a rule $r$ and an object $u$ is measured by $MA_A(r, u)$,

$$MA_A(r, u) = \text{Max} (\text{Min} (r^A, u^A))$$

where $K$ means attributes, MK or HK.

2) The degree of condition match between a rule $r$ and an object $u$ is measured by $MA_C(r, u)$.

$$MA_C(r, u) = \text{Min} (MA_A(r, u))$$

3) The degree of conclusion match between a rule $r$ and an object $u$ is measured by $MC(r, u)$.

$$MC(r, u) = \text{Min} (r^C, u^C)$$

where $C$ means a set of classes.

As an example of a rule $r_{12}$ and the above case $u_1$, we have attribute match: $MA_{attr}(r_{12}, u_1) = 0.8520$, $MA_{attr}(r_{12}, u_1) = 0.5625$; condition match: $MA(r_{12}, u_1) = 0.5625$; and conclusion match: $MC(r_{12}, u_1) = 0.5390$.

4) The accuracy of a rule $r$ at the predefined significance level $\alpha$ (herein 0.5) is defined by $Ac(r)$:

$$Ac(r) = \frac{\sum_{u_1} \text{Min} (f_A(MA_A(r, u)), f_C(MC(r, u)))}{\sum_{u_1} \text{Min} (f_A(MA_A(r, u)))}$$

It measures the degree to which the condition set is the subset of the conclusion set, or the truth level that the condition implies the conclusion [18].

5. Experiments and Discussion

The rule with the highest accuracy is select from the competitive rules as the domain knowledge and represented in the FAM to resolve the conflicts as shown in Table 1.
MK is derived from UK expressed as fuzzy membership function for three evaluation categories 'good', 'not good', and 'bad'. HK is the integration of UK and EK, which are described in one of five evaluation categories for external factors. Although they use UK as the primary knowledge source, MK (0.5962) and HK (0.625) show the better performance than UK (0.6923). This means that MK and EK can be used as intelligent guidance for supporting users' decision making. This result is supported by the findings that intelligent systems with expertise and model base can support users' decision making [22].

CK knowledge derived from fuzzy logic-driven framework (0.5577) has the higher level of predictive performance than MK and HK. This result shows that the knowledge integration can provide the robust knowledge with an intelligent system. When compared with the result of CK, the reasoning performance of the cooperated knowledge derived from the FAM-driven approach (0.4904) outperforms than CK.

We use Wilcoxon matched-pairs signed-ranks test to examine whether the predictive performance of the FAM is significantly higher than that of other techniques. Statistical results show that the FAM-driven approach performs significantly better than any other knowledge at 5% level.

6. Concluding Remarks

This paper presents the knowledge integration mechanism using the FAM-driven approach, which derive the domain knowledge from case-specific knowledge of fuzzy logic-driven framework. The domain knowledge consists of a set of fuzzy rules to generalize CKs by resolving the conflicts between them. We demonstrate that the FAM-driven approach can enhance the reasoning performance by generalizing case-specific CKs. In this sense, the FAM-driven approach can be a robust knowledge guidance for fuzzy logic-driven framework.

However, there remain further research issues as following: One issue for further research is related to automatic tuning of membership function. One solution for tuning fuzzy membership function is using some AI techniques such as genetic algorithms [15] and neural networks [23]. Another is that the FAM approach is promising in knowledge integration, however it may not an optimized integration for providing the better or the best performance. Recently, genetic algorithms have been applied to
optimizing the FAM-based control systems [26, 29].

References