The Hybrid Knowledge Integration Using the Fuzzy Genetic Algorithm

Myoung-Jong Kim[1], Ingoo Han[1], Kun Chang Lee[2],

[1]Graduate School of Management, Korea Advanced Institute of Science & Technology, mjkim@kgsm.kaist.ac.kr

[2]School of Business Administration, Sungkyunkwan University, Seoul 110-745, Korea

Abstract

An intelligent system embedded with multiple sources of knowledge may provide more robust intelligence with highly ill structured problems than the system with a single source of knowledge. This paper proposes the hybrid knowledge integration mechanism that yields the cooperated knowledge by integrating expert, user, and machine knowledge within the fuzzy logic-driven framework, and then refines it with a genetic algorithm (GA) to enhance the reasoning performance. The proposed knowledge integration mechanism is applied for the prediction of Korea stock price index (KOSPI). Empirical results show that the proposed mechanism can make an intelligent system with the more adaptable and robust intelligence.

Keywords: Hybrid knowledge integration, Fuzzy genetic algorithm, Human knowledge, Machine knowledge, The cooperated knowledge

1. Introduction

The common recognition in the field of decision support is that an intelligent system with integration of multiple sources of knowledge can deal with the ill-structured problems more effectively rather than with a single sources of knowledge (Blattberg and Hoch, 1990; Lee et al.; Kim et al., 1998). The typical types of knowledge which can be used for an intelligent decision support are machine knowledge, expert knowledge, and user knowledge (Lee et al.). Both expert and user knowledge means human knowledge. Experts have expertise or expert knowledge which is highly organized while users have shallow knowledge which is not well-organized. Therefore, decision making with expert knowledge can outperform that with user knowledge. Several researchers in behavioral science literature compared and investigated the characteristics of expert knowledge and user knowledge (Larkin et al., 1980; Lesgold et al., 1988; Chi et al., 1981; Einhorn, 1974; Goldberg, 1959; Hoch, 1987).

Machine knowledge is consistent because it is derived automatically by applying machine learning techniques 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 (Goldberg, 1970; Hongarth, 1978; Granger and Ramanthan, 1984; Lawrence et al., 1986; Conroy and Harris, 1987; Bunn, 1988; Clemen, 1989; Jo and Han, 1996). More recently, several works propose the fuzzy logic-based knowledge integration mechanisms that yields the cooperated knowledge by integrating multiple sources of knowledge including fuzzy logic-driven framework (Lee et al.) and fuzzy associative memory-based approach (Kim et al., 1998). Their findings show that they are promising in integrating multiple knowledge. However they may not ensure the best performance from the cooperated knowledge because the underlying ideas of them are not for the optimized integration. This may degrade the performance of the cooperated knowledge when applied to solving an ill-structured problem.

In this sense, this paper proposes a hybrid approach using a fuzzy genetic algorithm (FGA) for the optimized knowledge integration. In the proposed mechanism, fuzzy logic-driven framework yields the cooperated knowledge by integrating machine, user and expert knowledge. A genetic algorithm (GA) refines the cooperated knowledge by assigning the weights to machine knowledge and human knowledge derived from fuzzy logic-driven framework in terms of the relative importance of two kinds of knowledge. 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 cooperated knowledge is presented in section 3. The knowledge refinement by the GA 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 have shown the better reasoning performance by combining multiple results obtained from different models (Clemen, 1989; Jo and Han, 1996), 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 (Granger and Ramanathan, 1984). 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 (Granger and Ramanathan, 1984;. Bunn, 1988; Conroy and Harris, 1987; Goldberg, 1970; Hogarth, 1978; Lawrence et al., 1986; Pankoff, and Roberts, 1968). This type of combination means that integration of machine knowledge and human knowledge can reduce uncertainty involved with data and produce more elaborate results (Blattberg and Hoch, 1990).

Also there are some reasons for knowledge integration from the standpoint of behavioral science. In general, experts in a domain are assumed in behavioral literature to have expertise or expert knowledge which is highly organized and domain-specific enough to encode complex information (Larkin et al., 1980; Lesgold et al., 1988) and result in faster and more accurate performance (Chi et al., 1981). In addition, the behavioral science literature suggests that experts are better at knowing what questions to ask (diagnosis) than at predicting the future (Einhorn, 1974; Goldberg, 1959; Hoch, 1987). 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. Based on this claim, expert knowledge can be used for an intelligent guidance of user knowledge by providing diagnosis about the external factors that might affect stock market.

More recently, the fuzzy logic-based knowledge integration mechanisms have been proposed to deal with the uncertainty involved in the highly ill-structured problems including fuzzy logic-driven framework (Lee et al.) and fuzzy associative memory-based approach (Kim et al., 1998). Lee et al. propose fuzzy logic-driven framework which is capable of generating the cooperated knowledge by integrating multiple sources of knowledge. The findings show that the cooperated knowledge outperforms single sources of knowledge in terms of the reasoning performance. However, fuzzy logic-driven framework generates the cooperated knowledge pertaining to a case or an object so that it suffers from the conflicts between multiple sources of knowledge (Kim et al., 1998). This may degrade the robustness of an intelligent system with multiple sources of knowledge and result in local optima.

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

The cooperated knowledge, in which expert, user, and machine knowledge are integrated, can be created first by applying the same fuzzy logic-driven approach (Lee et al.). It consists of three phases: (1) machine knowledge-based inference (MKBI) phase (2) expert knowledge-based inference (EKBI) phase, and (3) combining phase. There is no explicit phase for user knowledge because it is incorporated into both MKBI phase and EKBI phase during the operation of inference. User knowledge is expressed in MKBI phase as one of three evaluation categories 'good', 'not good or not bad', and 'bad' which represents the current state of a specific technical indicator. In EKBI phase, user knowledge is described as one of five evaluation categories for external factors: very bad (0), bad (1), not good or not bad (2), good (3), and very good (4). In this way, user knowledge is incorporated into both MKBI and EKBI phases.

3.1 MKBI Phase

Variables and Data selection

Machine knowledge (MK) is obtained by applying backpropagation neural network model to technical indicators obtained from KOSPI. We collect 9 technical indicators indicating dynamic trends of stock price index including Moving Average (MA), Relative Strength Index (RSI), PSYchology (PSY), MOMentum (MOM), STOchastic %D (STOD), Volume Ratio (VR), On Balance Volume (OBV), DISparity (DIS), and Rate Of Change (ROC). The formulas for technical indicators are shown in Table 1.

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. If a return of the next month is greater than 3%, the corresponding stock market level is regarded as Bull. Similarly, Edged-Up, Edged-Down, and Bear when the KOSPI return is between 0% and 3%, between -3% and 0%, and less than -3%, respectively.

We apply two stages of input selection process. At the first stage, we select 7 variables by one-way ANOVA between technical indicators and stock market level. In the second stage, we select 5 technical indicators, MA, RSI, STOD, DIS, and ROC, by a stepwise method. Each technical indicator is expressed in fuzzy membership function for three categories: 'good', 'not good' and 'bad', which are determined by users' decision. The results of statistical analysis are given in Table 2.

Variables	Formula
MA	6 day moving average of the closing price
RSI	(the sum of closing values in positive KOSPI change for 25 days / the sum of closing values for 25 days) × 100
PSY	(days of positive KOSPI change for 12 days / 12) × 100
MOM	(the latest closing price – the closing price 6 days ago)
STOD	[(the most recent closing price – the lowest of low price for the last 6 days) / (the highest of high price for the last 6 days – the lowest of low price for the last 6 days)]
VR	[(the sum of volume in positive KOSPI change for 6 days – the sum of volume in negative KOSPI change for 6 days) / the sum of volume for 6 days] × 100
OBV	(the sum of volume in positive KOSPI change – the sum of volume in negative KOSPI change
DIS	(the most recent closing price / 6 day MA of price)× 100
ROC	(the price of the latest day / the price of 6 days ago)

Table 1. Formulas of Technical Indicators

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), genetic learning (52 weeks from January 1993 to December 1993), and validation (104 weeks from January 1994 to December 1995). The descriptive statistics of data sets are listed in Table 3.

Variables	Wilks' λ	ANOVA
MA	.979	5.777*
RSI	.964	18.121**
PSY		2.254
MOM		19.924**
STOD	.971	9.600**
VR		19.351**
OBV		2.399
DIS	.881	22.978**
ROC	.953	13.702**

(* : significant at 5% ** : significant at 1%)

Table 2. The Statistics of One-Way ANOVA and the Stepwise MDA

Neural Network Architecture

Input variables selected for the neural network are the five technical indicators such as MA, RSI, STOD, DIS, and ROC whose selection criterion is shown in Table 2. Specific values for each technical indicator are 'good', 'not good or not bad', and 'bad', which are determined by users' knowledge. Output values from the neural network are the four categories such as Bull, Edged-up, Edged-down, and Bear.

MKBI Phase

We can define a fuzzy prediction vector of MK (FPV^{MK}) as fuzzy membership function derived from a neural network.

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

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

$$FPV^{MK} = (.5390\ .8520\ .1222\ .1012)$$

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

Level / Data	Bear		Edged-Down		Edged-Up		Bull	
	Sample Size	Average Returns	Sample Size	Average Returns	Sample Size	Average Returns	Sample Size	Average Returns
Neural Net Training	36	-3.98	209	-1.20	184	1.20	64	5.92
GA Learning	2	-5.13	19	-1.31	26	1.48	5	4.40
Validation	8	-3.58	51	-1.23	34	1.40	11	3.9
	46	-3.96	279	-1.12	244	1.26	80	5.55

Table 3. The Descriptive Statistics of Data Sets

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 (CFU). EP means the forecast about economic situation, which is determined by the composite effects of export, GNP, and inflation, etc. Those factors affecting SSD include capital-increase of listed firms, new stock supply, and the investment activities of institution. AOC is determined by bond yield, call rate of interest, the amount of cash deposited in stockbrokers, and monetary policy of government. CFU is a broad concept. For example, the political situations, domestic or international, might affect the stock market trend. Also the news background with respect to stock market could influence the investor's decision to buy or sell the corresponding stocks. Therefore CFU covers from macro factors to micro factors. 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-typed 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 typed-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 $\mu^i_{UK}(x)$ and $\mu^i_{EK}(y)$ 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_i^K) as follows:

$$\mathrm{FEV}_i^K = (\mu^i_K(\mathrm{very\_bad}), \mu^i_K(\mathrm{bad}), \mu^i_K(\mathrm{not\_bad}), \mu^i_K(\mathrm{good}), \mu^i_K(\mathrm{very\_good}))$$

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

$$\mathrm{FEM}^K = [\mathrm{FEV}_i^K], i = 1,2,3,4$$

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 $FEMs$ for UK and EK respectively as follows:

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

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

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

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

Let us define the weight vector to consider the effect of each factor on five evaluation categories as $W=(W_1, \ldots, W_m)$, $0<w<1$ 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 $CFEM^{HK}$, HK-based combined fuzzy evaluation vector ($CFEV^{HK}$) can be calculated as follows: $CFEV^{HK} = W \times CFEM^{HK} = (.5425\ .5825\ .25\ .0\ .0)$.

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 (FPV^{HK})

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

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:

$$FPV^{CK} = (FPV^{HK} \wedge FPV^{MK})$$
$$= (.5390\ .8520\ .1222\ .1012) \wedge (.5625\ .4163\ .1250\ .0)$$
$$= (.5390\ .4163\ .1222\ .0).$$

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 Knowledge Refinement by the Genetic Algorithm

The central task of the GA in this paper is to find the optimal or near optimal weights vector to be assigned for machine and human knowledge derived from fuzzy logic-driven framework. The genetic evolution process for knowledge integration consists of the following steps:

Step 1: The genetic coding

Step 2: Population initialization

Step 3: The fitness evaluation of a set of weights.

Step 4: Operation of fundamental operators: reproduction, crossover and mutation.

Step 5: Iteration

In step 1, we can define fuzzy vector of actual outcomes (FV^{AO}) to convert actual outcome into the triangular typed-fuzzy membership function with the width 2. For example, a particular actual outcome, Edged-Down, can be converted as following:

$$FV^{AO} = (0.5\ 1.0\ 0.5\ 0.0).$$

We use a real numbered string to represent a particular case or object, where each real number represents the membership of FPV^{MK}, FPV^{HK} and FV^{AO}. If the actual outcome of the case illustrated in the previous section is Edged-Down, then we can encode it as followings:

(.5390 .8520 .1222 .1012; .5625 .4163 .1250 .0: 0.5 1.0 0.5 0.0)

Let us define the weights vector assigned to FPV^{HK} and FPV^{MK} as followings:

$$W = [W_K(i)]$$
$$\text{for } K = MK, HK$$
$$i = \text{bear}, \text{edged}-\text{down}, \text{edged}-\text{up}, \text{bull}$$

where the weight vectors to determine the importance between machine and human knowledge in each level is ranged between 0 and 1.

The GA generates the weights vector which is used to adjust FPV^{CK} to FV^{AO} approximately so that the fuzzy prediction vector of the cooperated (FPV^{CK}) refined by the GA is calculated as the weighted sum of fuzzy membership functions of FPV^{HK} and FPV^{MK}. FPV^{CK} is mathematically represented as followings:

$$FPV^{CK} = \begin{bmatrix} W_{MK}(\text{Bear}) \times \mu_{MK}(\text{Bear}) + W_{HK}(\text{Bear}) \times \mu_{HK}(\text{Bear}), \\ W_{MK}(\text{Edged-Down}) \times \mu_{MK}(\text{Edged-Down}) \\ \quad + W_{HK}(\text{Edged-Down}) \times \mu_{HK}(\text{Edged-Down}), \\ W_{MK}(\text{Edged-Up}) \times \times \mu_{MK}(\text{Edged-Up}) \\ \quad + W_{HK}(\text{Edged-Up}) \times \mu_{HK}(\text{Edged-Up}), \\ W_{MK}(\text{Bull}) \times \mu_{MK}(\text{Bull}) + W_{HK}(\text{Bull}) \times \mu_{HK}(\text{Bull}) \end{bmatrix}$$

With FPV^{CK} related to a given case, the market level with maximum value is regarded as the level of next week. The weights assigned to machine and human knowledge means the relative importance of two kinds of knowledge. We can calculate the value of importance (*VOI*) as followings:

$$VOI_i = \frac{W^{HK}(i)}{W^{MK}(i)}$$

where i means stock market levels. The higher value of importance in a given level indicates that human knowledge is more important and useful in predicting the corresponding level than machine knowledge, and vice versa. One way to assign the weights to machine and human knowledge is often performed by human experts if the experts are expected to have appropriate expertise to decide which type of knowledge is more important in predicting each level. However, it is a difficult and time-consuming task. Instead of using human experts, we use the machine learning approach using GAs to learn optimal weight vector from historical instances by evaluating the fitness of different sets of weights.

In step 2, genetic search begins by randomly generating an initial population of strings. The population size is a compromising factor. Large population size increases the possibility of including the solution in the first few generations but decreases the running speed of the GAs. We use population size of 50.

In step 3, the fitness of a set of weights vector in the population is evaluated together against the entire training cases. The fitness function is for the minimization of the average deviation (*AD*) which is the performance measure of each weights vector in terms of accuracy. It is the average of the absolute differences between the actual levels (*AL*) and the predicted levels (*PL*) where both levels are represented as one of four values, Bear (1), Edged-Down (2), Edged-Up(3), and Bull (4). It is. The fitness function is mathematically expressed as followings:

$$\text{Min} \quad AD = \frac{1}{n}\sum_{i=1}^{n}|AL_i - PL_i|$$

$$\text{s.t} \quad AL_i \text{ or } PL_i = \begin{cases} 1 & \text{if each level is Bear} \\ 2 & \text{if each level is Edged-Down} \\ 3 & \text{if each level is Edged-Up} \\ 4 & \text{if each level is Bull} \end{cases}$$

$$PL_i = \text{Max}(FPV^{CK})$$
$$\text{for a given } i \ (i = 1, 2, \ldots, n)$$

In step 4, a subset of the weights vector with good fitness value is selected as parents, while the poor ones will be eliminated. Parents can reproduce offspring by using fundamental operators such as crossover and mutation to continue the search for optimal solution. A new population pool of the same size as the original is then created with a higher average fitness value. We set the crossover and mutation rate as the range of 0.5-0.7 and 0.06-0.1, respectively.

In step 5, the GA run iteratively repeating the steps 2-4 until it arrives at a predetermined stopping condition. We also set 1,000 generation as the stopping condition.

5. Empirical Results

The weights vector derived from GAs and the value of importance are as followings:

$$W = [w^{Bear}_{MK}, w^{ED}_{MK}, w^{EU}_{MK}, w^{Bull}_{MK}, w^{Bear}_{HK}, w^{ED}_{HK}, w^{EU}_{HK}, w^{Bull}_{HK}]$$
$$= [0.085 \quad 0.7512 \quad 0.9476 \quad 0.1212 \quad 0.687 \quad 0.1796 \quad 0.1512 \quad 0.7453]$$

$$\textbf{The value of Importance} = [VOI^{Bear}, VOI^{Edged\text{-}Down}, VOI^{Edged\text{-}Up}, VOI^{Bull}]$$
$$= [8.80,\ 0.24,\ 0.16,\ 6.15]$$

The values of importance of Bear and Bull levels are relatively higher than those of Edged-Down and Edged-Up levels. Bear and bull levels are regarded as the turbulent situation that stock price is rapidly changing and the regularities are hardly found. On the contrary, the values of importance of Edged-down and edged-up level are smaller than those of bear and bull level. These findings are supported by the arguments that human knowledge is more adaptable to the changing environment than machine knowledge, while machine knowledge performs well in dealing with the regularities (Ren and Sheridan, 1995).

The comparative analysis of multiple sources of knowledge, the cooperated knowledge, and the refined cooperated knowledge with the GA is shown in Table 4. The deviation column indicates the difference between the actual status and the predicted status.

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 user's decision making (Lee et al., 1989).

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 user's decision making (Lee et al., 1989).

Deviation	UK		EK		HK		MK		CK		CK-FGA	
	No.	Propo-rtion	No	Propo-rtion	No.	Propo-rtion	No.	Propo-rtion	No.	Propo-rtion	No.	Propo-rtion
0	41	39%	46	44%	48	46%	53	51%	54	52%	66	66
1	55	53%	49	47%	48	46%	40	38%	43	42%	31	30
2	7	7%	8	8%	7	7%	11	11%	6	5%	7	4
3	1	1%	1	1%	1	1%	0	0%	1	1%	0	0
Average Deviation	0.6923		0.6538		0.625		0.5962		0.5577		0.4327	

Table 4. The Performance of Different Sources of Knowledge (n=104)

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 FGA (CK-FGA:) outperforms than CK.

We use Wilcoxon matched-pairs signed-ranks test to examine whether the predictive performance of the FGA is significantly higher than that of other techniques. Table 5 shows the results of Wilcoxon matched-pairs signed-ranks test to compare the prediction performance between benchmark models and the FGA approach. As shown in Table 5, statistical results show that the FGA performs significantly better than any other knowledge at 1% level.

	UK	EK	HK	MK	CK
EK	1.269				
HK	1.305	1.342			
MK	1.366	1.380	0.447		
CK	2.858***	2.236**	1.658*	0.632	
CK-FGA	4.536***	4.321***	3.832***	505***	2.968***

(*: significant at 10% ** : at 5% *** : at 1%)

Table 5. The results of Wilcoxon matched-pairs signed-ranks test

Based on the empirical results, we concluded that the hybrid knowledge integration mechanism to assign weights derived from the genetic search process can be the most effective one as GAs find optimal or near optimal solution for the specified objective function.

6. Concluding Remarks

Machine knowledge and human knowledge coexist in real decision-making environment. These two types of knowledge are not exclusively independent but interact with each other. How to integrate and coordinate different types of knowledge together is an open research area.

Fuzzy logic-based knowledge integration mechanisms have shown the improved reasoning performance by integrating multiple sources of knowledge. However, they may inappropriate for the optimized knowledge integration.

This paper presents the hybrid approach using fuzzy logic and the GA to the optimized knowledge integration. In this approach, fuzzy logic-driven framework generates the cooperated knowledge by integrating user knowledge, expert knowledge, and machine knowledge. The GA assigns the weights to multiple sources of knowledge by means of their relative importance. We demonstrate that the proposed knowledge integration mechanism can improve overall predictive performance significantly as well as provide the more adaptable and robust intelligence for the prediction of KOSPI.

However, more research is needed to further improve the reasoning performance. One issue for further research is related to tuning membership function. For instance, we assume that the triangular-typed membership functions are given for UK and EK. Therefore we need to find out what is the most effective way to fuzzify these variables into linguistic terms and defuzzify them back to numerical numbers. One solution for tuning fuzzy membership function is using some AI techniques such as genetic algorithms (Karr, 1991) and neural networks (Lee et al., 1996). Another is the development of the refined knowledge integration mechanisms using other techniques. For instance, fuzzy rule-based system can be applied to knowledge integration, providing the explanatory capability to an intelligent system.

References

Blattberg, R.C. and Hoch, S.J., 'Database models and managerial intuition: 50% model + 50% manager', *Management Science*, **36**, 8, 1990, 887-899.

Bunn, D.W., 'Combining forecasts', *European Journal of Operations Research* **33**, 1988, 223-229.

Chi, M.T.H. Glaser, R. and Rees, E., 'Expertise in problem-solving'. in R.J. Sternberg (ed.), *Advances in the Psychology of Intelligence*, Erbaum, Hilldale, NJ, 1981.

Clemen, R.T., 'Combining forecasts: a review and annotated bibliography', *International Journal of Forecasting* **4**, 1989, 559-584.

Colin, A.M., 'Genetic algorithms for financial modeling, In G.J. Deboeck(Ed.). *Trading on the egdge*, New York: Wiley, 1994, pp. 174-188.

Conroy, R. and Harris, R., 'Consensus forecasts of corporate earnings: analysts' forecasts and time series methods', *Management Science* **33**, 1987, 725-738.

Einhorn, H.J., 'Cue definition and residual judgment', *Organizational Behavior and Human Performance* **12**, 1974, 30-49.

Goldberg, L., 'The effectiveness of clinician's judgments: the diagnosis of organic brain damage from the Bender-Gestalt test', *Journal of Consulting Psychology* **23**, 1959, 25-33.

Goldberg, L., 'Man vs. model of man: rationale, plus some evidence, for a method of improving on clinical inferences', *Psychological Bulletin* **73**, 1970, 422-432.

Granger, C.W.J. and Ramanathan, R., 'Improved methods of forecasting', *Journal of Forecasting* **3**, 1984, 197-204.

Hoch, S.J., 'Perceived consensus and predictive accuracy: the pros and cons of projection', *Journal of Personality and Social Psychology* **53**, 1987, 221-234

Hogarth, R.M., 'A note on aggregating opinions', *Organizational Behavior and Human Performance* **21**, 1978, 40-46.

Holland, J.H., '*Adaptation in Natural and Artificial Systems*', Univ. of Michigan press, Ann Arbor, MI, 1975.

Ishibuchi, H., Nozaki, K., Yamamoto, N. and Tanaka, H., 'Construction of fuzzy classification systems with rectangular fuzzy rules using genetic algorithms', *Fuzzy Sets and Systems* **65**, 1994, 237-253.

Jo, H.K. and Han, I.G., 'Integration of case-based forecasting, neural network, and discriminant analysis for bankruptcy prediction', *Expert Systems With Applications* **11**(4), 1996, 415-422.

Karr, C.L., 'Design of an adaptive fuzzy logic controller using a genetic algorithm', *Proceedings of 4th International Conference on Genetic Algorithms*, San Diego, CA, July 1991, 450-457.

Kim, M.J., Han, I.G. and Lee, K.C., 'The integration of machine and human knowledge by fuzzy logic for the prediction of stock price index', *5th Pacific Rim International Conference on Artificial Intelligence*, Singapore, 1998, 260-271.

Koza, J., '*Genetic programming*', Cambridge, MA: The

MIT Press., 1993

Larkin, J., McDermott, J., Simon, D.P. and Simon, H.A. 'Expert and novice performance in solving physics problems', *Science* **208**, 1980, 1335-1342.

Lawrence, M.J., Edmundson, R.H. and O'CONNOR, M.J., 'The accuracy of combining judgmental and statistical forecasts', *Management Science* **32**, 1986, 1521-1532.

Lee, J.K., Kim, H.S. and Chu, S.C., 'Intelligent stock protfolio management system', *Expert Systems* **6**(2), 1989, 74-87.

Lee, K.C., Han, J.H. and Song, Y.U., 'A fuzzy logic-driven multiple knowledge integration framework for improving the performance of expert systems', forthcoming in *Intelligent Systems in Accounting, Finance and Management.*

Lee, K.M. Kwak, D.H. and Kwang, H.L., 'Fuzzy inference neural network for fuzzy model tuning', *IEEE Transactions on Systems, Man, and Cybernetics.* **26**(4), August 1996, 637-645.

Lesgold, A., Rubinson, H., Feltovich, P., Glaser, R. and Klopfer, D., 'Expertise in complex skills: diagnosing X-ray pictures'. in M.T.H. Chi, R. Glaser and M.J. Farr (ed.), *The Nature of Expertise*, Erlbaum, Hilssdale, NJ, 1988.

Normura, H., Hayashi, I. and Wakami, N., 'A Self-tuning method of fuzzy reasoning by genetic algorithm', *Proceedings of 1992 International. Fuzzy Systems and Intelligent Control Conference*. Louisville, KY, 16-18. March 1992, 236-245.

Pankoff, L.D. and Roberts, H.V., 'Bayesian synthesis of clinical and statistical prediction', *Psychological Bulletin* **80**, 1968, 762-773.

Ren, J. and Sheridan, B., 'An adaptive decision aid for real enviornments', *IEEE Transactions on systems, Man, and Cybernetics*, **25** (10), October 1995, 1384-1391

Rumelhart, D., Hinton, S.S. and Williams, R., 'Learning internal representation by error propagation'. in *Parallel Distributed Processing*, Vol. I, MIT Press, Cambridge, MA, 1986.

Yuan, Y. and Zhuang, H., 'A genetic algorithm for generating fuzzy classification rules', *Fuzzy Sets and Systems*, **84**, 1996, 1-19