

Integrated Model for Informal Inference Based on Neural Networks*

Kyung-Joong Kim and Sung-Bae Cho

Department of Computer Science, Yonsei University
134 Shinchon-dong, Sudaemoon-ku, Seoul 120-749, South Korea
kjkim@cs.yonsei.ac.kr, sbcho@cs.yonsei.ac.kr

Abstract. Inference is one of human's high-level functionalities and it is not easy to implement in machine. It is believed that inference is not results of single neuron's activity. Instead, it is a complex activity generated by multiple neural networks. Unlike computer, it is more flexible and concludes differently even for the similar situations in case of human. In this paper, these characteristics are defined as "informality." Informality in inference can be implemented using the interaction of multiple neural networks with the inclusion of internal or subjective properties. Simple inference tasks such as pattern recognition and robot control are solved based on the informal inference ideas. Especially, fuzzy integral and behavior network methods are adopted to realize that. Experimental results show that the informal inference can perform better with more flexibility compared to the previous static approaches.

Keywords: Informal Inference, Neural Networks, Robot Control, Pattern Recognition, Fuzzy Integral, Behavior Network.

1 Introduction

Biological neural networks are composed of a number of modules specialized to a specific task and integrate them for high-level functions. In an engineering perspective, there are three practical problems in realizing this biological anatomy: the selection of appropriate neural network architecture, the learning mechanism for each model and the integration method. It is very important that a solution for each problem has to show the similar characteristics with the biological counterpart. Meanwhile, it also has to be useful in an engineering perspective: the system might satisfy the performance requirement.

Inference is one of the critical brain's functions and many researchers have been attempted to mimic it by using symbolic or connectionist models [1][2]. Using the function, human reaches a conclusion from evidences (observations and facts). Unlike the artifacts, human's inference is not always the same and sometimes it is flexible and dynamic. In short, it is informal. The traditional engineering methods are well defined in the perspective of formality and it is required to adopt a new method for the informality. Figure 1 shows the difference between two inference models.

* This research was supported by Brain Science and Engineering Research Program sponsored by Korean Ministry of Commerce, Industry and Energy.

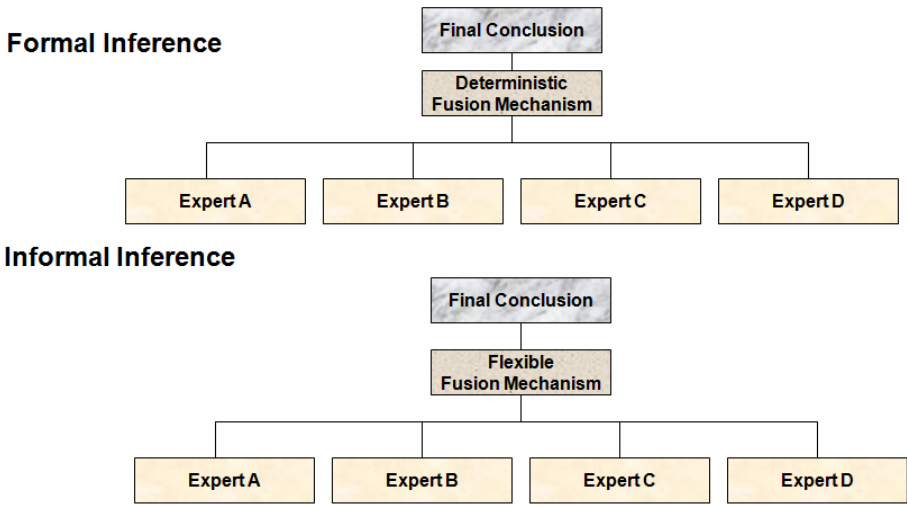


Fig. 1. Comparison of two inference methods

The easiest way to model the informality on a set of neural networks is to derive computational method for the connections among the neural models. Based on [3], the integration within and among specialized areas of human's brain is mediated by effective connectivity. Behavior network is a computational model that is derived from the research of action selection of animals [4]. It connects a number of behavioral modules using symbolic connections and selects the best action for each sensory inputs and motivation by propagating activation through the network. It is more flexible and dynamic than strict rule sets.

Sometimes, a human expert can incorporate his or her subjectivity on the preference of the neural modules and the final inference becomes more informal than one without user's preference. Fuzzy integral is a mathematical model that provides a systematic way to incorporate user's subjective preference on the neural models [5]. It is more flexible than the static combination rules.

In this paper, the behavior network and fuzzy integral are used to realize the informal inference in the real-world applications. They are controlling mobile robot using multiple neural behavior modules [6] and classifying web documents preference based on user's subjectivity on the neural modules [7]. For each problem, we have used different neural models that are integrated with more flexible methods.

2 Related Works

The term "informal inference" is not common word and not frequently used in engineering domain. Traditional symbolic logic is one of the examples of formal logic but it is not flexible to deal with human's generalization capability. There are many works to implement the inference capability using connectionist models like neural networks [1]. In probability theory society, the "informal inference" is used as the same meaning with plausible inference (probabilistic inference) [8]. In statistics

education society, the term is used as the concept compared to the “formal statistics” [9]. Usually, adults capture statistical concepts using statistical formula and logics. But it is not easy for children and they use pictures, trends and patterns to grasp high-level statistical information. These inference processes are called as “informal inference.”

There are a number of ensemble methods for neural networks and they are well summarized in [10]. Verikas *et al.* used “soft combination” to refer combination methods with more flexibility. They include fuzzy integral as one of the group of soft combination. In robotics community, they tried to incorporate motivation of the agent to the multi-module controller [4]. This allows the ensemble of multiple modules can be flexible to the change of internal motivation and user’s subjective preference on each module.

3 Informal Inference Based on Neural Networks

In this paper, two computational methods are used to implement informal inference that shows different behaviors based on user’s internal preference and agent’s internal motivation. Fuzzy integral combines a number of neural networks with user’s preference on each module. Behavior network combines multiple neural networks for high-level behaviors reflecting internal motivation of agents.

3.1 Fuzzy Integral

The definition of fuzzy integral is as follows. Let $C=\{c_1, c_2, c_3, \dots, c_N\}$ be a set of classes, where binary classification problem has $|C|=2$. Let $Y=\{y_1, y_2, \dots, y_n\}$ be a set of classifiers. $h_k(y_i)$ is an indication of how certain we are in the classification of web page to be in class c_k using the network y_i . Usually, the $h_k(y_i)$ is assigned as the real value from k th output neuron (normalized from 0 to 1) for the input. $A_i=\{y_1, y_2, \dots, y_i\}$ is a partition of the set Y . $G=\{g_1, g_2, \dots, g_n\}$ is a set of user’s preference value on each classifier. Based on user’s evaluation, we calculate λ .

$$\lambda + 1 = \prod_{i=1}^n (1 + \lambda g^i) \quad \lambda \in (-1, +\infty) \text{ and } \lambda \neq 0. \tag{1}$$

From calculated λ , the fuzzy measure value of each A_i is decided based on following equation.

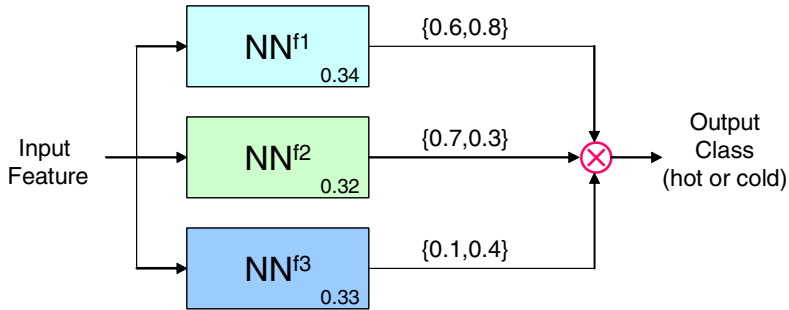
$$\begin{aligned} g(A_1) &= g(\{y_1\}) = g^1 \\ g(A_i) &= g^i + g(A_{i-1}) + \lambda g^i g(A_{i-1}), \text{ for } 1 < i \leq n. \end{aligned} \tag{2}$$

The final classification of multiple neural networks is determined based on user’s evaluation and the fuzzy measure values for each classifier set.

$$\text{Final class} = \operatorname{argmax}_{c_k \in C} \left[\max_{i=1}^n [\min(h_k(y_i), g(A_{ki}))] \right] \tag{3}$$

For example, each neural network (NN) determines the class label of an unknown document as “0” or “1” (in the binary classification problem). If $SASOM_1$ classifies

the document as “0”, $h_0(NN_1)=1.0$ and $h_1(NN_1)=0.0$. Supposed that there are three NN's, user evaluates classifiers as g^1, g^2 and g^3 , respectively. λ is calculated from g^1, g^2 , and g^3 . It is easily determined from the 2nd degree polynomial based on (1). For each class k , classifiers are sorted by $h_k(NN_i)$. By the sorted order, they are labeled as y_1, y_2 and y_3 . With $g(y_1), g(y_1, y_2)$ and $g(y_1, y_2, y_3)$, the class label of the unknown document is determined using formula (3). Figure 2 shows an example of fuzzy integral calculation



class	$h(y_i)$	$g(A_i)$	$H(E)$	Max[$H(E)$]
hot	0.7	$g(\{y_2\})=g^2=0.32$	0.32	V
	0.6	$g(\{y_2, y_1\})=g^2+g^1+\lambda g^2 g^1=0.66$	0.6	
	0.1	$g(\{y_2, y_1, y_3\})=1.0$	0.1	
cold	0.8	$g(\{y_1\})=g^1=0.34$	0.34	V
	0.4	$g(\{y_1, y_3\})=g^1+g^3+\lambda g^1 g^3=0.67$	0.4	
	0.3	$g(\{y_1, y_3, y_2\})=1.0$	0.3	

Fig. 2. An example of fuzzy integration

3.2 Behavior Network

A behavior network is defined as a set of topology and the parameters. The topology means the relationships among nodes, sensors and goals. If the sensor is the precondition of the behavior, there is a link between the behavior node and the sensor node. It means that the sensor has to be true if the robot wants to execute the behavior. If a behavior is directly related to achieve the goals, there is a link between the behavior node and the goal node. Among behaviors, there are two different kinds of connections and they are called as predecessor and successor links. If one behavior is possible to help the other behavior's future execution, there is a predecessor link between the two behaviors. The successor link is the reverse of the predecessor link if one behavior can help other behavior right now (executable without delay). The topology is manually designed by experts.

The parameter part of the behavior network is related to the strength of the links. For each link type, there is different weight value and it controls the strength of the

inputs, goals and the internal relationships among behaviors. The inputs from the environmental sensors are multiplied by ϕ and those from goals are readjusted by multiplying γ to the original input value. The weight of the successor link is defined as ϕ/γ and the one of the predecessor link is 1. The parameters are determined by designer manually.

After constructing the architecture (topology + parameters) of the behavior network, it is required to define the selection procedure based on it. If the architecture is fixed, we can propagate the activation of each behavior. The first step is to input the values from sensors into the behavior through environmental links. After then, inputs from goals that are representing the internal status of the robot are inputted to the behavioral node through goal links. Then, the activation of each behavior node is updated by propagating activation through links among behaviors. Based on the type of link, it adds activation to other node or reduces it from other's one. After finishing all updating, we choose candidate behaviors if they satisfy all preconditions and have larger activation than threshold value. Among candidates, the one with the highest activation is finally selected as a final winner. If there is no candidate, the threshold is reduced by 10% and the activation update is repeated. Figure 3 shows the algorithm in a pseudo code manner.

```

WHILE (1) {
  initialization();
  spreading activation();
  normalization();
  FOR all behaviors {
    IF (all preconditions are true
      && activation (behavior) > threshold) {
      candidate (behavior);
    }
  }
  IF (candidate () = NULL) {
    threshold = 0.9 * threshold;
  }
  ELSE{
    select();
    break;
  }
}

```

Fig. 3. A pseudo code for behavior network action selection

4 Experimental Results

The two methods are tested in different two applications. The first method is evaluated on pattern classification problem which classify HTML document into one of “hot” and “cold” class. “Hot” means that the page is preferable by the user. The dataset is downloaded from UCI benchmakr repository and its name is “Syskill & Webert.” There

are two different types of groups in HTML webpages: Goats and Bands. The second method is evaluated using robot controlling problem. The controller is composed of multiple neural modules evolved. For each time step, one of modules can get the control of robot and behavior network is used to choose of the multi-modules. Internal goal of agent is modelled in the network of action selection. In both cases, we tried to compare them with static methods (rules and statical combination scheme).

4.1 Web Document Classification

From the UCI KDD database, Syskill & Webert data that have web documents and user's preference value ("hot" or "cold") are accessible. Syskill & Webert data have four different topics "Bands," "Biomedical," "Goats," and "Sheep," among which we use "Goats" and "Bands" data.

"Goats" data have 70 HTML documents and "Bands" 61 HTML documents. Each document has the class label of "hot" or "cold." Each HTML file contains texts related with the topic. Rating file contains file name, rating, URL, date and title orderly. Preprocessing of web documents constructs input vector with selected features and class label. From training data, we extract k important features using three different feature selection methods. Each method ranks all features by different manner. Document $D = \langle v_1, v_2, v_3, \dots, v_{128}, c \rangle$ has three different input vectors that are used to train SASOM (Structure-Adaptive Self-Organizing Map) [12].

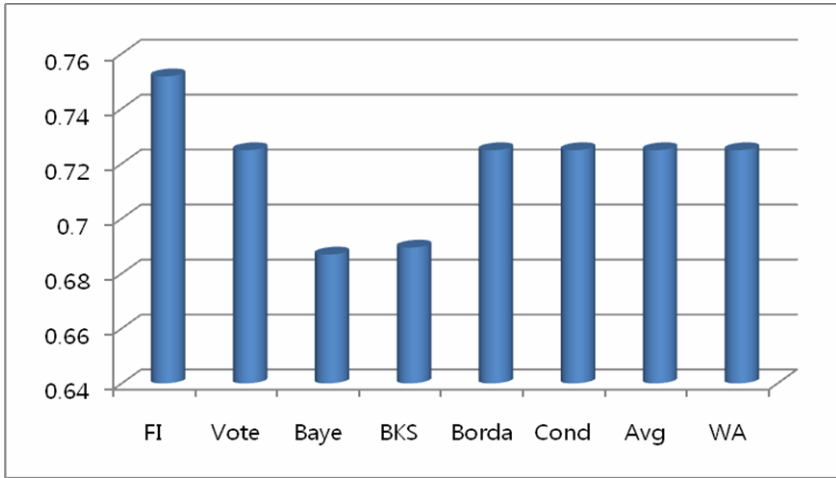
The problem to solve is to predict unknown documents' classes using known web documents with fuzzy integration of three different SASOM's trained using the input vectors. Experiments are repeated 10 times and the result is the average of them. For comparison, representative combination methods are used. Figure 4 shows the classification accuracy comparison (FI=Fuzzy Integral, BKS=Behavior Knowledge Space, WA = Weighted Average). It shows that the proposed fuzzy integral outperforms the other static combination methods for both dataset. Also, it allows user assign the preference for each classifier.

SOM is very powerful to visualize the classification results as a map of 2D structures and SASOM also has such capability. After training each SASOM, user can watch each SASOM's classification results (clusters) on 2D maps. Based on this, user can give preference on each SASOM. It allows user tune his system by changing preference for each SASOM. This integrated system can progressively improve his performance with the interaction of user. In this context, visualization of classifier's structure and results is quite important.

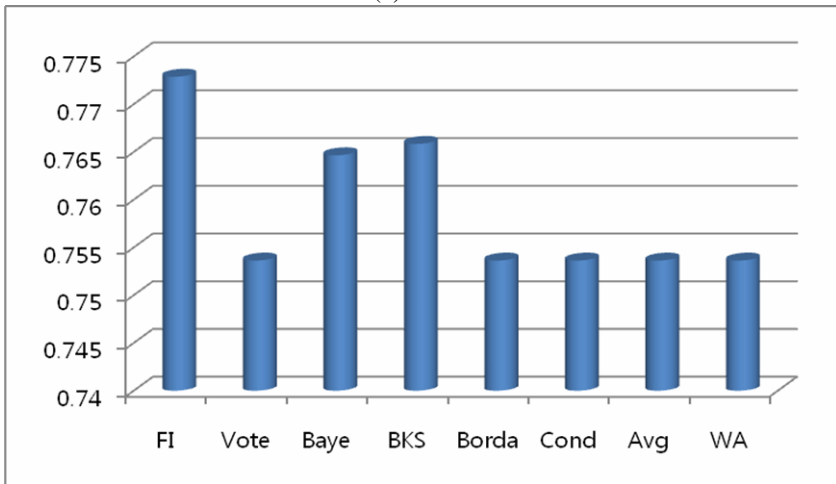
4.2 Robot Control

In robot control problem, four different modules are used. Among them, following light and avoiding obstacle behaviors are implemented using evolutionary neural networks. The four behavior modules are as follows.

- Recharging Battery: If a robot arrives at a battery recharge area, the battery is recharged. This behavior enables the robot to operate for as long as possible.
- Following Light: The robot goes to a stronger light. This behavior can be used to make the robot go to the battery recharge area, because the light source is located in that area.



(a) Goats



(b) Bands

Fig. 4. Comparison of classification accuracies for both dataset

- Avoiding Obstacles: If there are obstacles around the robot, it avoids them without bumping against them.
- Going Straight: If there is nothing around the robot, it goes straight ahead. This behavior allows it to move continuously without stopping.

The problem to solve is that maintaining robot’s life in simulation environment by minimizing bumping to the walls. Robot has initial battery level (2500) and it decreases 1 if it moves. To survive, robot has to go to battery recharge area (upper left area) and executes the battery recharge behavior automatically. The coordination task is choosing one of four behaviors at each time appropriately based on robot’s sensory

and internal motivation. The robot's motivation is to survive long time by minimizing bumping. In previous research, the coordination task is implemented using rule-based logic [13] and behavior network [6]. In this experiment, we focus on the adaptability of the two coordination methods. In the environment, light source is placed in the battery recharge area (represented as black arc). It allows robot find the area automatically. The coordination mechanism is evaluated in 3 different new environments (figure 5).

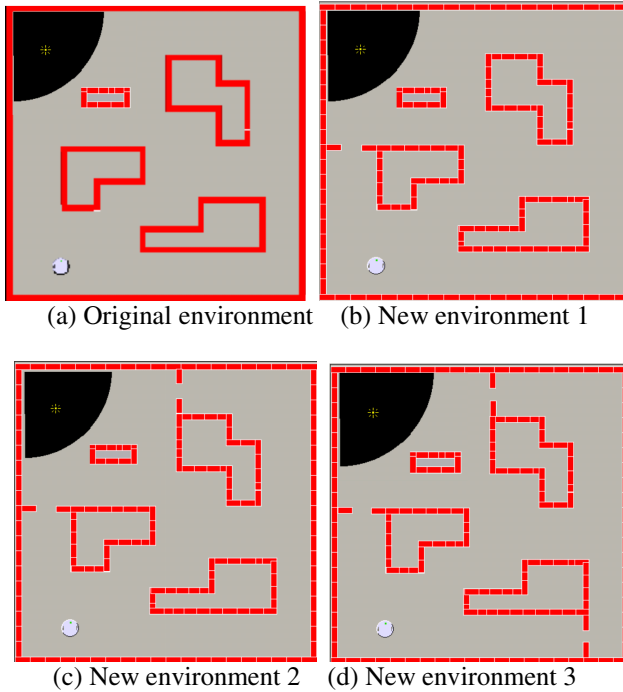


Fig. 5. Original and 3 different new environments

Table 1. Comparison of performance for 3 new environments (100 runs for each model, N =Number of cases that recharge battery is done more than 1, A =Average battery usage for the successful cases)

		Rule-based Model	Behavior Network
New environment 1	N	3	2
	A	4210.333	5156.5
New environment 2	N	8	10
	A	4293.75	4861.2
New environment 3	N	4	3
	A	4100.75	5104.333

Table 1 shows statistics on the experimental results. Because the environment is changed, the successful rate is not high. For the number of successful cases, there is no significant difference between the two models. However, the average life time (battery usage) is larger in behavior network than one in rule-based model. This shows that the behavior network can work better than the rule-based model in changed environment.

5 Conclusion and Future Works

In this paper, informal inference is defined as a function of human's flexible inference that response differently to the similar external sensory inputs. This is because human has his own preference or motivation inside of his brain. Two computational models are used to simulate this behavior and compared with static models. Experimental results show that the proposed method performs well compared to the rule-based and statistics-based model.

There are still some parameters on forming flexible inference systems and it makes difficult to use of the models. The learning mechanism for the flexible inference models are required to minimize human's intervention in deciding parameters of the system. In fuzzy integral, system can recommend some preference parameters to the user based on some statistical information. It can save user's fatigue to use the system.

References

- [1] Browne, A., Son, R.: Connectionist inference models. *Neural Networks* 14(10), 1331–1355 (2001)
- [2] D'Ambrosio, B.: Inference in Bayesian networks. *AI Magazine* 20(2), 21–35 (1999)
- [3] Friston, K.: Learning and inference in the brain. *Neural Networks* 16(9), 1325–1352 (2003)
- [4] Meas, P.: How to do the right thing. *Connection Science* 1(3), 291–323 (1989)
- [5] Cho, S.-B., Kim, J.-H.: Combining multiple neural networks by fuzzy integral for robust classification. *IEEE Transactions on Systems, Man and Cybernetics* 25(2), 380–384 (1995)
- [6] Kim, K.-J., Cho, S.-B.: A unified architecture for agent behaviors with selection of evolved neural network modules. *Applied Intelligence* 25(3), 253–268 (2006)
- [7] Kim, K.-J., Cho, S.-B.: Fuzzy integration of structure adaptive SOM's for web content mining. *Fuzzy Sets and Systems* 148(1), 43–60 (2004)
- [8] Pearl, J.: *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Francisco (1997)
- [9] Rubin, A., Hammerman, J., Konold, C.: Exploring informal inference with interactive visualization software. In: *Proceedings of the 7th International Conference on Teaching Statistics* (2006)
- [10] Polikar, R.: Ensemble based systems in decision making. *IEEE Circuits and Systems Magazine* 6(3), 21–45 (2006)

- [11] Verikas, A., Lipnickas, A., Malmqvist, K., Bacauskiene, M., Gelzinis, A.: Soft combination of neural classifiers: A comparative study. *Pattern Recognition Letters* 20, 429–444 (1999)
- [12] Cho, S.-B.: Self-organizing map with dynamical node splitting: Application to handwritten digit recognition. *Neural Computation* 9(6), 1343–1353 (1997)
- [13] Kim, K.-J., Cho, S.-B.: Evolved neural networks based on cellular automata for sensory-motor controller. *Neurocomputing* 69(16-18), 2193–2207 (2006)