Original Article

Healthc Inform Res. 2012 June;18(2):105-114.
http://dx.doi.org/10.4258/hir.2012.18.2.105
pISSN 2093-3681 • eISSN 2093-369X

HIR
Healthcare Informatics Research

Design of Activation Functions for Inference of Fuzzy Cognitive Maps: Application to Clinical Decision Making in Diagnosis of Pulmonary Infection

In Keun Lee, PhD[1], Hwa Sun Kim, RN, PhD[2], Hune Cho, PhD[1]

[1]Department of Medical Informatics, Kyungpook National University School of Medicine, Daegu; [2]Department of Medical Information Technology, Daegu Haany University, Daegu, Korea

Objectives: Fuzzy cognitive maps (FCMs) representing causal knowledge of relationships between medical concepts have been used as prediction tools for clinical decision making. Activation functions used for inferences of FCMs are very important factors in helping physicians make correct decision. Therefore, in order to increase the visibility of inference results, we propose a method for designing certain types of activation functions by considering the characteristics of FCMs. **Methods:** The activation functions, such as the sinusoidal-type function and linear function, are designed by calculating the domain range of the functions to be reached during the inference process of FCMs. Moreover, the designed activation functions were applied to the decision making process with the inference of an FCM model representing the causal knowledge of pulmonary infections. **Results:** Even though sinusoidal-type functions oscillate and linear functions monotonously increase within the entire range of the domain, the designed activation functions make the inference stable because the proposed method notices where the function is used in the inference. And, the designed functions provide more visible numeric results than do other functions. **Conclusions:** Comparing inference results derived using activation functions designed with the proposed method and results derived using activation functions designed with the existing method, we confirmed that the proposed method could be more appropriately used for designing activation functions for the inference process of an FCM for clinical decision making.

Keywords: Decision Making, Computer Reasoning, Fuzzy Cognitive Maps, Activation Function

Submitted: May 9, 2012
Revised: June 21, 2012
Accepted: June 21, 2012

Corresponding Author
Hune Cho, PhD
Department of Medical Informatics, Kyungpook National University School of Medicine, 680, Gukchaebosang-ro, Jung-gu, Daegu 700-842, Korea. Tel: +82-53-420-4899, Fax: +82-53-423-1242, E-mail: hunecho@knu.ac.kr

I. Introduction

Various artificial intelligence techniques have been used for decision making in medical activities, such as diagnosis and therapy recommendations [1]. However, these techniques can be useful when the physician's knowledge is well represented in terms of computer realization and use. However, a large number of parameters such as medical symptoms and laboratory test results make it difficult for these techniques to be implemented using computers [2]. To overcome this problem, over the past few decades, various methods, e.g., neural networks [3,4], Bayesian networks [5], ontology [6,7], fuzzy cognitive maps (FCMs) [2,8,9], have been proposed

to represent a physician's knowledge and to support clinical decision making. In particular, FCMs [2,8-13] can efficiently handle complex modeling problems when assessing clinical decision making tasks [2]. FCMs represent causal knowledge between events and are used as tools that can predict results for current states of events by inference. FCMs have been used in not only clinical decision making but also system control, game theory, information analysis, etc. [10]. The advantages of FCMs are that FCMs can be represented in a matrix form and their inference process involves numerical matrix computations. During the inference process, the value of nodes in FCMs could be out of the range of [0, 1]; therefore, activation functions are used to keep the value of the nodes within the range. Therefore, the activation function used for the inference of FCMs is an important factor in determining the results of the inference.

Several research efforts have been conducted on activation functions for the inference process of FCMs [12-14] where a sigmoid function, hyperbolic tangent function, step function, and threshold linear function have been considered as activation functions, and [13] showed that the sigmoid function offers significantly greater advantages than the other functions. Moreover, Lee and Kwon [12] suggested a method for determining the λ value of a sigmoid function, as shown in Equation (1), to design an activation function that adapts to an FCM model.

$$f(x) = 1/\left(1 + e^{-\lambda x}\right). \tag{1}$$

During inference, the concept values of FCMs are restricted to $v \in (0, 1)$ by the sigmoid function. One of the characteristics of a sigmoid function is that its domain is $(-\infty, \infty)$, whereas the range of the function is (0, 1), i.e., we cannot obtain a "0" or "1" as a concept value after inference. Moreover, while a sigmoid function using $\lambda = 5$ [13] is known to be a good degree for normalization in [0, 1], the slope of a function of around $x = 0$ is greatly different from that of $x = 1$. Therefore, a sigmoid function is not suitable for use as a normalization function.

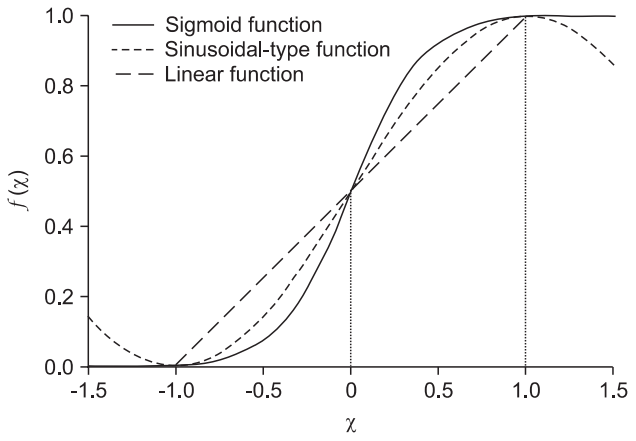

Figure 1. Graph of sigmoid function, where $\lambda = 5$; sinusoidal-type function, where $\beta = 1.5708$; and linear function, where $\alpha = 1$.

As shown in Figure 1, the sinusoidal-type function shown in Equation (2) at interval $[-\beta\pi/2, \beta\pi/2]$, and the linear function shown in Equation (3) are better normalization functions than a sigmoid function. Moreover, the range of a sinusoidal-type function is [0, 1], where the domain is restricted to the interval $[-\beta\pi/2, \beta\pi/2]$, and therefore, we can obtain "0" and "1" as the concept values after inference.

$$f(x) = 0.5\left(\sin(\beta x) + 1\right) \tag{2}$$

$$f(x) = 0.5\alpha x + 0.5 \tag{3}$$

Therefore, in this paper, we propose a method for designing activation functions for the inference of FCMs, which is different from the method suggested in [12]. Moreover, we apply the designed function to a clinical decision regarding the prediction of a pulmonary infection model [2].

II. Methods

1. Model Description and Preliminaries

For convenience, we will use the following notations and definitions throughout this paper.

Notations. $\mathbf{N}$, $\mathbf{R}$, $\mathbf{R}^n$, and $\mathbf{R}^{n\times m}$ denote a set of natural numbers, a real number space, a real n-space, and a set of real $n\times m$ matrices, respectively. The superscript "T" denotes a vector and matrix transposition (i.e., if $\mathbf{u}\in\mathbf{R}^n$, then $\mathbf{u}^{\mathrm{T}}=[u_i]_{1\le i\le n}$, and if $\mathbf{A}=[a_{ij}]_{n\times m}\in\mathbf{R}^{n\times m}$, then $\mathbf{A}^{\mathrm{T}} = [a_{ji}]_{m\times n}$, where $1\le i\le n$, $1\le j\le m$, and $n, m\in\mathbf{N}$). For all $\mathbf{u}\in\mathbf{R}^n$, let $\|\mathbf{u}\|$ denote the Euclidean vector norm (i.e., $\|\mathbf{u}\|=(\mathbf{u}^{\mathrm{T}}\cdot\mathbf{u})^{1/2}$). For all $\mathbf{A}\in\mathbf{R}^{n\times m}$, let $\|\mathbf{A}\|$ denote the spectral norm (i.e., $\|\mathbf{A}\|$ = (the maximum eigenvalue of $\mathbf{A}^{\mathrm{T}}\cdot\mathbf{A}$)$^{1/2}$). If $\mathbf{u}=[u_i]^{\mathrm{T}}_{1\le i\le n}$ is a state vector of a system, then $\mathbf{u}^*=[u_i^*]^{\mathrm{T}}_{1\le i\le n}$ denotes an equilibrium state vector of the system. If $f:\mathbf{R}\to\mathbf{R}$, then $f'(\cdot)$ and $f^{-1}(\cdot)$ are the first derivative and inverse function of $f(\cdot)$, respectively. If $u\in[u_1,u_2]$ for any $u, u_1, u_2\in\mathbf{R}$, then $I_{\max}(u)$ and $I_{\min}(u)$ stand for the maximum and minimum values of u, respectively, i.e., $I_{\max}(u)=u_2$ and $I_{\min}(u)=u_1$.

The following descriptions show mathematical models representing the characteristics and inference process of an FCM as defined in [9,12,15].

Definition 1 (Components of FCM). (refer to [15]) Suppose C_i and C_j are concepts in an FCM, and v_i and v_j are

http://dx.doi.org/10.4258/hir.2012.18.2.105

the values of C_i and C_j belonging to [0, 1], respectively, when $i, j \in N = \{1,2,\dots, n\}$ and $n \in \mathbf{N}$ is the number of concepts. Then, weight w_{ij} is defined as a real number in [−1, 1]. We deem the weight as positive, negative, or having no causality from C_i to C_j when $w_{ij} > 0$, $w_{ij} < 0$, and $w_{ij} = 0$, respectively.

Definition 2 (Inference process of FCM). (refer to [15]) For every $i \in N$ and any $j \in N$, let C_i be the causal concepts that influence concept C_j. Then, for every $j \in N$ and all iteration steps of $k \geq 0$ during inference process of the FCM,

$$v_j^{(k+1)} = f\left(\rho_1 v_j^{(k)} + \rho_2 \sum_{l=1}^{n} \left(w_{lj} v_l^{(k)}\right)\right), \tag{4}$$

where $\rho_1, \rho_2 \in (0, 1]$ and $v_j^{(k)} \in [0, 1]$ represents the value of C_j at the k-th iteration step. Moreover, $f: \mathbf{R} \to \mathbf{R}$ is an activation function to restrict $v_j^{(k+1)}$ into the interval [0, 1]. Equation (4) is also represented in vector form as

$$\mathbf{v}^{(k+1)} = f\left(\rho_1 \mathbf{v}^{(k)} + \rho_2 \mathbf{w}^{\mathrm{T}} \cdot \mathbf{v}^{(k)}\right), \tag{5}$$

where $\mathbf{v}^{(k)} = [v_1^{(k)}\ v_2^{(k)} \cdots\ v_n^{(k)}]^{\mathrm{T}}$ and $\mathbf{w} = [w_{ij}]_{n \times n}$, where $1 \leq i, j \leq n$, and are called a state vector and weight matrix, respectively. Moreover, $f(\mathbf{v}^{(k)}) = \left[f(v_1^{(k)}) \ \cdots\ f(v_n^{(k)})\right]^{\mathrm{T}}$.

As in [12], we transform the model from Equations. (4) and (5) into the form described in the following definition.

Definition 3 (Transformation). (refer to [12]) For every $j \in N$ and all $k \geq 0$, let $x_j^{(k)} = \rho_1 v_j^{(k)} + \rho_2 \sum_{l=1}^{n} \left(w_{lj} v_l^{(k)}\right)$ and $v_j^{(k+1)} = f(x_j^{(k)})$ in Equation (4); then,

$$\begin{aligned} \mathbf{x}^{(k+1)} &= \rho_1 \mathbf{v}^{(k+1)} + \rho_2 \mathbf{w}^{\mathrm{T}} \cdot \mathbf{v}^{(k+1)} \\ &= \rho_1 f(\mathbf{x}^{(k)}) + \rho_2 \mathbf{w}^{\mathrm{T}} \cdot f(\mathbf{x}^{(k)}), \end{aligned} \tag{6}$$

where $\mathbf{x}^{(k)} = [x_1^{(k)}\ x_2^{(k)} \cdots\ x_n^{(k)}]^{\mathrm{T}}$.

We also consider unipolar, nonlinear, and continuous functions as activation functions of FCMs. Therefore, we assume that the activation functions satisfy the following conditions:

Assumption 1. The function $f: \mathbf{R} \to \mathbf{R}$ is bounded; i.e., $0 \leq f(u) \leq M$ for all $u \in \mathbf{R}$ and any $M \in \mathbf{R}$ such that $M > 0$.

Assumption 2. The function $f: \mathbf{R} \to \mathbf{R}$ satisfies the Lipschitz condition with a Lipschitz constant, $L > 0$; i.e., $|f(u_1) - f(u_2)| \leq L|u_1 - u_2|$ for all $u_1, u_2 \in \mathbf{R}$.

Lee and Kwon [12] suggested a bound of L of activation functions satisfying Assumptions 1 and 2, as shown in Equation (7), which guarantees the global exponential stability of Equation (5) during an inference process of an FCM.

$$0 < L < \frac{1}{\rho_1 + \rho_2 \|\mathbf{w}\|} \tag{7}$$

Moreover, as shown in Equation (8), the bound of λ was derived using Equation (7) and the property in which the maximum value of the derivative of a sigmoid function occurs when $x = 0$.

$$0 < \lambda < \frac{4}{\rho_1 + \rho_2 \|\mathbf{w}\|} \tag{8}$$

Consequently, the λ value are determined by adapting the weight matrix $\mathbf{w}$ as an FCM model and the sigmoid function whose λ value satisfies inequality (8) guarantees the stability of Equation (5).

2. Design of Sinusoidal-Type Activation Functions

As mentioned in the introduction, a sinusoidal-type function may not be appropriate to an activation function because it is oscillated in the bound of domain $(-\infty, \infty)$. From another viewpoint, the sinusoidal-type function, shown in Equation (2), within the bound of domain $[-\beta\pi/2, \beta\pi/2]$, could be better than a sigmoid function as a normalization function, because a sinusoidal-type function is a monotonous increase function that has a gentler slope than a sigmoid function does. Also, the range of a sinusoidal-type function is [0, 1], which is different than a sigmoid function whose range is (0, 1). Therefore, we need to find the value of β. Intuitively, $\beta\pi/2$ and $-\beta\pi/2$ may be the maximum and minimum values that x can reach during inference, respectively. Consequently, finding the maximum and minimum values of x is a way to find the value of β. Since a sinusoidal-type function will be used as an activation function for Equation (5) in this paper, the domain values will be the elements of vector $\mathbf{v}^{(k)}$ in Definition 3. To find the bound of $\mathbf{x}^{(k)}$ as the result of inference of an FCM using Equation (5), we give following lemmas.

Lemma 1. Let $\mathbf{x}^{(k)}$ and M be the same as in Definitions 3 and Assumption 1, respectively. Then, for all $k \geq 0$, the following inequality is satisfied.

$$\|\mathbf{x}^{(k)}\| \leq \left(\rho_1 + \rho_2 \|\mathbf{w}\|\right) n^{1/2} M \tag{9}$$

Proof. Let the right term of inequality (6) be $\varphi(x)$ as follows:

$$\varphi(\mathbf{x}) = \rho_1 f\left(\mathbf{x}^{(k)}\right) + \rho_2 \mathbf{w}^{\mathrm{T}} \cdot f\left(\mathbf{x}^{(k)}\right). \tag{10}$$

If we use the Euclidean norm for both terms of Equation (10), we can then derive the following inequality.

$$\begin{aligned}\|\varphi(\mathbf{x})\| &= \|\rho_1 f(\mathbf{x}^{(k)}) + \rho_2 \mathbf{w}^{\mathrm{T}} \cdot f(\mathbf{x}^{(k)})\| \\ &\le \rho_1 \|f(\mathbf{x}^{(k)})\| + \rho_2 \|\mathbf{w}\| \|f(\mathbf{x}^{(k)})\| \\ &= (\rho_1 + \rho_2 \|\mathbf{w}\|) \|f(\mathbf{x}^{(k)})\| \\ &= (\rho_1 + \rho_2 \|\mathbf{w}\|) \left(\sum_{l=1}^{n} \left(f(x_l^{(k)})\right)^2\right)^{1/2} \\ &\le (\rho_1 + \rho_2 \|\mathbf{w}\|) \left(\sum_{l=1}^{n} M^2\right)^{1/2} \\ &= (\rho_1 + \rho_2 \|\mathbf{w}\|) n^{1/2} M\end{aligned}$$

Therefore, we have

$$\|\mathbf{x}^{(k+1)}\| \le (\rho_1 + \rho_2 \|\mathbf{w}\|) n^{1/2} M . \quad (11)$$

If we suppose there exists $\mathbf{x}^{(-1)}$ such that $\mathbf{v}^{(0)} = f(\mathbf{x}^{(-1)})$, then inequality (11) is represented as Equation (9) because of $0 \le f(x_j^{(-1)}) \le M$ from Assumption 1. □

Lemma 2. Let $\mathbf{x}^{(k)} = \left[x_j^{(k)}\right]_{1 \le j \le n}^{\mathrm{T}}$ be the same as in Definition 3. Then, for all $1 \le j \le n$, $\left|x_j^{(k)}\right| \le (\rho_1 + \rho_2 \|\mathbf{w}\|) n^{1/2} M$.

Proof. Consider a vector $\mathbf{x}^{(k)} = \left[0, \ldots, x_j^{(k)}, \ldots, 0\right]^{\mathrm{T}}$. Then, $\left|x_j^{(k)}\right| = \left\|\mathbf{x}^{(k)}\right\|$. Here, $\left|x_j^{(k)}\right|$ is the maximum absolute value among other elements within the unit circle in the vector norm. Thus, from inequality (11), we have $\left|x_j^{(k)}\right| \le (\rho_1 + \rho_2 \|\mathbf{w}\|) n^{1/2} M$. □

We can know the domain range of a sinusoidal-type activation function used in the inference process of FCMs through Lemmas 1 and 2. Therefore, we give following theorem for the design of the sinusoidal-type activation function.

Theorem 1. Let $\mathbf{x}^{(k)} = \left[x_j^{(k)}\right]_{1 \le j \le n}^{\mathrm{T}}$ and M be the same as in Definition 3 and Assumption 1. If there exists an inverse function, $f^{-1}(\cdot)$, of an activation function, $f(\cdot)$, the following equation is satisfied for all $k \ge 0$ and any $j \in N$.

$$\left|f^{-1}\left(v_j^{(k+1)}\right)\right| \le (\rho_1 + \rho_2 \|\mathbf{w}\|) n^{1/2} M \quad (12)$$

Proof. If $x_j^{(k)}$ is the j-th element of $\mathbf{x}^{(k)}$, by Lemma 2 we know the range of $x_j^{(k)}$ to be $\left|x_j^{(k)}\right| \le (\rho_1 + \rho_2 \|\mathbf{w}\|) n^{1/2} M$. Also, $v_j^{(k+1)} = f(x_j^{(k)})$ can be represented as $f^{-1}\left(v_j^{(k+1)}\right) = x_j^{(k)}$ because $f(\cdot)$ is invertible. Therefore, we finally have $\left|f^{-1}\left(v_j^{(k+1)}\right)\right| \le (\rho_1 + \rho_2 \|\mathbf{w}\|) n^{1/2} M$. □

Note 1. If we know the range of $x_j^{(k)}$ and $v_j^{(k+1)}$ for all $k \ge 0$ and $j \in N$, we can design activation functions satisfying Assumptions 1 and 2 by assigning $I_{\max}\left(x_j^{(k)}\right)$ and $I_{\min}\left(x_j^{(k)}\right)$ to $I_{\max}\left(v_j^{(k+1)}\right)$ and $I_{\min}\left(v_j^{(k+1)}\right)$, respectively.

The following corollary shows how to actually design a sinusoidal-type activation function using Theorem 1.

Corollary 1. Let β be the same as in Equation (2) and ρ_1, ρ_2, $\mathbf{w}$, n, and M be the same as in Theorem 1. Then, in the designed sinusoidal-type activation function, the value of β is

$$\beta = \frac{1.5708}{(\rho_1 + \rho_2 \|\mathbf{w}\|) n^{1/2} M} . \quad (13)$$

Proof. We can easily derive the following equation, which is the inverse function of Equation (2).

$$f^{-1}\left(v_j^{(k+1)}\right) = x_j^{(k)} = \frac{1}{\beta} \sin^{-1}\left(2 v_j^{(k+1)} - 1\right)$$

Thus, we have

$$\beta = \frac{1}{x_j^{(k)}} \sin^{-1}\left(2 v_j^{(k+1)} - 1\right).$$

If $I_{\max}\left(x_j^{(k)}\right)$ and $I_{\min}\left(x_j^{(k)}\right)$ are assigned to $I_{\max}\left(v_j^{(k+1)}\right)=1$ and $I_{\min}\left(v_j^{(k+1)}\right)=0$, respectively, then β is computed as

$$\begin{aligned}\beta &= \frac{1}{I_{\max}\left(x_j^{(k)}\right)} \sin^{-1}(2-1) = \frac{1}{I_{\min}\left(x_j^{(k)}\right)} \sin^{-1}(0-1) \\ &\approx \frac{1.5708}{(\rho_1 + \rho_2 \|\mathbf{w}\|) n^{1/2} M} .\end{aligned}$$ □

The following corollary shows that the designed sinusoidal-type activation function guarantees the global exponential stability of the inference process of an FCM.

Lemma 3. (refer to [12]) If $x_j^{(k)}$ is the same as in Definition 3, and L is a Lipschitz constant as shown in Assumption 2, then for all $j \in N$, $\left|f'(x_j^{(k)})\right| \le L$.

Corollary 2. The inference process of an FCM using Equation (5) is globally exponentially stable when the activation function is a sinusoidal-type, as in Equation (2), where $\beta = 1.5708 / (\rho_1 + \rho_2 \|\mathbf{w}\|) n^{1/2}$.

Proof. The sinusoidal-type function of Equation (2) satisfies Assumption 1, and the maximum value of the first derivative of the function occurs when $x = 0$. Therefore, the range of β is calculated using inequality (7) and Lemma 3 as follows:

$$\begin{aligned}f'(x) &= \frac{1}{2}\beta \cos(\beta x) \\ \left|f'(0)\right| &= \frac{1}{2}\beta \cos(0) = \frac{1}{2}\beta \le L < \frac{1}{\rho_1 + \rho_2 \|\mathbf{w}\|} \\ 0 &< \beta < \frac{2}{\rho_1 + \rho_2 \|\mathbf{w}\|} .\end{aligned} \quad (14)$$

http://dx.doi.org/10.4258/hir.2012.18.2.105

This guarantees the global exponential stability of Equation (5). If $M = 1$ in inequality (13), we finally have

$$0 < \frac{1.5708}{(\rho_1+\rho_2\|\mathbf{w}\|)n^{1/2}} < \frac{2}{\rho_1+\rho_2\|\mathbf{w}\|}.$$

This inequality is satisfied for all $n\in\mathbf{N}$. Therefore, Equation (2), where $\beta=1.5708/(\rho_1+\rho_2\|\mathbf{w}\|)n^{1/2}$, also guarantees the global exponential stability of Equation (5). □

We next give an example that confirms the stability of the inference process using the designed sinusoidal-type activation function.

Example 1. Let $\rho_1=\rho_2=1$ and $M=1$ in Definition 2 and Assumption 1. Also, suppose that weight matrix $\mathbf{w}$ and initial state vector $\mathbf{v}^{(0)}$ are

$$\mathbf{w}=\begin{bmatrix}1.0 & -0.8\\ 0.95 & 0.7\end{bmatrix},\quad \mathbf{v}^{(0)}=\begin{bmatrix}0.5\\ 0.7\end{bmatrix}.$$

The following three kinds of sinusoidal-type activation functions with different values of β are considered in this example:

(*i*) $\beta_1 = 0.4652$, calculated by the proposed method, where $n = 2$:

$$\beta_1 \approx \frac{1.5708}{(\rho_1+\rho_2\|\mathbf{w}\|)n^{1/2}M} \approx \frac{1.5708}{3.3766} \approx 0.4652.$$

(*ii*) $\beta_2 = 0.8376$, which is within the range of β calculated by inequality (14), and guarantees the global exponential stability of the inference process,

$$0<\beta<0.83766392031338.$$

(*iii*) $\beta_3 = 1.0870$, which is out of the range of β, and does not guarantee the global exponential stability of the inference process.

Using the designed sinusoidal-type activation functions, the inferences in Equation (5) based on $\mathbf{w}$ and $\mathbf{v}^{(0)}$ are performed. After inference, the following results are obtained: vectors saturated to (*i*) $\mathbf{v}_1^*$ = [0.9501 0.5346] and (*ii*) $\mathbf{v}_2^*$ = [0.9369 0.6396], and vectors oscillated between (*iii*) $\mathbf{v}_3^{(k)}$ = [0.7706 0.9451] and $\mathbf{v}_3^{(k+1)}$ = [0.7355 0.9401].

Figure 2 shows the designed activation functions and trajectories of the concept values during inference. The results in (*i*) and (*ii*) are stable, but some oscillation is observed at the start of the trajectory in the result in (*ii*), as shown in Figure 2D. Otherwise, the result in (*iii*) is not stable. Comparing the result in (*i*) with those in (*ii*) and (*iii*), it is reasonable to conclude that this oscillation in the trajectories is caused by the activation functions, which are $I_{\max}\left(f\left(x_j^{(k)}\right)\right) < \left|x_j^{(k)}\right|$, where $j\in\{1, 2\}$.

3. Design of Linear Activation Functions

A linear function, as shown in Equation (3), is not appropriate for an activation function for inference of FCMs, because it monotonously increases for the domain and range of $(-\infty, \infty)$. However, if we know the domain range of the linear function to be reached during inference process, we can design the linear function as an activation function. That is, the following corollary shows a way to design a linear-type activation function that satisfies the following condition.

Assumption 3. The function $f:\mathbf{R}\rightarrow\mathbf{R}$ is bounded; i.e., $0\le f(u)\le M$ for all $u, u_1, u_2\in\mathbf{R}$ such that $u\in[u_1, u_2]$ and $u_1<u_2$, and any $M\in\mathbf{R}$ such that $M>0$.

Corollary 3. Let α be the same as in Equation (3) and ρ_1, ρ_2, $\mathbf{w}$, n, and M be the same as in Theorem 1. Then, in the designed linear-type activation function, the value of α is

$$\alpha = \frac{0.5}{(\rho_1+\rho_2\|\mathbf{w}\|)n^{1/2}M}.$$

Proof. We derive the following equation, which is the inverse function of Equation (3).

$$f^{-1}\left(v_j^{(k+1)}\right) = x_j^{(k)} = \frac{1}{\alpha}\left(v_j^{(k+1)}-0.5\right)$$

Thus, we have

$$\alpha = \frac{1}{x_j^{(k)}}\left(v_j^{(k+1)}-0.5\right).$$

If $I_{\max}\left(x_j^{(k)}\right)$ and $I_{\min}\left(x_j^{(k)}\right)$ are assigned to $I_{\max}\left(v_j^{(k+1)}\right)=1$ and $I_{\min}\left(v_j^{(k+1)}\right)=0$, respectively, then α is computed as

$$\alpha = \frac{1}{x_j^{(k)}}\left(v_j^{(k+1)}-0.5\right) = \frac{0.5}{I_{\max}\left(x_j^{(k)}\right)} = \frac{-0.5}{I_{\min}\left(x_j^{(k)}\right)} = \frac{0.5}{(\rho_1+\rho_2\|\mathbf{w}\|)n^{1/2}M}.$$ □

4. Design of FCM on Pulmonary Infection

To apply he designed sinusoidal-type and linear activation functions to an FCM model for a clinical decision, we refer the FCM model designed in [2] representing causal knowledge of pulmonary infections. However, the characteristics of the FCM in [2] are different with those of the FCM con-

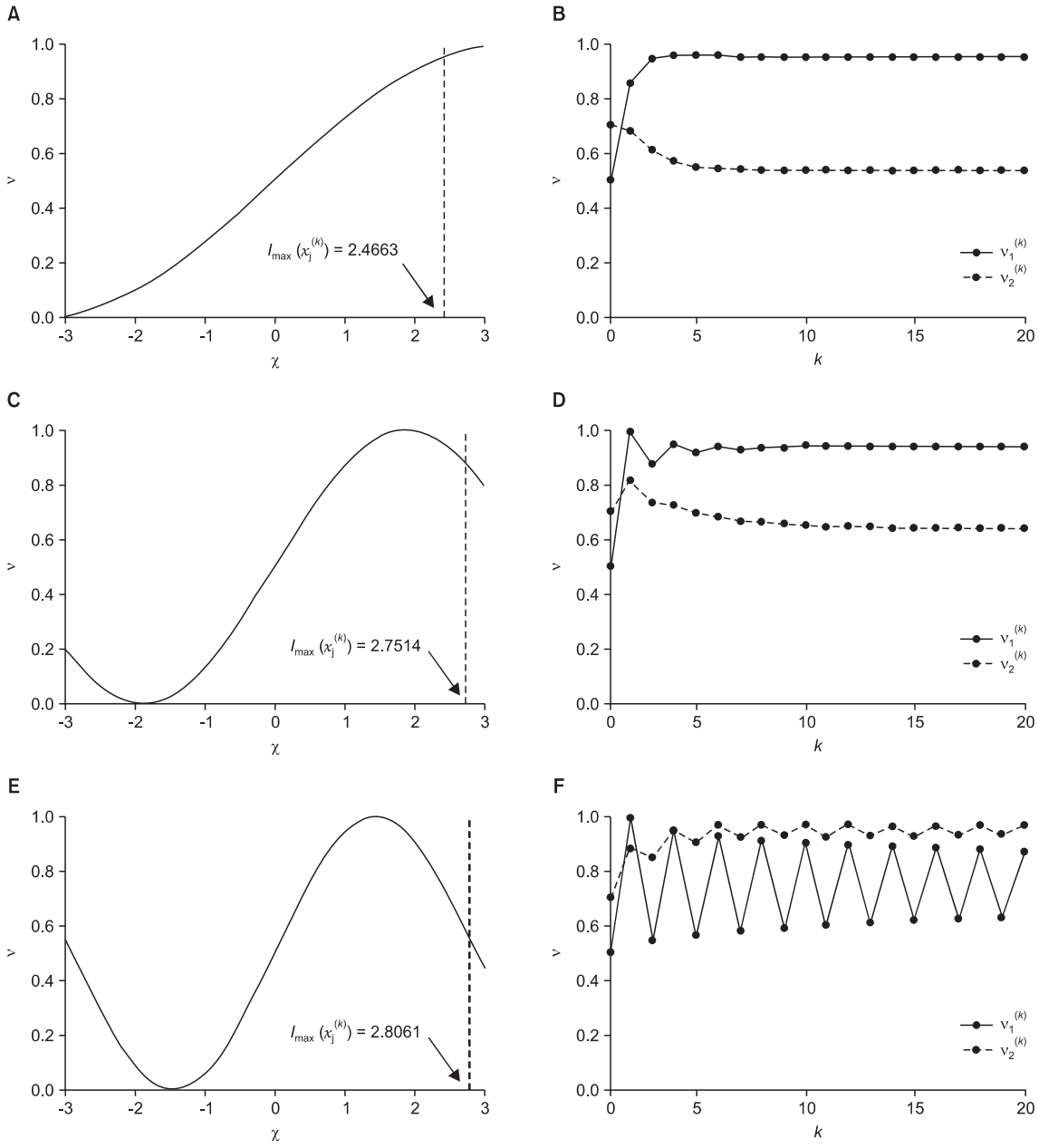

Figure 2. Results of inference process of an fuzzy cognitive map using various activation functions: (A) sinusoidal function where β = 0.4652; (B) trajectory of state vector in (*i*); (C) sinusoidal function where β = 0.8376; (D) trajectory of state vector in (*ii*); (E) sinusoidal function where β = 1.0870; and (F) trajectory of state vector in (*iii*).

sidered in this paper. That is, different from Definition 1, the concept values in [2] were bounded into the interval [-1, 1], and a bipolar activation function was used for the inference process. Therefore, we customize the FCM model designed in [2], as shown Figure 3, by adding seven concepts (C26-1, C27-1, C28-1, C29-1, C30-1, C31-1, and C32-1) representing the negative values of the concepts (C26, C27, C28, C29, C30, C31, and C32, respectively), as was done in [11]. For instance, if the concept value of C26 is "$v_{C26} = -1$," then D1 is affected by the amount from $v_{C26} \times w_{C26,D1} = -1 \times 0.7$. However, according to Definition 1, we cannot represent the concept value of "-1." To affect negative influence on D1, we create a new concept, C26-1, which applies a negative value to concept C26, i.e., if "sputum culture" is "-1" then "$v_{C26-1} = 1$" because C26-1 involves the meaning of the negative concept value. Moreover, we give the weight "$w_{C26-1,D1} = -0.7$" between concepts C26-1 and D1, i.e., even though the value of C26-1 is positive, D1 is affected negatively by $v_{C26-1} \times w_{C26-1,D1} = 1 \times (-0.7)$.

We deal with the two scenarios described in [2]. That is, this experiment aims to show the process for physician's decision that which patient is more serious in pulmonary

http://dx.doi.org/10.4258/hir.2012.18.2.105

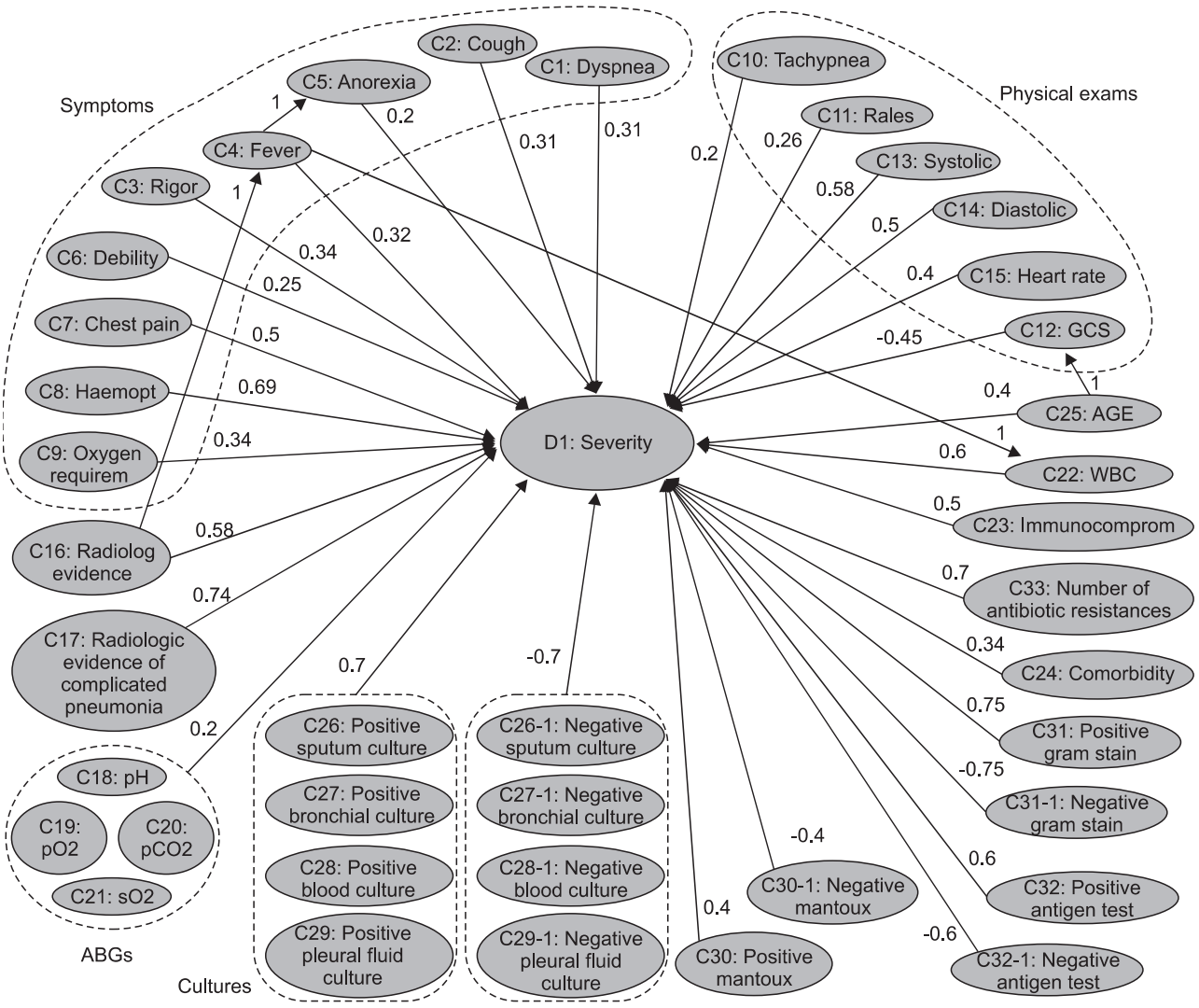

Figure 3. Customized fuzzy cognitive map model for predicting severity index of pulmonary infection. ABGs: arterial blood gases, WBC: white blood cell, GCS: glasgow coma scale.

infection, based on observed symptoms and laboratory test results of patients in the following scenarios.

Scenario 1: An immunocompromised patient (A23 = 1) with a high fever (A4 = 0.7), loss of appetite (A5 = 1), and high systolic blood pressure (A13 = 0.7), with radiologic evidence present in his/her chest x-rays (A16 = 1), a small number of WBCs (A22 = 0.4), a negative sputum culture (A26-1 = 1), and negative antigen (A32-1 = 1).

Scenario 2: An older patient (A25 = 0.8) with a low fever (A4 = 0.3), altered mental status (A12 = 0.4), high oxygen requirements (A9 = 0.8), a normal number of leukocytes-white blood cell (A22 = 0), positive sputum culture (A26 = 1), negative blood culture (A28-1 = 1), and negative gram stain (A31-1 = 1).

III. Results

According to [2], physicians made decision that the patient in scenario 1 is more serious in pulmonary infection than that in scenario 2, and the inference of FCM also shows the same result with the physicians' decision. In this experiment, therefore, we compared the results of the inference process of the FCMs with four activation functions, (*i*) the sigmoid function designed in [12], (*ii*) a sinusoidal-type function designed using the method proposed in [12], (*iii*) a sinusoidal-type function and (*iv*) a linear function designed using the proposed method. In these inferences, we provided the same stimulus as in [16].

According to the two scenarios, we created the following initial state vectors:

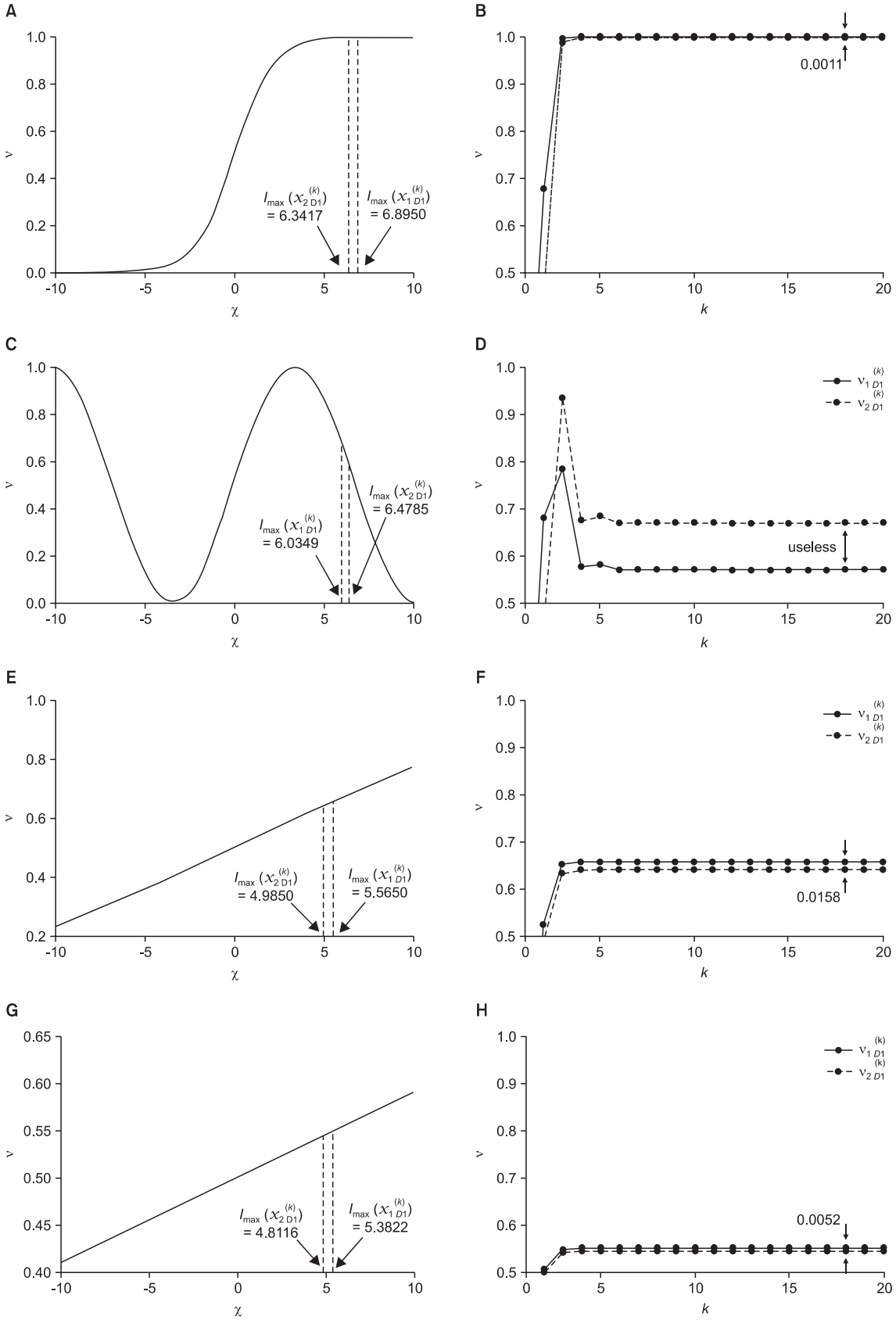

Figure 4. Designed activation functions and trajectories of values of concept D1 regarding two scenarios during inference: (A) sigmoid function where λ = 0.9270; (B) trajectory of values in (*i*); (C) sinusoidal function where β = 0.4635; (D) trajectories of values in (*ii*); (E) sinusoidal-type function where β = 0.0569; and (F) trajectory of values in (*iii*); (G) linear function where α = 0.0181; and (H) trajectory of values in (*iv*).

http://dx.doi.org/10.4258/hir.2012.18.2.105

$\mathbf{v}_1^{(0)}$ = [0 0 0 0.7 1 0 0 0 0 0 0 0 0 0 0.7 0 1 0 0 0 0 0 0 0.4 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0],

$\mathbf{v}_2^{(0)}$ = [0 0 0 0.3 0 0 0 0 0.8 0 0 0.4 0 0 0 0 0 0 0 0 0 0 0 0 0 0.8 1 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0].

Figure 4 shows the designed activation function and trajectories of the values of concept "D1: Severity" at each inference of the two scenarios.

For (*i*), the values of concept D1 regarding scenarios 1 and 2 were converged to $v_{1\,D1}^{*} = 0.9983$ and $v_{2\,D1}^{*} = 0.9972$, respectively. As shown in Figure 4A, the slope is almost flat around the maximum values of $x_{1\,D1}^{(k)}$ and $x_{2\,D1}^{(k)}$, and thus the gap between the converged concept values is very small, at $\left|v_{1\,D1}^{*} - v_{2\,D1}^{*}\right| = 0.0011$. Even though it was difficult to make a decision based on the results in (*i*), the results showed that the patient in scenario 1 has a severer condition than the patient in scenario 2, which is the same result determined in [2].

For (*ii*), the values of concept D1 regarding scenarios 1 and 2 were converged to $v_{1\,D1}^{*} = 0.5691$ and $v_{2\,D1}^{*} = 0.6688$, respectively. However, in the designed activation function, as shown in Figure 4C, the maximum values of $x_{1\,D1}^{(k)}$ and $x_{2\,D1}^{(k)}$ exceeded the value that yielded $I_{\max}(v)$. As a result, the patient in scenario 2 has a severer condition than the patient in scenario 1, which is contrary to (*i*). That is, the result is not useful for decision making.

For (*iii*), the values of concept D1 regarding scenarios 1 and 2 were converged to $v_{1\,D1}^{*} = 0.6556$ and $v_{2\,D1}^{*} = 0.6398$, respectively. Differing with cases (*i*) and (*ii*), the designed activation function looks almost like a linear function around the maximum values of $x_{1\,D1}^{(k)}$ and $x_{2\,D1}^{(k)}$. Moreover, the gap between the converged concept values is the largest among the results, at $\left|v_{1\,D1}^{*} - v_{2\,D1}^{*}\right| = 0.0158$. That is, comparing (*i*) and (*iii*), we can see that the results of the latter make it more convenient for a physician to make a decision and with more correct results.

For (*iv*), the values of concept D1 regarding scenarios 1 and 2 were converged to $v_{1\,D1}^{*} = 0.5487$ and $v_{2\,D1}^{*} = 0.5435$, respectively. The gap between the converged concept values is the better than the result in (*i*), at $\left|v_{1\,D1}^{*} - v_{2\,D1}^{*}\right| = 0.0052$.

As a result, we can see that although the sinusoidal-type activation functions was designed using the method proposed in [12], it occasionally provides incorrect results for decision making. Therefore, we can determine that the method proposed in this paper is more appropriate for designing sinusoidal-type activation functions than the method proposed in [12].

IV. Discussion

There exist various methods to support physicians' clinical decision making such as fuzzy, neural networks, decision tree, and FCMs. However, the usability of the methods is strongly dependent on the features of a clinical field; because their knowledge models are different from each other and each method has its own strength and weakness. And even in the same method, the results of decision making may be different according to its knowledge model. Thus, the users in clinical field (e.g., physicians) only refer to the results from the methods when they make clinical decisions. The aim of the methods in clinical field is how clearly shows the results to the users.

In this paper, we focused on a clinical decision making based on FCMs, which are good models of the causal knowledge of relationships between medical concepts and provide prediction results based on the current status of a concept through an inference process. Therefore, activation functions used for the inference process are very important factors that support physicians in making the right decision. In other words, for physicians to make a final decision, how well the physicians' knowledge is represented as an FCM model is not the only important factor. The inference process of that model is also important in the application of clinical decision making. In general, sigmoid functions have been used as activation functions for the inference process of FCMs; the design of an activation function is greatly dependent on the experience of experts because, during inference, the slope varies considerably within the domain range of the function.

Therefore, we proposed a method for designing sinusoidal-type and linear activation functions by calculating the domain range of the activation function to be reached during the inference process of FCMs. Even though sinusoidal-type functions are oscillated and linear functions are monotonously increased within the entire range of the domain, the designed activation functions make the inference stable because the proposed method notices where the function is used in the inference. Moreover, because a sinusoidal-type function designed by the proposed method provides a gentler slope than a sigmoid function does, it can be used as a normalization function. We applied the designed functions to an FCM model that represents the causal knowledge of pulmonary infections. Comparing the activation function designed using the proposed method with activation functions designed using an existing method, we confirmed that the proposed method can be appropriately used for designing the activation functions for the inference process of an FCM for clinical decision making.

This study dealt with only two kinds of functions and limited their adaption into an example of decision making with

the designed knowledge model in other study. In future research, we will consider another type of functions such as hyperbolic tangent function and apply the functions to more various FCM models in medical field.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2011-0026260).

References

1. Clancey WJ, Shortliffe EH. Readings in medical artificial intelligence: the first decade. Reading, MA: Addison-Wesley; 1984.
2. Papageorgiou EI, Papandrianos NI, Karagianni G, Kyriazopoulos GC, Sfyras D. A fuzzy cognitive map based tool for prediction of infectious diseases. In: Proceedings of the IEEE International Conference on Fuzzy Systems, 2009. p.2094-9.
3. Assadi SM, Yucel M, Pantelis C. Dopamine modulates neural networks involved in effort-based decision-making. Neurosci Biobehav Rev 2009;33:383-93.
4. Mangalampalli A, Mangalampalli SM, Chakravarthy R, Jain AK. A neural network based clinical decision-support system for efficient diagnosis and fuzzy-based prescription of gynecological diseases using homoeopathic medicinal system. Expert Syst Appl 2006;30:109-16.
5. Cruz-Ramirez N, Acosta-Mesa HG, Carrillo-Calvet H, Nava-Fernandez LA, Barrientos-Martínez RE. Diagnosis of breast cancer using Bayesian networks: a case study. Comput Biol Med 2007;37:1553-64.
6. Farion K, Michalowski W, Wilk S, O'Sullivan D, Rubin S, Weiss D. Clinical decision support system for point of care use: ontology-driven design and software implementation. Methods Inf Med 2009;48:381-90.
7. Kawazoe Y, Ohe K. An ontology-based mediator of clinical information for decision support systems: a prototype of a clinical alert system for prescription. Methods Inf Med 2008;47:549-59.
8. Froelich W, Wakulicz-Deja A. Associational cognitive maps for medical diagnosis support. In: Proceedings of the International Intelligent Information Systems Conference, 2008. p.387-96.
9. Stylios CD, Groumpos PP. Fuzzy cognitive maps in modeling supervisory control systems. J Intell Fuzzy Syst 2000;8:83-98.
10. Aguilar J. A survey about fuzzy cognitive maps papers (invited paper). Int J Comput Cogn 2005;3:27-33.
11. Kosko B. Fuzzy cognitive maps. Int J Man-Machine Stud 1986;24:65-75.
12. Lee IK, Kwon SH. Design of sigmoid activation functions for fuzzy cognitive maps via Lyapunov stability analysis. IEICE Trans Inf Syst 2010;93:2883-6.
13. Bueno S, Salmeron JL. Benchmarking main activation functions in fuzzy cognitive maps. Expert Syst Appl 2009;36:5221-9.
14. Tsadiras AK. Comparing the inference capabilities of binary, trivalent and sigmoid fuzzy cognitive maps. Inf Sci 2008;178:3880-94.
15. Stylios CD, Groumpos PP. Mathematical formulation of fuzzy cognitive maps. In: Proceedings of the 7th Mediterranean Conference on Control and Automation, 1999. p.2251-61.
16. Dickerson JA, Kosko B. Virtual worlds as fuzzy cognitive maps. In: Proceedings of the IEEE Virtual Reality Annul International Symposium, 1993. p.471-7.

http://dx.doi.org/10.4258/hir.2012.18.2.105