Modeling with Words Based on Hedge Algebra

Nguyen Van Han[1(✉)] and Phan Cong Vinh[2]

[1] Faculty of Information Technology, College of Science, Hue University, 77 Nguyen Hue street, Phu Nhuan ward, Hue city, Vietnam
nvhan@fit-hitu.edu.vn

[2] Faculty of Information Technology, Nguyen Tat Thanh University, 300A Nguyen Tat Thanh street, Ward 13 District 4, Ho Chi Minh city, Vietnam
pcvinh@ntt.edu.vn

Abstract. In this paper, we introduce a method for modeling with words based on hedge algebra using fuzzy cognitive map. Our model, called linguistic cognitive map, consists of set of vertices and edges with value to be linguistic variables. We figure out relationship between the length of linguistic variables for fuzzifying data and a number of partition from unit interval. We also prove finite properties of state space, generating from linguistic cognitive map.

Keywords: Fuzzy logics · Linguistic variable · Hedge algebra
Fuzzy cognitive map

1 Introduction

In everyday life, people use natural language (NL) for analysing, reasoning, and finally, make their decisions. Computing with words (CWW) [5] is a mathematical solution of computational problems stated in an NL. CWW based on fuzzy set and fuzzy logic, introduced by L.A. Zadeh is an approximate method on interval [0,1]. In linguistic domain, linguistic hedges play an important role for generating set of linguistic variables. A well known application of fuzzy logic (FL) is fuzzy cognitive map ($\mathbb{FCM}$), introduced by Kosko [1], combined fuzzy logic with neural network. $\mathbb{FCM}$ has a lots of applications in both modeling and reasoning fuzzy knowledge [3,4] on interval [0,1] but not in linguistic values, However, many applications cannot model in numerical domain [5], for example, linguistic summarization problems [6]. To solve this problem, in the paper, we use an abstract algebra, called hedge algebra ($\mathbb{HA}$) as a tool for computing with words.

The remainder of paper is organized as follows. Section 2 reviews some main concepts of computing with words based on $\mathbb{HA}$ in Subsect. 2.1 and describes several primary concepts for $\mathbb{FCM}$ in Subsect. 2.2. In Sect. 3, we introduce an approach technique to modeling with words using $\mathbb{HA}$. Sect. 4 outlines discussion and future work. Section 5 concludes the paper.

P. Cong Vinh and V. Alagar (Eds.): ICCASA 2018/ICTCC 2018, LNICST 266, pp. 211–217, 2019.
https://doi.org/10.1007/978-3-030-06152-4_18

2 Preliminaries

This section presents basic concepts of $\mathbb{HA}$ and $\mathbb{FCM}$ used in the paper.

2.1 Hedge Algebra

In this section, we review some $\mathbb{HA}$ knowledges related to our research paper and give basic definitions. First definition of a $\mathbb{HA}$ is specified by 3-Tuple $\mathbb{HA} = (X, H, \leq)$ in [7]. In [8] to easily simulate fuzzy knowledge, two terms G and C are inserted to 3-Tuple so $\mathbb{HA} = (X, G, C, H, \leq)$ where $H \neq \emptyset$, $G = \{c^+,\ c^-\}$, $C = \{0, W, 1\}$. Domain of X is $\mathbb{L} = Dom(X) = \{\delta c |\ c \in G,\ \delta \in H^*(\text{hedge string over H})\}$, $\{\mathbb{L}, \leq\}$ is a POSET (partial order set) and $x = h_n h_{n-1} \ldots h_1 c$ is said to be a canonical string of linguistic variable x.

Example 1. Fuzzy subset *X is Age*, $G = \{c^+ = young;\ c^- = old\}$, $H = \{less; more; very\}$ so term-set of linguistic variable Age X is $\mathbb{L}$(X) or $\mathbb{L}$ for short: $\mathbb{L}$ = {*very less young ; less young ; young ; more young ; very young ; very very young* ...}

Fuzziness properties of elements in $\mathbb{HA}$, specified by *fm* (fuzziness measure) [8] as follows:

Definition 2.1. A mapping $fm : \mathbb{L} \rightarrow [0, 1]$ is said to be the fuzziness measure of $\mathbb{L}$ if:

1. $\sum_{c \in \{c^+, c^-\}} fm(c) = 1,\ fm(0) = fm(w) = fm(1) = 0.$
2. $\sum_{h_i \in H} fm(h_i x) = fm(x)$, $x = h_n h_{n-1} \ldots h_1 c$, the canonical form.
3. $fm(h_n h_{n-1} \ldots h_1 c) = \prod_{i=1}^{n} fm(h_i) \times \mu(x)$.

2.2 Fuzzy Cognitive Map

Fuzzy cognitive map ($\mathbb{FCM}$) is feedback dynamical system for modeling fuzzy causal knowledge, introduced by Kosko [1]. $\mathbb{FCM}$ is a set of nodes, which present concepts and a set of directed edges to link nodes. The edges represent the causal links between these concepts. Mathematically, a $\mathbb{FCM}$ bis defined by.

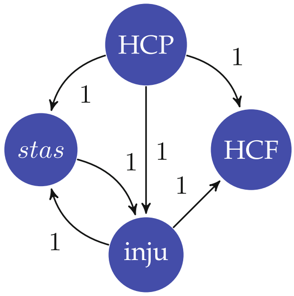
Definition 2.2. A $\mathbb{FCM}$ is a 4-Tuple:

$$\mathbb{FCM} = \{C, E, \mathcal{C}, f\} \tag{1}$$

In which:

1. $C = \{C_1, C_2, \ldots, C_n\}$ is the set of N concepts forming the nodes of a graph.
2. $E : (C_i, C_j) \longrightarrow e_{ij} \in \{-1,\ 0,\ 1\}$ is a function associating e_{ij} with a pair of concepts (C_i, C_j), so that e_{ij} = "weight of edge directed from C_i to C_j. The connection matrix $E(N \times N) = \{e_{ij}\}_{N \times N}$
3. The map: $\mathcal{C} : C_i \longrightarrow C_i(t) \in [0, 1], t \in N$
4. With $C(0) = [C_1(0, C_2(0), \ldots, C_n(0)] \in [0, 1]^N$ is the initial vector, recurring transformation function f defined as:

$$C_j(t+1) = f(\sum_{i=1}^{N} e_{ij} C_i(t)) \tag{2}$$



Example 2. Fig.1 shows a medical problem from expert domain of strokes and blood clotting involving. Concepts C={blood stasis (stas), endothelial injury (inju), hypercoagulation factors (HCP and HCF)} [2]. The conection matrix is:

$$E = (e_{ij})_{4\times4} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Fig. 1. A simple $\mathbb{FCM}$

$\mathbb{FCM}$s have played a vital role in the applications of scientific areas, including expert system, robotics, medicine, education, information technology, prediction, etc. [3,4] (Fig. 1).

3 Modeling with Words

Our model, based on linguistic variables, is constructed from linguistic hedge of $\mathbb{HA}$. The following are definitions in our research paper.

Definition 3.1 (Linguistic lattice)**.** With $\mathbb{L}$ as in the Sect. 2.1, set $\{\wedge, \vee\}$ are logical operators, defined in [7,8], a linguistic lattice $\mathcal{L}$ is a tuple:

$$\mathcal{L} = (\mathbb{L}, \vee, \wedge, 0, 1) \tag{3}$$

Property 3.1. *The following are some properties for* $\mathcal{L}$*:*

1. $\mathcal{L}$ *is a linguistic-bounded lattice.*
2. $(\mathbb{L}, \vee)$ *and* $(\mathbb{L}, \wedge)$ *are semigroups.*

Proof. Without loss of generality, let $\mathbb{L} = \{\rho\, c^+ | \rho \in (H^+)^* \wedge c^+ \in G\}$. W is the neutral element in $\mathbb{HA}$, we have:

1. $0 < w < c^+ < 1$ and for $\forall\ \rho \in (H^+)^*$: $\rho 0 < \rho w < \rho c^+ < \rho 1$. Because $\rho 0 = 0; \rho w = w; \rho 1 = 1$. This is equivalent to: $0 < \rho c^+ < 1$ or $\mathcal{L}$ is bounded
2. Let $\circ = \wedge$ or $\circ = \vee$ be operators in $\mathbb{HA}$ and $\{p, q, r\} \in X$. Applying definitions of operators $\wedge$ and $\vee$ from [9]:
   $p \circ (q \circ r) \wedge (\circ = \vee) = max\{p, max\{q, r\}\} = max\{p, q, r\} = (p \circ q) \circ r \wedge (\circ = \vee)$

$\square$

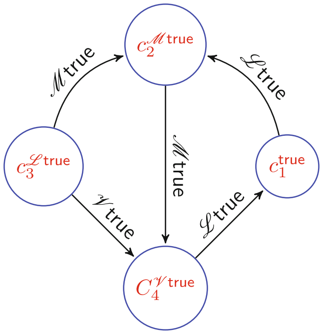
Definition 3.2. A linguistic cognitive map ($\mathbb{LCM}$) is a 4- Tuple:

$$\mathbb{LCM} = \{C, E, \mathcal{C}, f\} \tag{4}$$

In which:

1. $C = \{C_1, C_2, \ldots, C_n\}$ is the set of N concepts forming the nodes of a graph.
2. $E : (C_i, C_j) \longrightarrow e_{ij} \in \mathbb{L}$; $e_{ij} =$ "weight of edge directed from C_i to C_j. The connection matrix $E(N \times N) = \{e_{ij}\}_{N\times N} \in \mathbb{L}^{N\times N}$
3. The map: $\mathcal{C} : C_i \longrightarrow C_i(t) \in \mathbb{L}, t \in N$
4. With $C(0) = [C_1(0), C_2(0), \ldots, C_n(0)] \in \mathbb{L}^N$ is the initial vector, recurring transformation function f defined as:

$$C_j(t+1) = f(\sum_{i=1}^{N} e_{ij} C_i(t)) \in \mathbb{L} \tag{5}$$



Example 3. Fig. 2 shows a simple $\mathbb{LCM}$. Let

$$\mathbb{HA} = <\mathcal{X} = \mathsf{truth}; c^+ = \mathsf{true}; \mathcal{H} = \{\mathscr{L}, \mathscr{M}, \mathscr{V}\}> \tag{6}$$

be a $\mathbb{HA}$ with order as $\mathscr{L} < \mathscr{M} < \mathscr{V}$ ($\mathscr{L}$ for less, $\mathscr{M}$ for more and $\mathscr{V}$ for very are hedges).
$C = \{c_1, c_2, c_3, c_4\}$ is the set of 4 concepts with corresponding values
$\mathcal{C} = \{\mathsf{true}, \mathscr{M}\mathsf{true}, \mathscr{L}\mathsf{true}, \mathscr{V}\mathsf{true}\}$

Fig. 2. A simple $\mathbb{LCM}$

Square matrix:

$$M = (m_{ij} \in \mathbb{L})_{4\times 4} = \begin{vmatrix} 0 & \mathscr{L}\mathsf{true} & 0 & 0 \\ 0 & 0 & 0 & \mathscr{M}\mathsf{true} \\ 0 & \mathscr{M}\mathsf{true} & 0 & \mathscr{V}\mathsf{true} \\ \mathscr{L}\mathsf{true} & 0 & 0 & 0 \end{vmatrix}.$$

is the adjacency matrix of $\mathbb{LCM}$. Causal relation between c_i and c_j is m_{ij}, for example if $i = 1, j = 2$ then causal relation between c_1 and c_2 is: *"if* c_1 *is* true *then* c_2 *is* $\mathscr{M}\mathsf{true}$ *is* $\mathscr{L}\mathsf{true}$ "or let $\mathcal{P}$=" if c_1 is true then c_2 is $\mathscr{M}\mathsf{true}$" be a proposition then $\mathsf{truth}(\mathcal{P}) = \mathscr{L}\mathsf{true}$ (Fig. 2).

To have well fuzzify unit interval $\mathbb{I} = [0, 1]$ with $\mathbb{L}$, we need a number of hedges large enough. Assume $\hbar c = h_n h_{n-1} \ldots h_1 c$, accuracy in presenting interval $\mathbb{I}$ must be proportion to *length*($\hbar$). Set $|\hbar| = \mathsf{length}(\hbar)$, a question is what is the $|\hbar|$ value to be good enough for fuzzifying without increasing complexity. The following theorem clarifies this question.

Theorem 3.1. *Linguistic representations for unit interval*

- *Let sequences* $\{\mathbb{I}_j\}_{j=1}^k$ *be* k *partitions of* $\mathbb{I}$, I *is index set, for* $\forall j, l \in I \wedge j \neq l$
$$\mathbb{I} = \cup_{j=1}^k \mathbb{I}_j; \mathbb{I}_j \cap \mathbb{I}_l = \emptyset \tag{7}$$
- *Let*
$$\hbar c = h_n h_{n-1} \dots h_1 c \tag{8}$$
be a linguistic variable, $\hbar c \in \mathbb{L}$, $\bigvee$ *for max operator,* H *is a set of hedges* $h_i, i \in I$.

Then:
$$|\hbar| < \frac{\log(\frac{k-1}{k})}{\log(\bigvee_{h_i \in H} fm(h_i))} \tag{9}$$

Proof.

$$\begin{aligned}\mathbb{I} &= \{\mathbb{I}_1, \mathbb{I}_2, \dots, \mathbb{I}_k\} \\ &= \{(0, \frac{1}{k}], (\frac{1}{k}, \frac{2}{k}], \dots, (\frac{k-2}{k}, \frac{k-1}{k}], (\frac{k-1}{k}, 1)\}\end{aligned}$$

Let $\hbar_k c = h_n h_{n-1} \dots h_1 c \in \mathbb{L}$ to represent a point in $\mathbb{I}_k$, then

$$\begin{aligned} fm(\hbar_k c) &= \prod_{h_i \in H}^{n} fm(h_i) \times \mu(c) \\ &\leq (\bigvee_{h_i \in H} fm(h_i))^n \times \mu(c) < (\bigvee_{h_i \in H} fm(h_i))^n. \, because \; \mu(c) < 1 \\ so,\ \exists n \in \mathbb{N} : \frac{k-1}{k} &< (\bigvee_{h_i \in H} fm(h_i))^n < 1 \\ And \; : \log(\frac{k-1}{k}) &< \log((\bigvee_{h_i \in H} fm(h_i))^n) < \log(1) \\ Therefore \; : \frac{\log(\frac{k-1}{k})}{\log(\bigvee_{h_i \in H} fm(h_i)} &> n > 0 \end{aligned}$$

□

Theorem 3.1 is important in limiting the length of linguistic variables for fuzzifying knowledge. On the other hand, knowing $|\hbar|$ value, the complexity of computations will be decreased

From Eq. (5), let $\mathcal{C}^n = \{C(i)\}_{i=1}^n = \{C(0), C(1), \dots, C(n)\}$ be a set of state space, we want to know whether $\mathcal{C}^n$ finite or infinite. Finding $\mathcal{C}^n$ helps to limit searching space in many cases. Size of $\mathcal{C}^n$, say $|\mathcal{C}^n|$, proportions to $|\hbar|$ and size of vertices $|C|$ in $\mathbb{LCM}$.

Theorem 3.2. *State space $\mathcal{C}^n$, generating by $\mathbb{LCM}$ with $\mathbb{N}$ vertices, using $\hbar$ hedges is:*

$$|\mathcal{C}^n| = |\hbar|^{\mathbb{N}\times|\hbar|} \tag{10}$$

Proof. It is straightforward to prove Theorem 3.2 by using combinatory algebra. $\hbar$ hedges produce $|\hbar|^{|\hbar|}$ combinations which are hedges string with length $= |\hbar|$. $\mathbb{LCM}$ has $\mathbb{N}$ vertices, $|\hbar|^{|\hbar|}$ cases for each vertex, apply the rule of product, we have $\overbrace{|\hbar|^{|\hbar|} \times |\hbar|^{|\hbar|} \times \ldots \times |\hbar|^{|\hbar|}}^{\mathbb{N}\ times} = |\hbar|^{\mathbb{N}\times|\hbar|}$. □

4 Discussion and Future Work

We have introduced a new graphical model for representing fuzzy knowledge using linguistic variables from $\mathbb{HA}$. Our model, called $\mathbb{LCM}$, extended from $\mathbb{FCM}$, is a dynamical system with two properties: static and dynamic. Static properties allow forward or what-if inferencing between concepts on linguistic domain. Especially, we indicate inverse proportion relationship between length of hedges string and a number of partitions in representing fuzzy knowledge.

Dynamic behaviors are transformation states in state space $\mathcal{C}^n = \{\mathcal{C}\}_0^n = \{\mathcal{C}(0),\ \mathcal{C}(1), \ldots, \mathcal{C}(n)\}$, where $\mathcal{C}(i) = \{C_1(i),\ C_2(i), \ldots,\ C_{\mathbb{N}}(i)\}, i = \overline{0, n}$. We also prove the theorem about the number of states in state space is $|\mathcal{C}^n| = |\hbar|^{\mathbb{N}\times|\hbar|}$, this is the important theorem to decide whether or not installable computer programs.

Our next study is as follow: Let $A = \{\hbar^n :\ \hbar^n = h_n h_{n-1} \ldots h_1 h_0$ with $h_i \in H,\ i = \overline{0, n}\}$ be a string of hedges. Assume $I = \mathcal{C}(0)$, $T = \mathcal{C}(n)$ and $\mathcal{T} \subset \mathcal{C}\times A\times\mathcal{C}$ in order are initial, final and transition states. We will prove that $\mathbb{LCM}$ actions are fuzzy linguistic automata $\mathcal{A} = <A,\ \mathcal{C},\ I,\ \mathcal{T},\ T>$.

5 Conclusion

A new visual method for modeling with words representations of fuzzy knowledge based on $\mathbb{HA}$ has been proposed in this reseach paper. Our model has been shown easily reading, understanding and presenting of human. Theorem on limiting the hedges string length is clearly proved to reduce complexity in representing method and counting on the number of states in state space of fuzzy $\mathbb{LCM}$ is demonstrated with certainty.

References

1. Kosko, B.: Fuzzy cognitive maps. Int. J. Man Mach. Stud. **24**, 65–75 (1986)
2. Osoba, O.A., Kosko, B.: Fuzzy cognitive maps of public support for insurgency and terrorism. J. Defense Model. Simul. Appl. Methodol. **14**(I), 17–32 (2017)
3. Glykas, M.: Fuzzy Cognitive Maps, Advances in Theory Tools and Applications. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-03220-2

4. Papageorgiou, E.I.: Fuzzy Cognitive Maps for Applied Science and Engineering From Fundamentals to Extensions and Learning Algorithms. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-642-39739-4
5. Zadeh, L.A.: Computing with words - Principal Concepts and Ideas. Studies in Fuzziness and Soft Computing. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-27473-2
6. Kacprzyk, J., Wilbik, A., Zadrożny, S.: Linguistic summarization of trends: a fuzzy logic based approach. In: The 11th International Conference Information Processing and Management of Uncertainty in Knowledge-based Systems, pp. 2166–2172 (2006)
7. Ho, N.C., Wechler, W.: Hedge algebras: an algebraic approach to structure of sets of linguistic truth values. Fuzzy Sets Syst. **35**, 281–293 (1990)
8. Ho, N.C., Van Long, N.: Fuzziness measure on complete hedge algebras and quantifying semantics of terms in linear hedge algebras. Fuzzy Sets Syst. **158**(4), 452–471 (2007)
9. Phuong, L.A., Khang, T.D.: Generalized if... then... else... inference rules with linguistic modifiers for approximate reasoning. Int. J. Comput. Sci. Issues (IJCSI) **9**(6), 184–190 (2012). No 3