Tạp chí Tin học và Điều khiển học, T.27, S.3 (2011), 241–252

VẤN ĐỀ KẾT NHẬP THÔNG TIN BIỂU DIỄN BẰNG BỘ 4 VỚI NGỮ NGHĨA DỰA TRÊN ĐẠI SỐ GIA TỬ

NGUYỄN VĂN LONG, HOÀNG VĂN THÔNG

Đại học Giao thông vận tải

Tóm tắt. Bài toán kết nhập thông tin nói chung đóng vai trò quan trọng trong quá trình lấy quyết định và do đó nó có ý nghĩa ứng dụng rộng lớn, đặc biệt là bài toán kết nhập thông tin mờ vì con người thường quyết định thông qua thông tin mờ ngôn ngữ. Cho đến nay các phương pháp giải bài toán này chủ yếu dựa trên các tập mờ. Trong [8] và [7] các tác giả đã nghiên cứu giải bài toán này khi ứng dụng ngữ nghĩa ngôn ngữ dựa trên đại số gia tử (ĐSGT) với biểu diễn ngữ nghĩa tương ứng bằng dữ liệu bộ 2 và bộ 3. Trong bài báo này, thông tin mờ của thang đánh giá được biểu diễn bằng bộ 4 để khắc phục một số thiếu sót trong [7]. Nó cho phép biểu thị thông tin mờ đầy đủ hơn và cho phép tính toán trên các thông tin ngôn ngữ trong thang đánh giá trực tiếp hơn và do đó kết quả kết nhập tự nhiên và chính xác hơn.

Abstract. The role of information aggregation problems is in general very important in the decision process and hence they can find applications in many fields, particularly, for the *fuzzy information aggregation problem* since human makes decisions by means of vague linguistic terms. So far there are many methods have been exploited to solve these problems but mainly based on fuzzy sets. The works [8] and [7] were devpted to solve this problem, applying hedge-algebra-based semantics of linguistic terms to representing linguistic data by 2-tuples or 3-tuples. In this paper we introduce 4-tuple representation exploiting more deeply hedge-algebra-based semantics of linguistic terms for solving this problem. This allows represent the semantics of linguistic symbols in rating scale more completely and hence it produces output results more exactly and naturaly.

1. MỞ ĐẦU

Rất nhiều vấn đề của việc lấy quyết định hàng ngày của con người đều dựa trên các thông tin không chính xác có bản chất mờ, không chắc chắn. Lý thuyết tập mờ đã tạo ra các mô hình toán học và các công cụ cho phép đưa ra các cách tiếp cận khác nhau để giải các bài toán lấy quyết định trong những lĩnh vực khác nhau. Những thành tựu ứng dụng của lý thuyết tập mờ lại thúc đẩy sự phát triển của chính lý thuyết đó và do vậy đã ra đời các chuyên ngành như công nghệ logic mờ (Fuzzy logics technology), công nghệ kết nhập (Aggregation technology), tính toán với các từ ngôn ngữ (Computing with Words-CW) ...

Tính toán với các từ ngôn ngữ là một lý thuyết cho phép thao tác trên các từ để giải quyết các bài toán mà dữ liệu đầu vào và đầu ra cần biểu thị bằng các từ ngôn ngữ, đặc biệt đối với các bài toán lấy quyết định. Một bài toán quan trọng của quá trình lấy quyết định là tổng hợp, kết gộp các thông tin, dữ liệu ngôn ngữ từ các nguồn khác nhau (từ các chuyên gia chẳng hạn), để có một thông tin đánh giá đại diện chung, được gọi là *bài toán kết nhập* trên các từ ngôn ngữ. Bài toán có thể được mô tả như sau:

Giả sử người quyết định phải lấy quyết định chọn một phương án "tốt nhất" trong m phương án lựa chọn A_i, $i = 1, \ldots, m$, trên cơ sở lấy ý kiến đánh giá của n chuyên gia e_j, $j = 1, \ldots, n$. Trong môi trường thông tin ngôn ngữ, các chuyên gia biểu thị đánh giá của mình bằng các từ ngôn ngữ (thang đánh giá ngôn ngữ) lấy trong tập $\boldsymbol{S} = s_0, \ldots, s_g$. Ký hiệu x_{ij} là ý kiến đánh giá của chuyên gia j về phương án A_i.

Một yêu cầu tự nhiên là cần định giá ý kiến tổng hợp của các chuyên gia đối với từng phương án, nghĩa là ta cần sử dụng một phép toán kết nhập @ tích hợp các ý kiến $\{x_{ij} : j = 1, \ldots, n\}$ của các chuyên gia. Một cách hình thức toán học, toán tử kết nhập là một ánh xạ $@ : \{s_0, \ldots, s_g\}^n \to \{s_0, \ldots, s_g\}$. Ánh xạ này phải được xác định sao cho kết quả của phép toán $@(s_{i1}, \ldots, s_{in})$ có thể xem là biểu thị ý kiến tập thể của n chuyên gia.

Như vậy, bài toán kết nhập là một bài toán cốt yếu trong quá trình lấy quyết định, và vì giá trị của toán tử cũng là một từ ngôn ngữ, nên người ta cần sử dụng công cụ tính toán trên các từ ngôn ngữ. Có nhiều phương pháp tiếp cận tính toán khác nhau [1-5, 7, 8, 10] để giải quyết vấn đề này. Để làm rõ nội dung nghiên cứu của bài báo, ta sẽ phân tích tổng quan một vài phương pháp.

1.1. Phương pháp tính toán ngôn ngữ dựa trên nguyên lý mở rộng của tập mờ

Ý tưởng chính của phương pháp là các phép tính kết nhập kinh điển như phép trung bình số học, trung bình có trọng số ..., có thể chuyển thành các phép tính tương ứng trên các tập mờ, chẳng hạn phép lấy trung bình cộng mờ, trung bình cộng mờ có trọng số trên các tập mờ ... Khi đó, các từ ngôn ngữ trong tập $\boldsymbol{S}$ được xem là các nhãn của các tập mờ. Các phép kết nhập mờ thực hiện trên các tập mờ của các nhãn trong tập $\boldsymbol{S}$ sẽ cho kết quả là tập mờ. Nói chung tập mờ kết quả khác với các tập mờ của các nhãn, hay nó không biểu thị cho một nhãn ngôn ngữ nào trong $\boldsymbol{S}$. Điều này dẫn đến sự cần thiết phải phát triển các *phương pháp xấp xỉ ngôn ngữ*, tức là tìm một nhãn ngôn ngữ trong $\boldsymbol{S}$ có tập mờ xấp xỉ tập mờ kết quả nhất.

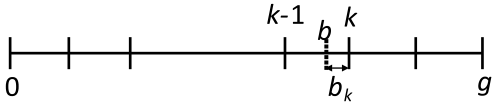
1.2. Phương pháp tính toán trên các ký hiệu ngôn ngữ

Giả sử ý kiến đánh giá theo một tiêu chí được biểu thị bằng các từ ngôn ngữ trong tập $\boldsymbol{S} = \{s_0, \ldots, s_g\}$ được sắp tuyến tính theo ngữ nghĩa của chúng sao cho: $s_i < s_j$ nếu và chỉ nếu $i < j$. Vì không thể tính trực tiếp trên các từ nên người ta mượn cấu trúc tính toán của đoạn $[0, g]$ bao hàm các chỉ số để thực hiện việc kết nhập số học. Ý tưởng này thể hiện như sau [3-5].

Giả sử ta lấy kết nhập tập các từ ngôn ngữ trong $A = \{a_1, \ldots, a_p\}$, $a_i \in \boldsymbol{S}$. Ta thực hiện một hoán vị các chỉ số của tập A, $A = \{a_{\pi 1}, \ldots, a_{\pi p}\}$, sao cho $a_{\pi i} \geq a_{\pi j}$ nếu $i \leq j$. Xét một phép kết nhập số học g nào đó, g sẽ cảm sinh một phép kết nhập g^* trên tập $\boldsymbol{S}$ được định nghĩa như sau: Tính $g(\pi_1, \ldots, \pi_p) \in [0, g]$, với $\pi_1, \ldots, \pi_p$ là các chỉ số của các phần tử trong A. Đặt $i^* = \text{round}(g(\pi_1, \ldots, \pi_p))$, trong đó round là phép làm tròn số học. Khi đó phần tử s_{i^*} được xem là kết quả kết nhập $g^*(a_{\pi_1}, \ldots, a_{\pi_p})$. Phép kết nhập g^* được định nghĩa như trên là một hạn chế, vì nó chịu một ràng buộc ít tự nhiên và mất mát nhiều thông tin. Hiển nhiên, thay vì dùng chỉ số các từ như trên, nếu ta sử dụng ngữ nghĩa định lượng của các từ ngôn ngữ trong thang đánh giá thì định nghĩa phép kết nhập $g*$ có cơ sở hợp lý hơn (xem [10]).

1.3. Phương pháp tính toán ngôn ngữ dựa trên biểu diễn dữ liệu bộ 2

Trong phương pháp trên ta cần làm tròn bằng biểu thức $i^* = \text{round}(g(\pi_1, \ldots, \pi_p))$ để kết quả là một từ ngôn ngữ a_{i^*} trong tập $\boldsymbol{S}$. Tuy nhiên việc làm tròn làm mất mát thông tin và các tác giả [4] đã đưa ra cách biểu diễn dữ liệu bộ 2 để khắc phục sự mất mát thông tin này.



Hình 1.1. Biểu diễn bộ 2

Ý tưởng của phương pháp như sau: giá trị $g(\pi_1, \ldots, \pi_p)$ chưa làm tròn sẽ là một số thực $b \in [0, g]$, $k - 0.5 \leq b < k + 0.5$. Để biểu thị rõ các thông tin các tác giả trong [5] đề xuất dạng biểu diễn bộ 2, (s_k, b_k), trong đó $b_k = b - k \in [-0.5, +0.5)$. Từ s_k được gọi là tâm của thông tin còn a_k biểu thị giá trị chuyển dịch từ giá trị gốc b về từ ngôn ngữ s_k. Biểu diễn này cũng xác định một ánh xạ sau:

$$\Delta : [0, g] \to \boldsymbol{S} \times [-0.5, 0.5); \Delta(b) = (s_k, b_k), \text{ với } b_k = b - k, k = \text{round}(b) \tag{1.1}$$

và, do đó, ta có

$$\Delta^{-1} : \boldsymbol{S} \times [-0.5, 0.5) \to [0, g]; \Delta^{-1}((s_k, b_k)) = k + b_k \in [0, g].$$

Với biểu diễn như vậy, các phép kết nhập trên các từ ngôn ngữ được định nghĩa thông qua các phép kết nhập số học nhờ việc chuyển đổi của ánh xạ Δ^{-1}.

1.4. Phương pháp tính toán ngôn ngữ dựa trên biểu diễn dữ liệu bộ 3

Trong [8,7], các tác giả đã nghiên cứu phát triển việc tính toán trên các từ dựa trên đại số gia tử (ĐSGT) với việc tận dụng các khoảng độ đo tính mờ của các từ ngôn ngữ. Theo cách tiếp cận của ĐSGT, ngữ nghĩa các từ ngôn ngữ có thể biểu thị qua các khoảng tính mờ mức l, với l chỉ độ dài của sâu biểu diễn các từ ngôn ngữ. Cho trước l và xét một từ s bất kỳ độ dài l. Biểu diễn ngữ nghĩa bộ 3 [7] của từ s là

$$(s; \upsilon(s); \Omega(s)) \tag{1.2}$$

trong đó $\upsilon(s)$ là giá trị định lượng của từ s, $\Omega(s)$ là khoảng tính mờ của từ s. Ta biết rằng $\upsilon(s) \in \Omega(s)$.

Lợi ích của phương pháp này là thay cho việc phải xét tập chỉ số của các từ, một đại lượng mang ít thông tin về ngữ nghĩa của từ, ta sử dụng giá trị thực $\upsilon(s)$ và khoảng tính mờ $\Omega(s)$ chứa nó mang ngữ nghĩa định lượng của từ s. Rõ ràng biểu diễn bộ 3 mang nhiều thông tin hơn và tự nhiên hơn so với biểu diễn bộ 2.

Nhược điểm của cách tiếp cận này là giá trị thực $\upsilon(s)$ được xác định bởi hàm định lượng υ là một ràng buộc chặt, vì chẳng hạn kết quả của phép kết nhập số học không nhất thiết trùng với giá trị thực của từ ngôn ngữ biểu thị kết quả kết nhập trên các bộ 3 dữ liệu. Ngoài ra trong [7] cũng đưa ra một ràng buộc chặt là các từ kết nhập được biểu thị bằng các bộ 3 dữ liệu trên cùng độ dài l.

Để khắc phục những thiếu sót nêu trên, trong bài báo này chúng tôi mở rộng biểu diễn bộ 3 thành một biểu diễn bộ 4 biểu thị nhiều thông tin hơn và cho phép linh hoạt hơn trong

việc giải các bài toán kết nhập trên các từ ngôn ngữ và trong giải quyết bài toán quyết định dựa trên ý kiến nhiều chuyên gia.

Nội dung còn lại của bài báo được tổ chức như sau. Mục 2 trình bầy phương pháp hình thức hóa biểu diễn ngữ nghĩa bằng bộ 4 các từ ngôn ngữ trong thang đánh giá và các phép kết nhập trên dữ liệu biểu diễn bộ 4. Mục 3 mô tả cách giải bài toán lấy quyết định bằng các phương pháp biểu diễn bộ 4 do bài báo đề xuất và so sách kết quả với các kết quả của các phương pháp khác. Việc phân tích so sánh các kết quả của các phương pháp sẽ chứng tỏ lợi ích và hiệu quả của phương pháp mới.

2. XÂY DỰNG THANG ĐIỂM NGÔN NGỮ VỚI NGỮ NGHĨA BIỂU DIỄN BẰNG BỘ 4

2.1. Lân cận ngữ nghĩa của các từ ngôn ngữ

Đặt $\boldsymbol{S} = \boldsymbol{X}_{(k)} = \boldsymbol{X}_1 \cup \ldots \cup \boldsymbol{X}_k$, nhớ rằng $\boldsymbol{X}_l$ là tập các từ có độ dài l.

Như đã trình bày trong Mục 1, đối với bài toán đặt ra ta cần xét thang đánh giá bằng các từ ngôn ngữ $\boldsymbol{S} = X_{(k)} = s_0, \ldots, s_g$. Để đưa ra cách biểu diễn ngữ nghĩa bằng bộ 4, trước hết chúng ta định nghĩa lân cận ngữ nghĩa của các từ ngôn ngữ. Ext. ĐSGT $\underline{\mathbf{AX}} = (\boldsymbol{X}, \boldsymbol{G}, \boldsymbol{C}, \boldsymbol{H}, \Sigma, \Phi, \leq)$. Để định nghĩa lân cận ngữ nghĩa mức k, ta xét tập các khoảng tính mờ mức $k+1$, I_{k+1}. Ta biết rằng I_{k+1} tạo thành một phân hoạch của đoạn [0,1].

Xét tập các khoảng tính mờ mức $k+1$, I_{k+1}. Ta sẽ phân cụm các khoảng tính mờ trong I_{k+1} như sau. Vì các khoảng tính mờ trong I_{k+1} rời nhau và kế tiếp nhau, ta sắp xếp chúng theo thứ tự tăng dần. Với $2 \leq p, q$, đặt

$$\boldsymbol{H}_1 = \{h_i, h_{-j} : 1 \leq i \leq [p/2] \& 1 \leq j \leq [q/2]\},\ \boldsymbol{H}_2 = \{h_i, h_{-j} : [p/2] < i \leq p \& [q/2] < j \leq q\}$$

Việc phân cụm sẽ được thực hiện như sau. Ta duyệt từ trái sang phải các khoảng tính mờ $\Im_{k+1}(h_i y_j)$, $h_i \in \boldsymbol{H}$ và $y_j \in \boldsymbol{X}_k$, trong I_{k+1}. Chúng sẽ được xếp vào cùng một cụm nếu các khoảng kế tiếp có các gia tử h_i cùng thuộc tập $\boldsymbol{H}_n$, với $n = 1$ hoặc $n = 2$. Có thể kiểm tra thấy rằng tập

$$\mathcal{C}_{y,1} = \{\Im_{k+1}(h_i y) : h_i \in \boldsymbol{H}_1\}, y \in \boldsymbol{X}_k, \tag{2.3}$$

chính là một cụm theo cách phân cụm này và các khoảng trong I_{k+1} nằm giữa hai cụm dạng (2.3) kế tiếp là cụm chứa một số khoảng $\Im_{k+1}(h_i y) : h_i \in \boldsymbol{H}_2$.

Ký hiệu $\mathcal{C}$ là tập tất cả các cụm theo cách phân cụm trên và mỗi cụm của nó được ký hiệu là $\mathcal{C}$. Có thể dễ dàng kiểm tra thấy rằng họ $\mathcal{C}$ là một phân cụm của tập I_{k+1}:

(i) Các cụm $\mathcal{C}$ là rời nhau;

(ii) Mọi khoảng tính mờ trong I_{k+1} đều thuộc vào một cụm $\mathcal{C}$ nào đó.

Đặt $\boldsymbol{X}_{(k)} = \boldsymbol{X}_1 \cup \ldots \cup X_k$ và

$$\mathcal{S}_k(\mathcal{C}) = \mathbf{Y}\{\Im_{k+1} : \Im_{k+1} \in C\} \tag{2.4}$$

Mệnh đề 2.1. (i) *$\mathcal{S}_k(\mathcal{C})$ là một khoảng mở ở đầu mút phải và đóng ở đầu mút trái, trừ cụm sinh ra khoảng tận cùng phải là đoạn đóng cả hai đầu mút, và họ $\mathcal{S}_k(\mathcal{C}) : \mathcal{C} \in \rfloor$ là một phân hoạch của đoạn [0,1];*

(ii) *Với mọi $y \in \boldsymbol{X}_{(k)}$, tồn tại một khoảng $\mathcal{S}_k(\mathcal{C})$ sao cho giá trị định lượng $\upsilon(y)$ nằm trong $\mathcal{S}_k(\mathcal{C})$ cùng với hai khoảng tính mờ mức $k+1$ kề với nó.*

Việc chứng minh mệnh đề trên là rất đơn giản, chúng tôi không trình bầy ở đây.

Định nghĩa 2.1. Với mọi $y \in \boldsymbol{X}_{(k)}$, khoảng (2.4) có tính chất $\upsilon(y) \in \mathcal{S}_k(\mathcal{C})$, và nó là điểm trong của khoảng $\mathcal{S}_k(\mathcal{C})$ với tôpô tự nhiên trên miền số thực, được gọi là lân cận ngữ nghĩa mức k của từ ngôn ngữ y.

2.2. Biểu diễn từ ngôn ngữ bằng bộ 4 ngữ nghĩa

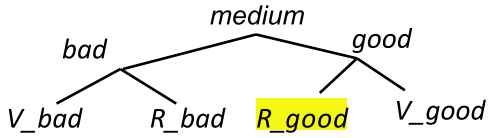
Cho một ĐSGT tuyến tính và đầy đủ $\mathbf{AX} = (\boldsymbol{X}, \boldsymbol{G}, \boldsymbol{C}, \boldsymbol{H}, \leq)$ và xét một tập hữu hạn các từ ngôn ngữ $S \in X$. Tập $\boldsymbol{S}$ được gọi là đầy đủ ngữ nghĩa, hay để đơn giản gọi là đầy đủ, nếu

$$hx \in \boldsymbol{S} \Rightarrow kx \in \boldsymbol{S}, vi \forall k \in H \tag{2.5}$$

Khái niệm đầy đủ bảo đảm rằng thang đánh giá X_0 ở mỗi mức sinh từ của nó có đủ mặt các từ ngôn ngữ sinh bằng gia tử. Rõ ràng rằng $\boldsymbol{X}_{(k)}$ là đầy đủ ngữ nghĩa.

Ví dụ, ta xét một ĐSGT $\mathbf{AX}$ với $\boldsymbol{G} = \{bad,\ good\}$, $\boldsymbol{H} = \{R, V\}$, trong đó các chữ cái là sự rút gọn tương ứng của các từ *Rather* và *Very*. Giả sử $\boldsymbol{S} = \{V_bad, bad, R_bad, medium, R_good, good, V_good\}$. Thang đánh giá này là đầy đủ và được biểu thị bằng dạng cây như trong Hình 2.2. Tính đầy đủ của tập S được kiểm chứng bằng cách xét các nút con của một nút cha, nếu chúng sinh từ cha bằng gia tử thì có đủ mặt các gia tử R và V.

Tập X_0 là tập sắp tuyến tính với quan hệ thứ tự là: $V_bad \leq bad \leq R_bad \leq medium \leq RR_good \leq R_good \leq VR_good \leq good \leq V_good$.



Hình 2.1. Cây biểu diễn tập **S**

Dưới đây chúng ta đưa ra cách biểu diễn thông tin bộ 4 của các từ ngôn ngữ trong thang đánh giá ngôn ngữ.

Định nghĩa 2.2. Xét một ĐSGT $\mathbf{AX} = (\boldsymbol{X}, \boldsymbol{G}, \boldsymbol{C}, \boldsymbol{H}, \leq)$ được trang bị một độ đo tính mờ fm. Giả sử l là số nguyên dương. Biểu diễn bộ 4 mức l của một từ ngôn ngữ $s \in \boldsymbol{X}$ là bộ 4 ngữ nghĩa $(s, \upsilon(s), r, \mathcal{S}_l(s))$, trong đó $\upsilon(s)$ là giá trị định lượng của từ s, $\mathcal{S}_l(s)$ là khoảng lân cận ngữ nghĩa mức l của s, gọi là mức lân cận, và $r, \upsilon(s) \in \mathcal{S}_l(s)$. Hiệu số $r - \upsilon(s)$ được gọi là độ lệch ngữ nghĩa của giá trị được chọn r.

Khác với biểu diễn bộ 3 trong [7], các từ s không bị ràng buộc phải có cùng độ dài và thêm thành phần mới r. Thành phần này bảo đảm các phép tính kết nhập trên bộ 4 là đóng. Ngoài ra, về thực tiễn nó cho phép các chuyên gia có thể cho điểm đánh giá bằng số thực trong miền tham chiếu của biến ngôn ngữ.

Vì ta không thể tính toán trực tiếp trên các từ s, cách biểu diễn ngữ nghĩa của từ như vậy cho phép ta tính toán trên giá trị thực r. Ý nghĩa của ánh xạ ngữ nghĩa định lượng $\upsilon(s)$ là giá trị thực mang nhiều thông tin của từ s nhất và có thể xem nó tương ứng với tâm điểm (core) của tập mờ biểu thị ngữ nghĩa của s. Với ý nghĩa đó, hiệu số $r - \upsilon(s)$ mang thông tin về độ phù hợp của r đối với từ s. Như vậy thay vì tính toán trên các chỉ số của tập $\boldsymbol{S}$, trong cách tiếp cận này chúng ta tính toán trên số thực r nằm trong lân cận $\mathcal{S}_l(s)$-khoảng lân cận ngữ nghĩa của s.

2.3. Xây dựng biểu diễn bộ 4 ngữ nghĩa cho thang đánh giá ngôn ngữ

Trong nhiều mô hình ứng dụng người ta mong muốn cho phép các chuyên gia đánh giá các tiêu chí đối với các đối tượng cần đánh giá để ra quyết định theo thang điểm ngôn ngữ. Xét một thang đánh giá ngôn ngữ $\boldsymbol{S}$ có thứ tự tuyến tính dựa trên ngữ nghĩa của chúng.

Giả sử tập thang đánh giá $\boldsymbol{S}$ có tính chất "kế tiếp", bảo đảm tính đa dạng của thang đánh giá ngôn ngữ, sau:

$$(\forall s \in \boldsymbol{S})(\exists s' \in \boldsymbol{S})(\exists h \in \boldsymbol{H})(s = hs') \tag{2.6}$$

Ý nghĩa của (2.6) là thang đánh giá ngôn ngữ liên tục về mức mô tả ngữ nghĩa sinh ra bởi các gia tử: *không có từ nào trong thang đánh giá mà phải sử dụng hai lần tác động gia tử kế tiếp vào một từ nào đó đã có trong thang đánh giá mới sinh được ra nó.* Nói khác đi, nếu cần hai lần tác động gia tử thì có nghĩa trong tập $\boldsymbol{S}$ còn thiếu một số từ ngôn ngữ nào đó.

Ví dụ, thang đánh giá ngôn ngữ S = bad, medium, *RR_good, VR_good, good, RV_good, VV_good* là không đầy vì "*RR_good*" chỉ có thể sinh được từ "*good*" trong $\boldsymbol{S}$ bằng việc sử dụng hai lần tác động liên tiếp gia tử R.

Thang đánh giá ngôn ngữ bằng bộ 4 ngữ nghĩa được định nghĩa như sau.

Định nghĩa 2.3. Cho một thang đánh giá ngôn ngữ $\boldsymbol{S} \subseteq \mathbf{X}$ gồm các từ ngôn ngữ, và một số nguyên dương l, một biểu diễn bộ 4 ngữ nghĩa của $\boldsymbol{S}$, $\{(s, \upsilon(s), r, \boldsymbol{S}(s)) : s \in \boldsymbol{S}, r \in \boldsymbol{S}(s)\}$, r là kí hiệu biến thực, được gọi là một thang đánh giá ngôn ngữ dựa trên bộ 4 ngữ nghĩa nếu tập $\boldsymbol{S}(s) : s \in \boldsymbol{S}$ lập thành một phân hoạch của đoạn [0,1], tức là ta có:

(i) $\boldsymbol{S}(s) \in \boldsymbol{S}(s') = \varnothing, \forall s, s' \in \boldsymbol{S}$ và $s \neq s'$;

(ii) $\boldsymbol{S}(s) = [0, 1]$.

Giá trị thực r trong bộ 4 ngữ nghĩa đóng vai trò như điểm thực trong thang đánh giá thực trong trường hợp kinh điển. Ở đây đoạn [0,1] là thang điểm thực được chuẩn hóa để tiện trình bày về phương pháp luận. Mối quan hệ định lượng giữa s và r thể hiện quan hệ ngữ nghĩa trong biểu diễn bộ 4. Điều kiện (i) bảo đảm rằng một giá trị thực r không thể là biểu thị cùng lúc cho hai từ ngôn ngữ khác nhau. (ii) nói rằng bất kỳ giá trị nào của thang điểm thực cũng có thể được chọn để biểu thị một từ nào đó trong $\boldsymbol{S}$ với độ sai lệch được đo bằng $r - \upsilon(s)$.

Với cách biểu diễn bộ 4 ngữ nghĩa việc so sánh hai bộ được xác định bởi thứ tự của các từ s trong $\boldsymbol{S}$ và *thứ tự của giá trị thực trong thành phần thứ* 3 của bộ 4. Cụ thể là:

$$(s_i, \upsilon(s_i), r_i, \boldsymbol{S}(s_i)) \leq_{df} (s_j, \upsilon(s_j), r_j, \boldsymbol{S}(sj)) \text{ nếu } s_i < s_j \text{ hoặc } s_i = s_j \ \&\ r_i \leq r_j.$$

và $\quad (s_i, \upsilon(s_i), r_i, \boldsymbol{S}(s_i)) =_{df} (s_j, \upsilon(s_j), r_j, \boldsymbol{S}(s_j))$ nếu $s_i = s_j \ \&\ r_i = r_j$.

Vì $\{\boldsymbol{S}(s) : s \in \boldsymbol{S}\}$ lập thành một phân hoạch, có thể thấy rằng $(s_i, \upsilon(s_i), r_i, \boldsymbol{S}(s_i)) \leq (s_j, \upsilon(s_j), r_j, \boldsymbol{S}(s_j)) \Leftrightarrow r_i \leq r_j$. Một bài toán đặt ra là, cho thang đánh giá ngôn ngữ $\boldsymbol{S}$, có thể hay không xây dựng được một thang đánh giá ngôn ngữ dựa trên bộ 4 ngữ nghĩa cho $\boldsymbol{S}$.

Giả sử k_S là độ dài lớn nhất của các từ trong S, nghĩa là k_S là cận trên đúng của các $|s|$ với mọi s trong $\boldsymbol{S}$. Ta sẽ dựa trên các khoảng tính mờ của $\boldsymbol{S}$ để xây dựng các khoảng $\boldsymbol{S}(s)$, với l là mức lân cận. Trước hết ta đưa ra khái niệm "Khoảng lân cận kề" của s.

Như trên, ta kí hiệu $\mathbf{X}_l$ là tập các từ trong $\boldsymbol{X}$ có độ dài l. Xét tập các khoảng tính mờ mức l, $\mathbf{I}_l = \{\mathcal{J}_l(x) : x \in \mathbf{X}_l\} = \bigcup\{\{\mathcal{J}_l(h_i u) : i \in [-q, p]\} : u \in \mathbf{X}_{l-1}\}$. Với $l > k_S$, như chúng ta biết các giá trị $\upsilon(s), s \in S$, đều là các đầu mút của các khoảng tính mờ mức l. Hợp của hai khoảng tính mờ mức l có cùng chung đầu mút $\upsilon(s)$ được gọi là *khoảng lân cận kề* của s, $s \in \boldsymbol{S}$.

Mệnh đề 2.1 chịu sự ràng buộc là $2 \leq p, q$ được mở rộng bằng mệnh đề sau.

Mệnh đề 2.2. *Giả sử tập $\boldsymbol{S}$ có tính chất kế tiếp. Khi đó mức l nhỏ nhất để các khoảng lân cận kề của $\upsilon(s)$, cũng kí hiệu là $\mathcal{S}_l(s)$, với $s \in \boldsymbol{S}$, rời nhau là:*

(i) $l = k_S + 1$, *với* $p, q \geq 2$;

(ii) $l = k_S + 2$ *với* $p = 1$ *hoặc* $q = 1$.

Mệnh đề 2.2 bảo đảm rằng đối với tập $\boldsymbol{S}$ có tính kế tiếp ta luôn luôn xây dựng được hệ khoảng lân cận mức l rời nhau. Việc mở rộng một hệ khoảng rời nhau như vậy để nó trở thành phân hoạch của $[0, 1]$ là không khó khăn về mặt thuật toán.

Sau đây ta đưa ra thuật toán xây dựng biểu diễn bộ 4 ngữ nghĩa của tập $\boldsymbol{S}$ các từ có tính kế tiếp.

Thuật toán 2.1. *Xây dựng biểu diễn bộ 4 ngữ nghĩa của* $\boldsymbol{S}$.

Xét danh sách các từ ngôn ngữ $\boldsymbol{S} = s_0, s_1, \ldots, s_m$ thỏa điều kiện (2.4) với thứ tự thỏa mãn $s_i < s_j$ nếu và chỉ nếu $i < j$.

Với l là một số nguyên dương thỏa điều kiện trong Mệnh đề 2.2, xét danh sách $\boldsymbol{I}_l = \{I_l(x_j) : x_j \in \boldsymbol{X}_l, j = 0, 1, \ldots, n\}$ gồm các khoảng tính mờ mức l thỏa điều kiện $I_l(x_j) < I_l(x'_j)$ khi và chỉ khi $j < j'$. Kí hiệu khoảng lân cận liền kề mức l là $\mathcal{S}_l(s), s \in \boldsymbol{S}$.

Xét vị từ $Incl(I_l(x_j), \mathcal{S}_l(s_i))$ nói rằng khoảng $I_l(x_j)$ chứa trong khoảng $\mathcal{S}_l(s_i)$.

Bước 1: Xuất phát từ $s_0 (i := 0)$ và từ khoảng tính mờ đầu tiên $\mathcal{J}_l(xj)(j := 0)$ trong danh sách $\boldsymbol{I}_l$. Lưu giá trị chỉ số 0 vào biến *Index*;

- Kiểm tra điều kiện $Incl(\mathcal{J}_l(x_j), \mathcal{S}_l(s_0))$ lần lượt đối với các $\mathcal{J}_l(x_j)$ cho đến khi tìm được j đầu tiên thỏa điều kiện $Connect(\mathcal{J}_l(x_j), \mathcal{S}_l(s_0)) =$ "true". Đặt $Temp_l(s_0) := \mathcal{S}_l(s_0) \cup \mathbf{Y}_{Index \leq k \leq j} \boldsymbol{I}_l(x_k)$. Thay j tìm được vào biến *Index* và tăng chỉ số j thêm $1, j := j + 1$.

Bước 2: Tăng $i := i + 1$.

- Kiểm tra điều kiện $Incl(\mathcal{J}_l(x_j), \mathcal{S}_l(s_i))$ lần lượt tiếp các $\mathcal{J}_l(x_j)$ cho đến khi tìm được j thỏa điều kiện $Incl(\mathcal{J}_l(x_j), \mathcal{S}_l(s_i)) =$ "true". Đặt $Center = round((j + Index)/2)$ và đặt $\boldsymbol{S}(s_{i-1}) := Temp_l(s_{i-1}) \cup \mathbf{Y}_{Index < k \leq Center} I_l(x_k)$ và đưa $\boldsymbol{S}(s_{i-1})$ vào danh sách NEIGH-SYS. Đặt $Temp_l(s_i) := \mathcal{S}_l(s_i) \cup \mathbf{Y}_{Center < k \leq j} \mathcal{J}_l(x_k)$. Thay j tìm được vào biến *Index* và tăng chỉ số j thêm $1 : j := j + 1$.

- Nếu $i < m$, quay về Bước 2.

Bước 3: Đặt $\boldsymbol{S}(s_m) := \mathcal{S}_l(s_m) \mathbf{Y}_{Index < k \leq n} I_l(x_k)$ và đưa vào danh sách NEIGH-SYS.

- Với mỗi s_i, tính $\upsilon(s_i)$ và thành lập tập bộ 4 ngữ nghĩa $(s_i, \upsilon(s_i), r, \boldsymbol{S}(s_i))$, với $r \in \boldsymbol{S}(s_i) \in$ NEIGH-SYS đã được xác định trong Bước 2, và đưa vào tập OUTPUT.

- Xuất danh sách kết quả OUTPUT và kết thúc thuật toán.

Định lý 2.1. . *Với l thỏa mãn Mệnh đề 2.2, thuật toán luôn luôn dừng, độ phức tạp tuyến tính và kết suất (ouput) của Thuật toán 2.1 là một phân hoạch của đoạn* $[0, 1]$. *Do đó, tập biểu diễn bộ 4* $\{(s, \upsilon(s), r, \boldsymbol{S}(s)) : s \in \boldsymbol{S}, r \in \boldsymbol{S}(s)\}$ *của* $\boldsymbol{S}$ *là thang đánh giá ngôn ngữ dựa trên bộ 4 ngữ nghĩa.*

Chứng minh. Theo giả thiết của l, các giá trị $\upsilon(s_i)$, $i = 0, 1, \ldots, m$, đều là các đầu mút của các khoảng tính mờ mức l trong $\boldsymbol{I}_l = \{\mathcal{J}_l(x_j) : x_j \in \boldsymbol{X}_l, j = 0, 1, \ldots, n\}$. Chú ý rằng họ $\boldsymbol{I}_l$ là một phân hoạch của đoạn $[0, 1]$ và, do vậy, luôn tôn tại khoảng $\mathcal{J}_l(x_j)$ thỏa điều kiện $Connect(\mathcal{J}_l(x_j), \mathcal{S}_l(s_i)) =$ "true". Theo Định lý 2.2, giữa hai giá trị liên tiếp bất kỳ, $\upsilon(s_i)$ và $\upsilon(s_{i+1})$, đều có ít nhất 2 khoảng tính mờ của $\boldsymbol{I}_l$. Một nửa số các khoảng đó được sử dụng

để xây dựng khoảng lân cận $\boldsymbol{S}(s_i)$ và nửa còn lại để xây dựng $\boldsymbol{S}(s_{i+1})$. Lưu ý rằng tất cả các khoảng tính mờ mức l mằm bên trái $\upsilon(s_0)$ được sử dụng để xây dựng $\boldsymbol{S}(s_0)$) và tất cả các khoảng tính mờ mức l mằm bên phải $\upsilon(s_m)$ được sử dụng để xây dựng $\boldsymbol{S}(s_m)$. Từ đó suy ra tập $\{\boldsymbol{S}(s_i) : i = 0, 1, \ldots, m\}$ tạo thành một phân hoạch của đoạn $[0, 1]$.

Theo cách xây dựng, luôn luôn có 2 khoảng tính mờ mức l kề bên trái và kề bên phải của $\upsilon(s_i)$ nằm trong $\mathcal{S}_l(s_i)$ và do đó $\upsilon(s_i) \in \boldsymbol{S}(s_i), i = 0, 1, \ldots, m$.

Ta thấy rằng thuật toán chỉ duyệt một lần danh sách $\boldsymbol{I}_l$ và do đó dễ dàng chứng minh các khẳng định còn lại trong phát biểu định lý trên. ■

3. CÁC PHÉP KẾT NHẬP TRÊN CÁC BỘ 4 NGỮ NGHĨA

3.1. Bài toán kết nhập và phép kết nhập

Trong thực tế có rất nhiều vấn đề dẫn đến bài toán tìm ý kiến đánh giá làm đại diện chung cho ý kiến của các chuyên gia về một vấn đề nào đó. Việc tổng hợp và kết hợp các ý kiến riêng của từng chuyên gia được gọi là *việc kết nhập* (aggregation) và phép tính thực hiện việc kết nhập các ý kiến riêng biệt được gọi là *phép kết nhập* (aggregation operator). Trong cách tiếp cận tính toán trên từ, thang đánh giá là thang các từ ngôn ngữ và phép tính kết nhập cũng phải cho kết quả là một từ trong thang điểm.

Tương tự như trong [4,5,7], chúng ta sẽ mở rộng các phép kết nhập thông thường trên số thực "đại diện" ngữ nghĩa định lượng trong biểu diễn bộ 4 ngữ nghĩa của các từ trong $\boldsymbol{S}$.

Trong phần này ta xét một thang điểm ngôn ngữ dựa trên bộ 4 ngữ nghĩa của $\boldsymbol{S} = s_0, s_1, \ldots, s_m$, $T_S = \{(s_i, \upsilon(s_i), r_i, \boldsymbol{S}(s_i)) : r_i \in \boldsymbol{S}(s_i), i = 0, 1, \ldots, m\}$.

Tương tự như trong [5], ta định nghĩa ánh xạ sau:

$$\Delta : [0, 1] \to \boldsymbol{T}_S, \Delta(r) = (s_i, \upsilon(s_i), r, \boldsymbol{S}(s_i)), r \in \boldsymbol{S}(s_i), \text{ với } r \in [0, 1] \tag{3.7}$$

Định nghĩa ánh xạ (3.1) là chỉnh vì họ $\{\boldsymbol{S}(s_i) : i = 0, 1, \ldots, m\}$ là một phân hoạch của [0,1] nên chỉ tồn tại duy nhất một từ s_i thỏa mãn (3.1).

Ánh xạ ngược $\Delta^{\breve{}1}$ được xác định bởi công thức

$$\Delta^{-1}((s_i, \upsilon(s_i), r, \boldsymbol{S}(s_i))) = r, r \in [0, 1] \tag{3.8}$$

Định nghĩa 3.1. Cho g là một phép kết nhập $p-$ngôi trên các số thực của đoạn $[0, 1]$, $g : [0, 1]^p \to [0, 1]$. Phép $p-$ngôi $g*, g* : (\boldsymbol{T}_S)^p \to \boldsymbol{T}_S$, được gọi là phép kết nhập mở rộng của phép kết nhập thông thường g sang miền các bộ 4 chứa các từ ngôn ngữ nếu nó được định nghĩa như sau:

Với vectơ dữ liệu $\langle(s_{ik}, \upsilon(s_{ik}), r_{ik}, \boldsymbol{S}(s_{ik})) : k = 1, \ldots, p\rangle$,

$$g^*((s_{i1}, \upsilon(s_{i1}), r_{i1}, \boldsymbol{S}(s_{i1})), \ldots, (s_{ip}, \upsilon(s_{ip}), r_{ip}, \boldsymbol{S}(s_{ip}))) = \Delta(g(r_{i1}, \ldots, g(r_{ip})) \tag{3.9}$$

Định nghĩa phép kết nhập trên từ dựa trên biểu diễn bộ 4 ngữ nghĩa của chúng mang nhiều thông tin ngữ nghĩa của từ hơn so với việc thực hiện phép kết nhập trên các chỉ số vì những lý do sau:

(i) Các khoảng lân cận $\boldsymbol{S}(s_i)$ của các từ mang ngữ nghĩa định lượng của các từ với mức tính mờ l. Chúng được xác định dựa trên độ đo tính mờ của các từ ngôn ngữ. Vì vậy, giá trị thực r_i cũng mang nhiều thông tin ngữ nghĩa của từ nhiều hơn các chỉ số của chúng;

(ii) Theo định nghĩa của bộ 4 ngữ nghĩa, giá trị ngữ nghĩa định lượng $\upsilon(s_i)$ của s_i có thể được xem là giá trị thực mang thông tin ngữ nghĩa phù hợp với si nhất và hiệu số $(r_{i^*} - \upsilon(s_{i^*}))$ cho ta thông tin về độ lệch của việc chuyển đổi giá trị thực về giá trị ngôn ngữ.

Sau đây ta chỉ ra một số phép kết nhập thông dụng.

1) *Phép trung bình cộng số học*

Cho các bộ 4 ngữ nghĩa trên thang điểm $\boldsymbol{T}_S$, $(s_{ik}, \upsilon(s_{ik}), r_{ik}, \boldsymbol{S}(s_{ik})), k = 1, \ldots, p$. Theo Định nghĩa 3.1, phép kết nhập trung bình số học

$$g_{arith}(a_1, \ldots, a_p) = \frac{1}{p} \sum_{1 \le k \le p} a_k$$

cảm sinh một phép kết nhập trên các từ với biểu diễn bộ 4 ngữ nghĩa sau:

$$g^*_{arith}((s_{i1}, \upsilon(s_{i1}), r_{i1}, \boldsymbol{S}(s_{i1})), \ldots, (s_{ip}, \upsilon(s_{ip}), r_{ip}, \boldsymbol{S}(s_{ip}))) = \Delta(\frac{1}{p} \sum_{1 \le k \le p} r_{ik}) \tag{3.10}$$

nghĩa là giá trị của g_{arith} là bộ 4 $(s_{i^*}, \upsilon(s_{i^*}), \frac{1}{p}\sum_{1 \le k \le p} r_{ik}), \boldsymbol{S}(s_{i^*}))$ ứng với i^* sao cho

$$\frac{1}{p} \sum_{1 \le k \le p} r_{ik} \in \boldsymbol{S}(s_{i^*}).$$

. 2) *Phép trung bình cộng có trọng số*

Cho các bộ 4 ngữ nghĩa trên thang điểm $\boldsymbol{T}_S, (s_{ik}, \upsilon(s_{ik}), r_{ik}, \boldsymbol{S}(s_{ik}))$, $k = 1, \ldots, p$. Phép trung bình có trọng số thông thường

$$g_{weight}(a_1, \ldots, a_p) = \sum_{1 \le k \le p} w_k a_k, w_k \ge 0 \text{ và } \sum_{1 \le k \le p} w_k = 1.$$

cảm sinh một phép kết nhập trung bình có trọng số g^*_{weight} sau:

$$g^*_{weight}((s_{i1}, \upsilon(s_{i1}), r_{i1}, \boldsymbol{S}(s_{i1})), \ldots, (s_{ip}, \upsilon(s_{ip}), r_{ip}, \boldsymbol{S}(s_{ip}))) = \Delta(\sum_{1 \le k \le p} w_k r_{ik})$$

nghĩa là giá trị của g^*_{weight} là bộ 4 $(s_{i^*}, \upsilon(s_{i^*}), \sum_{1 \le k \le p} w_k r_{ik}, \boldsymbol{S}(s_{i^*}))$ ứng với i^* sao cho

$$\sum_{1 \le k \le p} w_k r_{ik} \in \boldsymbol{S}(s_{i^*}).$$

3) *Phép kết nhập trọng số có thứ tự (ordered weighted aggregation operator)*

Cho các bộ 4 ngữ nghĩa trên thang điểm $\boldsymbol{T}_S, (s_{ik}, \upsilon(s_{ik}), r_{ik}, \boldsymbol{S}(s_{ik}))$, $k = 1, \ldots, p$. Phép kết nhập trọng số có thứ tự thông thường

$$g_{Or\text{-}weight}(a_1, \ldots, a_p) = \sum_{1 \le k \le p} w_k a^{\ge}_k, w_k \ge 0 \text{ và } \sum_{1 \le k \le p} w_k = 1,$$

trong đó vectơ $(a^{\ge}_1, \ldots, a^{\ge}_k)$ thu được từ việc sắp xếp dãy $a_1, \ldots, a_p$ theo thứ tự giảm dần, sẽ cảm sinh một phép kết nhập trọng số có thứ tự g^*_{weight} sau:

$$g^*_{Or\text{-}weight}((s_{i1}, \upsilon(s_{i1}), r_{i1}, \boldsymbol{S}(s_{i1})), \ldots, (s_{ip}, \upsilon(s_{ip}), r_{ip}, \boldsymbol{S}(s_{ip}))) = \Delta(\sum_{1 \le k \le p} w_k a^{\ge}_k)$$

nghĩa là giá trị của g^*_{weight} là bộ 4 $(s_{i^*}, \upsilon(s_{i^*}), \sum_{1\le k\le p} w_k a_k^{\geq}, \boldsymbol{S}(s_{i^*}))$ ứng với i^* sao cho

$$\sum_{1\le k\le p} w_k a_k^{\geq} \in \boldsymbol{S}(s_{i^*}).$$

3.2. Ví dụ minh họa mô tả các phương pháp tiếp cận

Ta hãy lấy ví dụ việc giải bài toán quyết định ngôn ngữ (Linguistic decision problem) để so sánh giữa hai phương pháp tiếp cận.

Trong [4,5] đưa ra ví dụ về bài toán ứng dụng như sau. Giả sử một công ty phân phối muốn ký hợp đồng với một công ty tư vấn làm một báo cáo phân tích tổng quan về các khả năng cung cấp các giải pháp xây dựng hệ máy tính hiện có trên thị trường để quyết định phương án phù hợp nhất cho nhu cầu của công ty. Những phương án lựa chọn bao gồm:

x_1	x_2	x_3	x_4
Unix	Windows-NT	AS/400	VMS

Công ty tư vấn lập một nhóm gồm 4 phòng thực hiện tư vấn này, bao gồm

p_1	p_2	p_3	p_4
Cost analysis	System analysis	Risk analysis	Technology analysis

Mỗi phòng cung cấp ý kiến đánh giá của mình đối với từng phương án (**Al**ternatives) bằng một vectơ đánh giá dựa trên ý kiến chuyên gia (**Ex**pert) biểu thị qua thang điểm ngôn ngữ $\boldsymbol{S} = \boldsymbol{X}_{(2)} = \{EL, VL, L, M, H, VH, EH\}$, trong đó các ký hiệu trong $\boldsymbol{S}$ là các chữ viết tắt của "Extremely Low", "Very Low", "Low", "Medium", "High", "Very High" và "Extremely High". Các ý kiến được cho trong bảng sau:

Al / **Ex**	x_1	x_2	x_3	x_4
p_1	V_l	$\boldsymbol{m}$	$\boldsymbol{m}$	$\boldsymbol{l}$
p_2	$\boldsymbol{m}$	$\boldsymbol{l}$	V_l	$\boldsymbol{h}$
p_3	$\boldsymbol{h}$	V_l	$\boldsymbol{m}$	$\boldsymbol{m}$
p_4	$\boldsymbol{h}$	$\boldsymbol{h}$	$\boldsymbol{l}$	$\boldsymbol{l}$

Chúng tôi thực hiện giải bài toán này với phương các phương pháp trình bầy ở trên và so sánh kết quả trong bảng 3.1. Dưới đây chúng tôi mô tả chi tiết phương pháp giải dựa trên biểu diễn các từ bằng bộ 4 ngữ nghĩa.

So với ví dụ trong [4], ở đây chúng ta đã thay từ N (Negative) bằng từ EL (Extremely Low) và P (Positive) bằng từ EH (Extremely High). Đối với 3 phương pháp trên, việc thay đổi này không ảnh hưởng đến bản chất của phương pháp luận vì thang điểm ngôn ngữ chỉ là một dãy các ký hiệu ngôn ngữ. Ý nghĩa các từ chỉ có tính gợi nhớ và thân thiện với người dùng. Tuy nhiên, theo cách tiếp cận dựa trên ngữ nghĩa của ĐSGT, ngữ nghĩa của các ký hiệu lại quan trọng. Việc thay đổi như trên để bảo đảm các từ sẽ nằm trong tập nền của một ĐSGT và, tất nhiên, việc chọn này sẽ ảnh hưởng đến kết quả tính toán.

Giả sử các từ trong $\boldsymbol{S} = \{E\boldsymbol{l}, V\boldsymbol{l}, \boldsymbol{l}, \boldsymbol{m}, \boldsymbol{h}, V\boldsymbol{h}, E\boldsymbol{h}\}$, trong đó $\boldsymbol{l}$:= Low, $\boldsymbol{h}$:= High và $\boldsymbol{m}$:= Medium, có ngữ nghĩa định lượng trên không gian tham chiếu $[0, 1]$. $\boldsymbol{S}$ là một tập con

trong một DSGT $\mathbf{AX} = (\boldsymbol{X}, \boldsymbol{G}, \boldsymbol{C}, \boldsymbol{H}, \leq)$, với $G = \{\boldsymbol{l}, \boldsymbol{h}\}$, $H^{\breve{}} = \{R, Lt\}$, trong đó R là viết tắt của từ "Rather" và Lt là viết tắt của từ "Little", $\boldsymbol{H}^{+} = \{V, E\}$.

Giả sử độ đo tính mờ của các từ được cho như sau: $fm(l) = 0.5 = fm(h)$, $\mu(R) = \mu(E) = 0.24$; $\mu(V) = \mu(Lt) = 0.26$. Theo Mệnh đề 2.1, ta có biểu diễn bộ 4 với mức ngữ nghĩa $l = 3$, nghĩa là việc tính các lân cận ngữ nghĩa của $\boldsymbol{S}$ dựa trên các khoảng tính mờ của các từ $\boldsymbol{X}_3$. Kết quả tính toán thu được như sau:

$$\boldsymbol{S}_3(E\boldsymbol{l}) = [0, (0.0288 + 0.0312) * 2) = [0, 0.12); \qquad \upsilon(E\boldsymbol{l}) = 0.06$$
$$\boldsymbol{S}_3(V\boldsymbol{l}) = [0.12, 0.0338 + 0.0312 * 2) = [0.12, 0.2162); \qquad \upsilon(V\boldsymbol{l}) = 0.185$$
$$\boldsymbol{S}_3(\boldsymbol{l}) = [0.2162, 0.0338 + 0.12) = [0.2114, 0.37); \qquad \upsilon(\boldsymbol{l}) = 0.25$$
$$\boldsymbol{S}_3(\boldsymbol{m}) = [0.37, 0.13 * 2) = [0.37, 0.63); \qquad \upsilon(\boldsymbol{m}) = 0.5$$
$$\boldsymbol{S}_3(\boldsymbol{h}) = [0.63, 0.12 + 0.0338) = [0.63, 0.7838); \qquad \upsilon(\boldsymbol{h}) = 0.75$$
$$\boldsymbol{S}_3(V\boldsymbol{h}) = [0.7838, 0.0312 * 2 + 0.0338) = [0.7838, 0.88); \qquad \upsilon(V\boldsymbol{h}) = 0.8488$$
$$\boldsymbol{S}_3(E\boldsymbol{h}) = [0.88, (0.0312 + 0.0288) * 2) = [0.88, 1.0]; \qquad \upsilon(E\boldsymbol{h}) = 0.94$$

Khi đó ta hoàn toàn chỉ ra được các giá trị bộ 4 của thang đánh giá của $\boldsymbol{S}$. Chẳng hạn ta có các giá trị: $(E\boldsymbol{l}, 0.06, r, [0.0, 0.12))$, $r \in [0.0, 0.12)$; $(V\boldsymbol{l}, 0.185, r, [0.12, 0.2162))$, $r \in [0.12, 0.2162)$; $(V\boldsymbol{h}, 0.8488, r, [0.7838, 0.88))$, $r \in [0.7838, 0.88)$ ...

Để so sánh, tương tự như đối với trường hợp biểu diễn bằng bộ 2 trong [5,8], ta sử dụng phép kết nhập trung bình số học:

Đối với phương án $x_1 : (0.185 + 0.5 + 0.75 + 0.75)/4 = 2.185/4 = 0.54625 \in \boldsymbol{S}_3(\boldsymbol{m})$

Đối với phương án $x_2 : (0.5 + 0.25 + 0.185 + 0.75)/4 = 1.685/4 = 0.42125 \in \boldsymbol{S}_3(\boldsymbol{m})$;

Đối với phương án $x_3 : (0.5 + 0.185 + 0.5 + 0.25)/4 = 1.435/4 = 0.35875 \in \boldsymbol{S}_3(\boldsymbol{m})$;

Đối với phương án $x_4 : (0.25 + 0.75 + 0.5 + 0.25)/4 = 1.75/4 = 0.4375 \in \boldsymbol{S}_3(\boldsymbol{m})$;

Theo định nghĩa thứ tự trên các bộ 4 ngữ nghĩa, x_1 là phương án được lựa chọn. Như vậy, phương án x_1 có điểm ngôn ngữ biểu thị bằng bộ 4 ngữ nghĩa $(\boldsymbol{m}, 0.5, 0.54625, [0.37, 0.63))$ là cao nhất và trùng với kết quả lựa chọn của phương pháp biểu diễn bằng bộ 2 được trình bày ở điểm 3).

Bảng 3.1 So sách các kết quả lời giải của 4 phương pháp, các ô chữ đậm là các phương án được lựa chọn.

Phương án	Phương pháp biểu diễn ngữ nghĩa các từ bằng tập mờ	Phương pháp dựa trên ký hiệu ngôn ngữ trong thang điểm	Phương pháp dựa trên biểu diễn diễn các từ bằng bộ 2	Phương pháp giải dựa trên biểu diễn các từ bằng bộ 4 ngữ nghĩa.
x_1	$\boldsymbol{app}_1(\boldsymbol{C}_{x1}) = \mathbf{m}$	**m**	**(m, 0.0)**	**(m, 0.5, 0.54625, [0.37, 0.63))**
x_2	$\boldsymbol{app}_1(\boldsymbol{C}_{x2}) = \mathbf{m}$	**m**	(m, –0.5)	(m, 0.5, 0.42125, [0.37, 0.63))
x_3	$app_1(C_{x3}) = \mathrm{l}$	l	(l, 0.25)	(l, 0.5, 0.54625, [0.2114, 0.37))
x_4	$\boldsymbol{app}_1(\boldsymbol{C}_{x4}) = \mathbf{m}$	**m**	(m, –0.25)	(m, 0.5, 0.4375, [0.37, 0.63))

Phương pháp biểu diễn bộ 4 và bộ 2 cho phép chọn một phương án tốt nhất trong khi đó phương pháp biểu diễn ngữ nghĩa các từ bằng tập mờ và bằng kí hiệu ngôn ngữ trong thang điểm lại cho phép chọn cả 3 phương án.

4. KẾT LUẬN

Bài toán lấy quyết định với thông tin mờ dẫn đến các cách tiếp cận khác nhau trong việc giải bài toán kết nhập các ý kiến đánh giá dựa trên thang điểm đánh giá ngôn ngữ. Các

phương pháp dựa trên biểu diễn ngữ nghĩa bằng tập mờ thường mất mát thông tin trong việc biểu diễn ngữ nghĩa các từ ngôn ngữ và trong tính toán trên các từ. Chẳng hạn việc tính toán trên các chỉ số của thang điểm ngôn ngữ đã được sắp tuyến tính là không tự nhiên, vì các chỉ số biểu thị quá ít thông tin về ngữ nghĩa của các từ ngôn ngữ. ĐSGT cung cấp cho ta công cụ biểu diễn thông tin ngôn ngữ với nhiều ưu điểm đáng kể: (i) Cấu trúc ĐSGT phản ánh khá trung thành ngữ nghĩa các từ ngôn ngữ; (ii) Có tồn tại mối quan hệ chặt chẽ và phong phú giữa ngữ nghĩa định tính và định lượng; (iii) Việc biểu diễn ngữ nghĩa nhờ bộ 4 ngữ nghĩa mang nhiều thông tin. Với những ưu điểm như vậy phương pháp mới để biểu diễn và kết nhập các thông tin ngôn ngữ được đề xuất trong bài báo sẽ bao hàm được những thông tin ngữ nghĩa hữu ích trong việc giải các bài toán ứng dụng.

Như ta đã phân tích ở trên, phương pháp mới chắt lọc được nhiều thông tin ngữ nghĩa của từ ngôn ngữ. Ngoài ra, phạm vi ứng dụng của phương pháp rộng và linh hoạt hơn vì nó tạo ra khả năng xây dựng phương pháp đánh giá cho phép vừa cho điểm ngôn ngữ, vừa cho điểm thực. Để so sánh chứng tỏ hiệu năng của phương pháp, sử dụng phương pháp được đề xuất để giải bài toán đã được giải bằng các phương pháp khác. Kết quả chỉ ra rằng phương pháp mới cho kết quả tốt và dễ sử dụng do tính mềm dẻo của nó (cho phép đánh giá bằng điểm thực hoặc ngôn ngữ, tính toán kết nhập trên các giá trị ngữ nghĩa định lượng, nghĩa là giá trị thực, ...).

TÀI LIỆU THAM KHẢO

[1] M. Delgado, F. Herrera, E. Herrera-Viedma, L. Marffnez, Combining numerical and linguistic information in group decision making, *Journal of Information Sciences* **107** (1998) 177-194.

[2] F. Herreraa, E. Herrera-Viedmaa, Luis Martnez, A fusion approach for managing multi-granularity linguistic term sets in decision making, *Fuzzy Sets and Systems* **114** (2000) 43-58.

[3] F. Herrera, E. Herrera-Viedma, Linguistic decision analysis: steps for solving decision problems under linguistic information, *Fuzzy Sets and Systems* **115** (2000) 67-82.

[4] F. Herrera and L. Martinez, A 2-Tuple Fuzzy Linguistic Reoresentation Model for Computing with Words, *IEEE Transactions on Systems, MAN, and Cybernetics* **8** (6) (2000), 746-752.

[5] F. Herrera and L. Martinez, A Model Based on Linguistic 2-Tuples for Dealing with Multigranular Hierachical Linguistic Contexts in Multi-Expert Decision-Making, *IEEE Transactions on Systems, MAN, and Cybernetics* **31** (2) (2001), 227-234.

[6] N.C. Ho, N.V. Long, Fuzziness Measure on Complete Hedge Algebras and Quantitative Semantics of Terms in Linear Hedge Algebras, *Fuzzy Sets and Systems* **158** (4) (2007) 452-471.

[7] Van Hung Le, Cat Ho Nguyen, Fei Liu, Semantics and Aggregation of Linguistic Information Based on Hedge Algebras, *The Third International Conference on Knowledge, Information and Creativity Support Systems*, KICSS 2008, Hanoi, Vietnam, Dec. 22-23, 2008 (128-135).

[8] Van Nam Huynh, Cat Ho Nguyen, Yoshiteru Nakamori, MEDM in General Multi-granular Hierarchical Linguistic Contexts Based on The 2-Tuples Linguistic Model, *Proc. of IEEE Int. Conf. on Granular Computing*, 2005 (482–487).

[9] Jun Liu, Da Ruan, Roland Carchon, Synthesis and Evaluation Analysis of the Indicator Information in Nuclear Safeguards Applications by Computing Withwords, *Int. J. Appl. Math. Comput. Sci.* **12** (3) (2002) 449–462.

[10] Nguyễn văn Long, Cơ sở phương pháp luận của phương pháp đánh giá bằng nhãn ngôn ngữ, *Tạp chí Tin học và Điều khiển học* **24**(1) (2008) 75-86.

Ngày nhận bài 26 - 7 - 2010
Nhận lại sau sửa 14 - 11 - 2011