Tạp chí Khoa học – Đại học Huế
ISSN 1859-1388
Tập 106, Số 07, 2015, Tr. 241-251

MỘT PHƯƠNG PHÁP XỬ LÝ KHÔNG NHẤT QUÁN TRI THỨC Ở MỨC CÚ PHÁP

Nguyễn Văn Trung[1], Hoàng Hữu Hạnh[2,*]

[1]Trường Đại học Khoa học - Đại học Huế
[2]Đại học Huế

Tóm tắt. Có thể phân các phương pháp xử lý sự không nhất quán *tri thức dựa trên logic* thành hai mức: mức cú pháp và mức ngữ nghĩa. Bài báo trình bày một phương pháp dựa trên lý thuyết đồng thuận để xử lý sự không nhất quán tri thức ở mức cú pháp, trong đó trạng thái tri thức được biểu diễn bằng một biểu thức dạng hội của các *literal* .

Từ khóa: không nhất quán, tri thức, đồng thuận, mức cú pháp

1 Giới thiệu

Người ta phân các phương pháp xử lý sự không nhất quán của tri thức dựa trên hình thức logic (logic-based knowledge) thành hai mức [5]: mức cú pháp (syntactic level) và mức ngữ nghĩa (semantic). Ở mức cú pháp, trạng thái tri thức của các chuyên gia hoặc các tác tử (agent) - được biểu diễn dưới dạng một biểu thức logic - không nhất quán với nhau khi cú pháp của các công thức là khác nhau; Ở mức ngữ nghĩa, sự không nhất quán còn được xét đến ở cả các diễn dịch (intepretion) của các biểu thức logic.

Bài báo này sẽ trình bày một phương pháp giải quyết sự không nhất quán tri thức ở mức cú pháp dựa theo phương pháp đồng thuận [1]. Một trong những yếu tố mấu chốt để giải quyết sự không nhất quán ở mức cú pháp là đánh giá mức độ sai khác (còn gọi là khoảng cách) giữa hai biểu thức logic. Về một mặt nào đó, nếu biết được độ tương tự, hay mức độ liên quan giữa hai biểu thức logic, người ta cũng suy ra được khoảng cách giữa chúng. Có nhiều phương án đánh giá các độ đo như vậy (độ sai khác, hoặc độ tương tự giữa hai biểu thức logic). Zhisheng Huang, Frank van Harmelen [3] đánh giá mức độ liên quan giữa hai biểu thức bằng cách xét các đối tượng chung và riêng giữa hai biểu thức này. Ferilli và cộng sự [2] bên cạnh việc đánh giá sự xuất hiện của các đối tượng, vị từ xuất hiện trong hai biểu thức còn xét đến mức độ liên quan giữa chúng dựa trên một cơ sở tri thức nền dạng cây phân cấp từ loại kiểu như Word Net. Với giả thiết, trạng thái tri thức được biểu diễn dưới dạng một biểu thức hội hoặc tuyển của các literal , tác giả Nguyễn Ngọc Thành [5] xây dựng độ đo khoảng cách giữa hai biểu thức logic dựa trên phép hiệu đối xứng và phép hợp của các tập hợp literal này.

Trong mô hình xử lý sự không nhất quán bằng phương pháp đồng thuận [4, 5], bên cạnh hàm khoảng cách, chúng ta cần xác định các tiêu chuẩn cho hàm chọn đồng thuận và mối liên

**Liên hệ: hhhanh@hueuni.edu.vn*

Nhận bài: 9-5-2015; Hoàn thành phản biện: 9-6-2015; Ngày nhận đăng: 31-8-2015.

hệ giữa các tiêu chuẩn này. Các tiêu chuẩn cũng có bị ảnh hưởng bởi việc chọn hàm khoảng cách.

Phạm vi bài báo này chỉ đề cập đến việc xử lý sự không nhất quán tri thức mức cú pháp, trong đó một trạng thái tri thức được biểu diễn dưới dạng hội của các literal . Cấu trúc của bài báo là như sau: Phần 2 trình bày một số khái niệm cơ bản. Các khái niệm này sẽ được dùng trong Phần 3 để hình thức hoá cho bài toán tìm đồng thuận từ hồ sơ xung đột gồm các cấu trúc hội của các literal . Phần 4 sẽ phân tích và chứng minh một số tính chất, mối liên hệ giữa các tiêu chuẩn cho bài toán chọn đồng thuận. Trên cơ sở các phân tích này, chúng tôi sẽ đưa ra thuật toán tìm đồng thuận ở Phần 5. Bài báo sẽ kết thúc với một số nhận xét về các kết quả quan trọng và hướng mở rộng.

2 Một số khái niệm cơ bản

Trong phần này, chúng tôi trình bày một số khái niệm cơ bản được dùng trực tiếp cho bài toán tìm đồng thuận của hồ sơ xung đột các cấu trúc hội. Các khái niệm này được dựa theo công trình [5] của tác giả Nguyễn Ngọc Thành. Giả sử rằng để biểu diễn tri thức, một tác tử (agent) hay một chuyên gia sử dụng một tập hợp hữu hạn **L** các ký hiệu để biểu diễn các giá trị logic khẳng định, tham chiếu đến các sự kiện và đối tượng cụ thể trong một thế giới thực. Mỗi ký hiệu biểu diễn cho một phần tử riêng biệt trong thế giới thực. Một literal tồn tại ở một trong hai dạng literal dương hoặc literal âm, trong đó literal dương là một ký hiệu lấy từ **L**, chẳng hạn a,b,… và literal âm là một ký hiệu lấy từ **L** theo sau bởi một ký hiệu phủ định “ $\neg$ ”, chẳng hạn, $\neg a, \neg b, \ldots$

Một tác tử có thể đưa ra ý kiến của nó về một chủ đề trong thế giới thực dưới dạng một cấu trúc hội của các literal như sau: $t_1 \wedge t_2 \wedge \ldots \wedge t_k$, trong đó $t_i \in \mathbf{L}$ hoặc $t_i = \neg t_i'$, với $t_i' \in \mathbf{L}, i = 1,2,\ldots,k$ và $t_i \neq t_j$ với $i \neq j; i,j = 1,2,\ldots,k$. Điều này được hiểu, theo ý kiến của tác tử này, các dữ kiện hay sự kiện được biểu diễn bởi các literal dương trong các literal $t_1, t_2, \ldots, t_k$ phải được diễn ra, và các dữ kiện hay sự kiện được biểu diễn bởi các literal âm phải không được diễn ra. Những dữ kiện hay sự kiện không được chứa trong các ý kiến của tác tử thì sẽ được bỏ qua: các tác tử hoặc các chuyên gia không được giả thiết là “biết tất cả mọi thứ”.

Ta ký hiệu $Conj(\mathbf{L})$ là tập hợp tất cả các cấu trúc hội với các ký hiệu lấy từ tập hợp **L**. Do **L** là hữu hạn, việc tham chiếu đến chúng cũng là bị giới hạn, nên tập hợp $Conj(\mathbf{L})$ cũng là hữu hạn.

Chúng ta biểu diễn một cấu trúc hội x bằng cặp (x^+, x^-), trong đó x^+ (tương ứng, x^-) gồm các ký hiệu trong cấu trúc hội thuộc về các literal dương (tương ứng, các literal âm). Ví dụ, biểu diễn của cấu trúc hội $x = a \wedge \neg b \wedge c$ với $a,b,c \in \mathbf{L}$ là cặp (x^+, x^-) với $x^+ = \{a,c\}$ và $x^- = \{b\}$.

Định nghĩa 2.1 Cấu trúc hội không xung đột [5]

Một cấu trúc hội (x^+, x^-) với $x^+, x^- \subseteq \mathbf{L}$ là không xung đột nếu $x^+ \cap x^- = \varnothing$.

Từ đây, ta ngầm định gọi ngắn gọn một cấu trúc hội là một cấu trúc hội không xung đột.

Định nghĩa 2.2. Hai cấu trúc hội không nhất quán, không nhất quán nghiêm ngặt [5]

Gọi $x = (x^+, x^-), x' = (x'^+, x'^-) \in Conj(\mathbf{L})$ là các cấu trúc hội không xung đột. Ta nói

a. Cấu trúc hội x là không nhất quán với cấu trúc hội x' nếu $x^+ \cap x'^- \neq \emptyset$ ho?c $x'^+ \cap x^- \neq \emptyset,$

b. Cấu trúc hội x được gọi là không nhất quán nghiêm ngặt với cấu trúc hội x' nếu chúng là không nhất quán và $x^+ \cap x'^+ = \emptyset$ và $x^- \cap x'^- = \emptyset,$

Định nghĩa 2.3. Tập cấu trúc hội không xung đột [5]

Một tập các cấu trúc hội không xung đột $\mathbf{X} = \left\{ x_i = (x^+, x^-) \in Conj(\mathbf{L}) : i = 1, 2, \ldots, n \right\}$ được gọi là không nhất quán nếu $\bigcup_{x \in \mathbf{X}} x^+ \cap \bigcup_{x \in \mathbf{X}} x^- \neq \emptyset,$ ngược lại, $\mathbf{X}$ được gọi là nhất quán.

Tiếp đến chúng ta định nghĩa hàm $d_\wedge$ giữa các cấu trúc hội. Hàm này sẽ giúp ích trong việc đánh giá sự khác biệt giữa các cấu trúc hội. Nhờ hàm này, ta có thể biết được mức độ không nhất quán của một hồ sơ xung đột và qua đó, xác định được đồng thuận cho hồ sơ xung đột. Trước hết ta cần định nghĩa khoảng cách giữa các tập hữu hạn.

Định nghĩa 2.4. Khoảng cách giữa hai tập hữu hạn $\mathbf{X}_1, \mathbf{X}_2$ là $\eta(\mathbf{X}_1, \mathbf{X}_2) = \frac{card(\mathbf{X}_1 \Delta \mathbf{X}_2)}{card(\mathbf{L})}$, trong đó, $card(\mathbf{L})$ biểu diễn số lượng phần tử trong tập hợp **L**, và $\mathbf{X}_1 \Delta \mathbf{X}_2$ biểu diễn tập hợp hiệu đối xứng của hai tập hợp $\mathbf{X}_1$ và $\mathbf{X}_2$.

Có thể nhận thấy một cách trực quan rằng, khoảng cách η giữa hai tập hợp sẽ đạt cực tiểu (= 0) khi chúng bằng nhau, và tăng dần khi số lượng phần tử khác nhau giữa chúng tăng lên. Điều này phản ánh tính hợp lý của độ đo khoảng cách giữa hai tập hợp. Dựa trên hàm η chúng ta định nghĩa khoảng cách giữa hai cấu trúc hội x_1 và x_2 như sau.

Định nghĩa 2.5. Khoảng cách giữa hai cấu trúc hội [5]

Khoảng cách giữa hai cấu trúc hội $x_1, x_2 \in Conj(\mathbf{L})$ là

$$d_\wedge(x_1, x_2) = w_1.\eta(x_1^+, x_2^+) + w_2.\eta(x_1^-, x_2^-),$$

trong đó

- $\eta(x_1^+, x_2^+)$ là khoảng cách giữa các tập các ký hiệu dương trong cấu trúc hội x_1, x_2.
- $\eta(x_1^-, x_2^-)$ là khoảng cách giữa các tập các ký hiệu âm trong cấu trúc hội x_1, x_2.
- w_1, w_2 là các trọng số của khoảng cách $\eta(x_1^+, x_2^+)$ và $\eta(x_1^-, x_2^-)$ trong khoảng cách $d_\wedge(x_1, x_2)$,

tương ứng, thoả điều kiện: $w_1 + w_2 = 1 \text{ và } 0 < w_1, w_2 < 1.$

Trong một cấu trúc hội, các literal dương có thể được xem như là phần tri thức khẳng định của các tác tử, và các literal âm có thể được xem như là tri thức phủ định của chúng. Dùng các giá trị w_1, w_2, chúng ta có thể phân biệt các trọng số của các khoảng cách giữa các phần khẳng định và phủ định của các trạng thái tri thức của các tác tử. Trong bài báo này, chúng ta xem vai trò của tri thức khẳng định và tri thức phủ định là như nhau, theo đó $w_1 = w_2 = \frac{1}{2}$.

Chúng ta có thể thấy hàm $d_\wedge$ thoả các tính chất được quy định cho một hàm khoảng cách giữa các phần tử thuộc tập vũ trụ $Conj(\mathbf{L})$, đó là tính không âm, phản xạ và đối xứng [5]. Giá trị của $d_\wedge$ cũng được chuẩn hoá, nghĩa là, thuộc về khoảng $[0,1]$.

3 Bài toán tìm đồng thuận của các cấu trúc hội và các tiêu chuẩn cho đồng thuận

Trước khi phát biểu bài toán tìm đồng thuận của các cấu trúc hội, chúng ta xác định một số thuật ngữ thường được dùng với bài toán này.

Gọi $\mathbf{U}$ là một tập hữu hạn các đối tượng biểu diễn các ý kiến tiềm năng về chủ đề xung đột (tập $\mathbf{U}$ thường được gọi là tập vũ trụ). Ký hiệu $2^{\mathbf{U}}$ biểu diễn tập hợp gồm tất cả các tập con lập được từ các phần tử thuộc tập hợp $\mathbf{U}$.

Gọi $\prod_k(\mathbf{U})$ là tập hợp tất cả các tập con có k phần tử có thể lặp của tập hợp $\mathbf{U}$ với $k \in \mathrm{N}$ (N là tập hợp các số tự nhiên). Mỗi phần tử trong $\prod(\mathbf{U}) = \bigcup_{k\in \mathrm{N}} \prod_k(\mathbf{U})$ được gọi là một hồ sơ xung đột. Bài toán tìm đồng thuận của các cấu trúc hội được phát biểu như sau [5]:

Cho một hồ sơ xung đột gồm các cấu trúc hội

$$X := \{x_i = (x_i^+, x_i^-) \in Conj(\mathbf{L}) : i = 1,2,\ldots,n\}.$$

Cần xác định một cấu trúc hội $x^* \in Conj(\mathbf{L})$ là đại diện tốt nhất cho các cấu trúc hội thuộc hồ sơ X (còn gọi là đồng thuận của hồ sơ X).

Ở đây, hồ sơ X có thể được tách ra thành các 2 hồ sơ con: hồ sơ khẳng định $X^+ = \{x_i^+ : i = 1,2,\ldots,n\}$ và hồ sơ phủ định $X^- = \{x_i^- : i = 1,2,\ldots,n\}$. Bây giờ ta định nghĩa hàm đồng thuận sau đây cho các hồ sơ các cấu trúc hội.

Định nghĩa 3.1. Hàm đồng thuận của hồ sơ cấu trúc hội [5]

Hàm đồng thuận của hồ sơ cấu trúc hội được định nghĩa như sau:

$$C : \prod(Conj(\mathbf{L})) \to 2^{Conj(\mathbf{L})}.$$

thoả một hoặc nhiều tiêu chuẩn sau đây.

P1. Mỗi cấu trúc hội $(x^{*+}, x^{*-}) \in C(X)$ phải có (a) $\bigcap_{x\in X} x^+ \subseteq x^{*+}$ và (b) $\bigcap_{x\in X} x^- \subseteq x^{*-}$.

P2. Mỗi cấu trúc hội $(x^{*+}, x^{*-}) \in C(X)$ phải có (a) $x^{*+} \subseteq \bigcup_{x\in X} x^+$ và (b) $x^{*-} \subseteq \bigcup_{x\in X} x^-$.

P3. Nếu X là nhất quán thì cấu trúc hội $(\bigcup_{x\in X} x^+, \bigcup_{x\in X} x^-)$ cũng là một đồng thuận của X.

P4. Với mọi $(x^{*+}, x^{*-}) \in C(X)$, ta đều có $x^{*+} \cap x^{*-} = \varnothing$.

P5. Một đồng thuận $x^* \in C(X)$ phải đạt cực tiểu tổng các khoảng cách:

$$\sum_{x\in X} d_{\wedge}(x^*, x) = min\left\{\sum_{x\in X} d_{\wedge}(x', x) \mid x' \in Conj(\mathbf{L})\right\}.$$

P6. Với mỗi ký hiệu $z \in \mathbf{L}$ và một đồng thuận $x^* \in C(X)$, dạng thức xuất hiện (tức là, dưới dạng literal âm hay dương) của z trong x^* là chỉ phụ thuộc vào dạng thức xuất hiện của nó trong các cấu trúc hội của X .

Trong một đồng thuận $(x^{*+}, x^{*-}) \in C(X)$, tập x^{*+} (tương ứng, tập x^{*-}) được gọi là thành phần khẳng định (tương ứng, thành phần phủ định).

Ta ký hiệu C_{co} là tập tất cả các hàm đồng thuận cho các hồ sơ cấu trúc hội. Mặc dù các tiêu chuẩn định nghĩa ở trên là dành cho các hàm đồng thuận nhưng chúng cũng có thể được xét để được thoả cho các đồng thuận cụ thể nào đó của một hồ sơ.

4 Phân tích các tiêu chuẩn đồng thuận

Trong phần này, chúng ta sẽ khảo sát các tính chất của các tiêu chuẩn cho hàm chọn đồng thuận đã nêu và mối quan hệ giữa các hàm đồng thuận thoả mãn các tiêu chuẩn này. Ta quy ước:

- Hàm đồng thuận C thoả tiêu chuẩn P cho 1 hồ sơ các cấu trúc hội X được viết là $C(X) \bullet P$.
- Hàm đồng thuận C thoả tiêu chuẩn P với mọi hồ sơ cấu trúc hội, được viết là $C \bullet P$.
- Tiêu chuẩn P được thoả với mọi hàm đồng thuận $C \in C_{co}$ được viết là $C_{co} \bullet P$.

Hai định lý dưới đây chỉ ra sự phụ thuộc của tiêu chuẩn $P5$ với hai tiêu chuẩn $P1$, $P2$ và tính độc lập của hai thành phần khẳng định, phủ định của đồng thuận thoả tiêu chuẩn $P5$. Do giới hạn số trang của bài báo, chúng tôi không trình bày phần chứng minh của các định lý này ở đây.

Định nghĩa 4.1. Một hàm đồng thuận thoả tiêu chuẩn $P5$ thì cũng thoả hai tiêu chuẩn $P1$ và $P2$; tức là $(C \bullet P5) \Rightarrow (C \bullet P1 \wedge C \bullet P2)$ với mọi $C \in C_{co}$.

Định nghĩa 4.2. Thành phần khẳng định và thành phần phủ định của một đồng thuận thoả tiêu chuẩn P5 có thể được xác định độc lập nhau; nghĩa là, một cấu trúc hội (x^{*+}, x^{*-}) là một đồng thuận của X nếu và chỉ nếu cấu trúc hội $(x^{*+}, \varnothing)$ là đồng thuận của $X' = \{(x_i^{+}, \varnothing) : i = 1, 2, \ldots, n\}$, và cấu trúc hội $(\varnothing, x^{*-})$ là đồng thuận của $X'' = \{(\varnothing, x_i^{-}) : i = 1, 2, \ldots, n\}$.

Một tính chất nữa của tiêu chuẩn P5 mà thông qua đó có thể chỉ ra cách xây dựng đồng thuận thoả tiêu chuẩn này cho một hồ sơ $X \in \prod(Conj(\mathbf{L}))$ được trình bày ở định lý sau đây:

Định nghĩa 4.3. Cho $X = \{x_i, i = 1, 2, \ldots, n\}$ là một hồ sơ xung đột gồm các cấu trúc hội, $X \in \prod(Conj(\mathbf{L}))$. Gọi:

• $\mathbf{Z}^{+}$ (tương ứng, $\mathbf{Z}^{-}$) là tập hợp tất cả các literal xuất hiện trong các thành phần dương (tương ứng, thành phần âm) của các cấu trúc hội thuộc hồ sơ X .

• $f^{+}(z)$ (tương ứng, $f^{-}(z)$) là tần số của phần tử z trong các thành phần dương (tương ứng, thành phần âm) của các cấu trúc hội thuộc hồ sơ X .

Giả sử $C(X)$ là một hàm chọn đồng thuận thoả tiêu chuẩn P5. Khi đó, $x^{*}=(x^{*+},x^{*-})\in C(X)$ nếu và chỉ nếu: (a) $x^{*+}=\{z\in\mathbf{Z}^{+}\mid f^{+}(z)>=\frac{n}{2}\}$ và (b) $x^{-+}=\{z\in\mathbf{Z}^{-}\mid f^{-}(z)>=\frac{n}{2}\}$.

Chứng minh.

Theo định lý 6.2, ta có thể xây dựng đồng thuận $x^{*}=(x^{*+},x^{*-})$ thoả tiêu chuẩn P5 bằng cách xây dựng một cách độc lập hai đồng thuận x^{*+} và x^{*-} của tương ứng hai hồ sơ $X^{+}=\{x_i^{+},i=1,2,\ldots,n\}$ và $X^{-}=\{x_i^{-},i=1,2,\ldots,n\}$. Ta sẽ phải chứng minh (a) $x^{*+}=\{z\in\mathbf{Z}^{+}\mid f^{+}(z)>=\frac{n}{2}\}$. Việc chứng minh (b) là hoàn toàn tương tự.

Trước hết, theo Định lý 6.1 ,x^{*+} chỉ có thể chứa các literal thuộc về $\mathbf{Z}^{+}$. Mặt khác, ta sẽ chứng minh thêm, với một cấu trúc hội bất kỳ $x\in Conj(\mathbf{L})$, ta có: [(i)]

(i) Nếu $z\in Z^{+}$ thoả $f^{+}(z)\geq\frac{n}{2}$ mà $z\notin x^{+}$ thì $d_{\wedge}((x^{+},x^{-}),X^{+})\leq d_{\wedge}((x^{+}\cup\{z\},x^{-}),X^{+})$

(ii) Nếu $z\in Z^{+}$ thoả $f^{+}(z)<\frac{n}{2}$ mà $z\notin x^{+}$ thì $d_{\wedge}((x^{+},x^{-}),X^{+})>d_{\wedge}((x^{+}\cup\{z\},x^{-}),X^{+})$

Thật vậy, với $X\in\prod(Conj(\mathbf{L}))$, $x\in Conj(\mathbf{L})$, $z\in\mathbf{L}$, xét khoảng cách $d_{\wedge}((x^{+}\cup\{z\},x^{-}),X)$:

$$d_{\wedge}((x^{+}\cup\{z\},x^{-}),X)=\sum_{y\in X}\left(\frac{1}{2}.\frac{card((x^{+}\cup\{z\})\Delta y^{+})}{card(\mathbf{L})}+\frac{1}{2}.\frac{card(x^{-}\Delta y^{-})}{card(\mathbf{L})}\right)$$

Nhận thấy:

$$\frac{card((x^{+}\cup\{z\})\Delta y^{+})}{card(\mathbf{L})}=\begin{cases}\dfrac{card(x^{+}\Delta y^{+})-1}{card(\mathbf{L})} & n?u\ \ y^{+}\ni z\\[2mm] \dfrac{card(x^{+}\Delta y^{+})+1}{card(\mathbf{L})} & n?u\ \ y^{+}\not\ni z\end{cases}$$

Gọi $X_z^{-}:=\{x\in X\mid x^{+}\ni z\}$ và $X_{\bar z}:=\{x\in X\mid x^{+}\not\ni z\}$. Ta có, $card(X_z)=f^{+}(z)$ và $card(X_{\bar z})=n-f^{+}(z)$. Như vậy:

$$\begin{aligned}\sum_{y\in X}\frac{card((x^{+}\cup\{z\})\Delta y^{+})}{card(\mathbf{L})}&=\sum_{y\in X_z}\frac{card((x^{+}\cup\{z\})\Delta y^{+})}{card(\mathbf{L})}+\sum_{y\in X_{\bar z}}\frac{card((x^{+}\cup\{z\})\Delta y^{+})}{card(\mathbf{L})}\\&=\sum_{y\in X_z}\frac{card(x^{+}\Delta y^{+})-1}{card(\mathbf{L})}+\sum_{y\in X_{\bar z}}\frac{card(x^{+}\Delta y^{+})+1}{card(\mathbf{L})}\\&=\sum_{y\in X}\frac{card(x^{+}\Delta y^{+})}{card(\mathbf{L})}+\frac{-f^{+}(z)+n-f^{+}(z)}{card(\mathbf{L})}\\&=\sum_{y\in X}\frac{card(x^{+}\Delta y^{+})}{card(\mathbf{L})}+\frac{n-2.f^{+}(z)}{card(\mathbf{L})}.\end{aligned}$$

Do đó:

$$d_{\wedge}((x^{+} \cup \{z\}, x^{-}), X) = \sum_{y \in X} (\frac{1}{2}.\frac{card((x^{+} \cup \{z\})\Delta y^{+})}{card(\mathbf{L})} + \frac{1}{2}.\frac{card(x^{-}\Delta y^{-})}{card(\mathbf{L})})$$
$$= \sum_{y \in X} (\frac{1}{2}.(\frac{card(x^{+}\Delta y^{+})}{card(\mathbf{L})} + \frac{n-2.f^{+}(z)}{card(\mathbf{L})}) + \frac{1}{2}.\frac{card(x^{-}\Delta y^{-})}{card(\mathbf{L})})$$
$$= d_{\wedge}(x, X) + \frac{1}{2}.\frac{n-2.f^{+}(z)}{card(\mathbf{L})}.$$

Như vậy, khi $n-2.f^{+}(z) \le 0$, hay $f^{+}(z) \ge \frac{n}{2}$ thì việc bổ sung z vào x^{+} sẽ không làm tăng tổng khoảng cách của x đến hồ sơ X. Ngược lại, khi $f^{+}(z) < \frac{n}{2}$ thì việc bổ sung z vào x^{+} sẽ làm tăng tổng khoảng cách của x đến hồ sơ X. Nói cách khác, (i) và (ii) được chứng minh.

Trở lại việc chứng minh (a). Có thể thấy rằng, xuất phát từ tập hợp $\{z \in \mathbf{Z}^{+} \mid f^{+}(z) >= \frac{n}{2}\}$, ta không thể loại bỏ bớt phần thuộc tập hợp này, hoặc bổ sung thêm phần tử khác thuộc $\mathbf{Z}^{+}$ trong quá trình thành lập thành phần dương của đồng thuận. Nói cách khác, đây chính là thành phần dương của đồng thuận. Do đó (a) là đúng (đpcm).

5 Thuật toán xác định đồng thuận

Dựa vào các tính chất của các tiêu chuẩn đã được phân tích ở Phần 6, phần này sẽ trình bày cách xây dựng đồng thuận $x^{*} = (x^{*+}, x^{*-})$ của một hồ sơ xung đột $X \in \prod(Conj(\mathbf{L}))$ theo chiến lược ưu tiên các tiêu chuẩn với thứ tự như sau: $P4$, $P1$, $P2$, $P3$, $P5$ và $P6$.

Định lý 4.1 cho thấy điều kiện được định nghĩa trong tiêu chuẩn $P5$ là rất quan trọng, bởi vì nói chung một đồng thuận thoả tiêu chuẩn này thì cũng thoả tiêu chuẩn $P1$ và $P2$. Ngoài ra, theo định lý 6.2, việc xác định các thành phần dương và âm của một đồng thuận có thể được thực hiện một cách độc lập. Vì thế, việc tính toán cấu trúc hội tối ưu $(x^{*+}, x^{*-}) \in C(X)$ mà $X = \{x_i = (x_i^{+}, x_i^{-}) \in Conj(\mathbf{L}) : i = \overline{1,n}\}$ có thể được chia ra làm 2 việc nhỏ tương tự nhau: xác định thành phần dương của đồng thuận và xác định thành phần âm của đồng thuận:

$$\sum_{x \in X} \eta(x^{*+}, x^{+}) = min\{\sum_{x \in X} \eta(x'^{+}, x^{+}) \mid x' \subseteq \mathbf{L}\}$$

và

$$\sum_{x \in X} \eta(x^{*-}, x^{-}) = min\{\sum_{x \in X} \eta(x'^{-}, x^{-}) \mid x' \subseteq \mathbf{L}\}.$$

Định lý 4.3 chỉ ra cách tìm các thành phần này, tuy nhiên, định lý này lại không đảm bảo được đồng thuận tìm được thoả tiêu chuẩn P4. Trên cơ sở của các phân tích này, chúng tôi đưa ra thuật toán xác định đồng thuận như sau đây (Thuật toán 1).

Chứng minh tính đúng của thuật toán.

Input: Hồ sơ xung đột $\mathsf{X} \in \prod(Conj(\mathbf{L}))$, $\mathsf{X} = \{(x_i{}^+, x_i{}^-), i = 1, 2, \ldots, n\}$,
$x_i{}^+ \cap x_i{}^- = \emptyset \ \forall i = 1, 2, \ldots, n$.
Output: Đồng thuận $x^* \in Conj(\mathbf{L})$ thoả một hoặc nhiều tiêu chuẩn trong tập hợp {P4, P1, P2, P3, P5, P6}.

begin
  $\mathbf{Z}^+ := \bigcup_{x \in \mathsf{X}} x^+$; $\mathbf{Z}^- := \bigcup_{x \in \mathsf{X}} x^-$;
  foreach $z \in \mathbf{Z}^+$ **do**
    $f^+(z) := card\{x \in \mathsf{X} \mid x^+ \ni z\}$;
  foreach $z \in \mathbf{Z}^-$ **do**
    $f^-(z) := card\{x \in \mathsf{X} \mid x^- \ni z\}$;
(a) $x^{*+} := \{z \in \mathbf{Z}^+ \mid f^+(z) \geq \frac{n}{2}\}$;
  $x^{*-} := \{z \in \mathbf{Z}^- \mid f^-(z) \geq \frac{n}{2}\}$;
  if $(x^{*+} \cup x^{*-} \neq \emptyset)$ **then**
(b)   **foreach** $z \in x^{*+} \cap x^{*-}$ **do**
      if $d_\wedge((x^{*+} \setminus \{z\}, x^{*-}), \mathsf{X}) < d_\wedge((x^{*+}, x^{*-} \setminus \{z\}), \mathsf{X})$ **then**
        $x^{*+} := x^{*+} \setminus \{z\}$;
      else
        $x^{*-} := x^{*-} \setminus \{z\}$;
  else
    if $(\mathbf{Z}^+ \cap \mathbf{Z}^- = \emptyset)$ **then**
(c) $x^* := (\mathbf{Z}^+, \mathbf{Z}^-)$;
    else
(d) $x^* := x_1$;
      for $i := 2$ *to* n **do**
        if $d_\wedge(x^*, \mathsf{X}) > d_\wedge(x, \mathsf{X})$ **then**
          $x^* := x_i$

Theo cách hoạt động được chỉ ra ở Thuật toán 1, chúng ta bắt đầu tìm đồng thuận thoả tiêu chuẩn P5 (phần (a)). Sau đó:

(i) Nếu cả hai thành phần dương và âm của đồng thuận P5 đều rỗng, thuật toán sẽ ưu tiên xét tìm đồng thuận thoả tiêu chuẩn P3 (phần (c)) nếu hồ sơ X là nhất quán. Trong trường hợp hồ sơ là không nhất quán, theo phần (d), chúng ta sẽ chọn từ hồ sơ X một phần tử có tổng khoảng cách đến các phần tử còn lại trong hồ sơ là cực tiểu. Đồng thuận trong trường hợp này luôn luôn thoả tiêu chuẩn P4 (do các x_i đều là các cấu trúc hội thoả $x_i{}^+ \cap x_i{}^- = \varnothing$, $\forall i = 1,2,\ldots,n$, theo giả thiết).

(ii) Nếu ở phần (a), chúng ta xác định được một trong hai thành phần phần dương và âm của đồng thuận là khác rỗng, chúng ta phải tìm cách làm mịn hai thành phần này để đảm bảo tiêu chuẩn P4 được thoả, đồng thời cũng đảm bảo tổng khoảng cách từ đồng thuận đến các phần tử trong hồ sơ là cực tiểu. Ngoài ra, trong tất cả các trường hợp xử lý thuộc hai nhánh phân tích (i) và (ii) ở trên, chúng ta đều xây dựng x^{*^+} (tương ứng, x^{*^-}) từ các phần tử thuộc z^+ (tương ứng, z^-). Vậy nên đồng thuận luôn luôn thoả tiêu chuẩn P2. Đồng thuận cũng luôn luôn thoả tiêu chuẩn P1 vì nó được xây dựng từ đồng thuận thoả tiêu chuẩn P5, sau đó, các phần tử bị loại đi chỉ là những phần tử có tần số xuất hiện ít hơn $\frac{n}{2}$.

Độ phức tạp của thuật toán là $O(n.m^2)$, với n là số lượng phần tử của hồ sơ xung đột, $m = max\{card(\bigcup_{x \in X} x^+), card(\bigcup_{x \in X} x^-)\}$.

Ví dụ minh hoạ cho Thuật toán 7

Giả sử có 6 tác tử $a_1, a_2, \ldots, a_6$ cùng tham gia mô tả tri thức với tập ký hiệu literal $\mathbf{L} = \{t_1, t_2, t_3, t_4\}$ như Bảng 1:

Bảng 1. Trạng thái tri thức của các tác tử cho bài toán minh hoạ

Tác tử	Trạng thái tri thức
a_1	$t_1 \wedge \neg t_2 \wedge t_3 \wedge t_4$
a_2	$t_1 \wedge \neg t_3 \wedge \neg t_4$
a_3	$t_1 \wedge \neg t_3$
a_4	$t_1 \wedge \neg t_3 \wedge \neg t_4$
a_5	$\neg t_1 \wedge t_3 \wedge \neg t_4$
a_6	t_3

Chúng ta sẽ áp dụng Thuật toán 7 để tìm tri thức đồng thuận từ các ý kiến của các tác tử. Trước tiên, chúng ta lập hồ sơ X gồm các cấu trúc hội:

$$X = \{(\{t_1, t_3, t_4\}, \{t_2\}), 2*(\{t_1\}, \{t_3, t_4\}), (\{t_1\}, \{t_3\}), (\{t_3\}, \{t_1\}), (\{t3\}, \varnothing)\}$$

Sau bước (a) của thuật toán, chúng ta có: $x^{*+} = \{t_1, t_3\}$ và $x^{*-} = \{t_3, t_4\}$. Do $x^{*+} \cup x^{*-} \neq \varnothing$ nên chúng ta sẽ tìm cách loại những literal chung ở một trong hai thành phần của đồng thuận (theo bước (b): Với $x^{*+} \cap x^{*-} = \{t_3\}$, ta xét hai tổng khoảng cách sau đây: $d_\wedge((\{t_1\}, \{t_3, t_4\}), X)$ và $d_\wedge((\{t_1, t_3\}, \{t_4\}), X)$

Với chú ý $card(\mathbf{L}) = 4$, ta lần lượt tính:

$$\begin{aligned} d_\wedge((\{t_1\}, \{t_3, t_4\}), (\{t_1, t_3, t_4\}, \{t_2\})) &= \frac{1}{2}.\frac{card(\{t_1\}\Delta\{t_1, t_3, t_4\})}{4} + \frac{1}{2}.\frac{card(\{t_3, t_4\}\Delta\{t_2\})}{4} \\ &= \frac{5}{8} \end{aligned}$$

Tương tự,

- $d_\wedge((\{t_1\}, \{t_3, t_4\}), (\{t_1\}, \{t_3, t_4\})) = 0$
- $d_\wedge((\{t_1\}, \{t_3, t_4\}), (\{t_1\}, \{t_3\})) = \frac{2}{8}$
- $d_\wedge((\{t_1\}, \{t_3, t_4\}), (\{t_3\}, \{t_1, t_4\})) = \frac{4}{8}$
- $d_\wedge((\{t_1\}, \{t_3, t_4\}), (\{t_3\}, \varnothing)) = \frac{4}{8}$
- $d_\wedge((\{t_1, t_3\}, \{t_4\}), (\{t_1, t_3, t_4\}, \{t_2\})) = \frac{2}{8}$

- $d_{\wedge}((\{t_1,t_3\},\{t_4\}),(\{t_1\},\{t_3,t_4\})) = \frac{2}{8}$
- $d_{\wedge}((\{t_1,t_3\},\{t_4\}),(\{t_1\},\{t_3\})) = \frac{3}{8}$
- $d_{\wedge}((\{t_1,t_3\},\{t_4\}),(\{t_3\},\{t_1,t_4\})) = \frac{2}{8}$
- $d_{\wedge}((\{t_1,t_3\},\{t_4\}),(\{t_3\},\varnothing)) = \frac{2}{8}$

Như vậy:

- $d_{\wedge}((\{t_1\},\{t_3,t_4\}), X) = \frac{5}{8} + 2*0 + \frac{2}{8} + \frac{4}{8} + \frac{4}{8} = \frac{15}{8}$
- $d_{\wedge}((\{t_1,t_3\},\{t_4\}), X) = \frac{2}{8} + 2*\frac{2}{8} + \frac{3}{8} + \frac{2}{8} + \frac{2}{8} = \frac{13}{8}$

Do $d_{\wedge}((\{t_1,t_3\},\{t_4\}), X) = \frac{13}{8} < \frac{15}{8} = d_{\wedge}((\{t_1\},\{t_3,t_4\}), X)$ nên ta kết luận, đồng thuận của hồ sơ X là $(\{t_1,t_3\},\{t_4\})$, hay là $t_1 \wedge t_3 \wedge \neg t_4$.

6 Kết luận và hướng phát triển

Trong bài báo này, chúng tôi đã trình bày mô hình xác định đồng thuận của tri thức biểu diễn dưới dạng hội của các literal. Cụ thể, chúng tôi đã đề xuất hàm tính khoảng cách cho hai biểu thức dạng hội của các literal, nêu một số định lý thể hiện mối quan hệ giữa các tiêu chuẩn xác định đồng thuận trong không gian khoảng cách được xây dựng. Trên cơ sở các định lý này, chúng tôi đã đưa ra thuật toán xác định đồng thuận cho hồ sơ xung đột của tri thức dạng hội. Các phụ thuộc và thuật toán này có thể được điều chỉnh để áp dụng cho cấu trúc tuyển của các literal (còn có thể được chuyển đổi thành dạng quy tắc), tuy nhiên, do giới hạn số trang của bài báo, chúng tôi không trình bày ở đây.

Mô hình xác định đồng thuận tri thức trong bài báo này có thể được áp dụng cho nhiều bài toán trong thực tế. Chẳng hạn, một tác tử cần tổng hợp tri thức từ nhiều nguồn khác nhau, phân tán để ra quyết định; hoặc người quản lý kho tri thức dạng wiki cần xác định thông tin tổng hợp từ nhiều phiên bản được cung cấp bởi nhiều người dùng khác nhau.

Tuy nhiên, có thể thấy rằng, cấu trúc dạng hội (hoặc tuyển) của các literal là khá đơn giản, khó có thể biểu diễn tri thức tổng quát trong thực tế. Hướng mở rộng của bài báo này là xây dựng mô hình để có thể làm việc với các biểu thức logic bậc nhất dạng chuẩn hội và chuẩn tuyển.

Tài liệu tham khảo

1. J. P. Barthélemy and M. F. Janowiĵ. A Formal Theory of Consensus. SIAM *Journal on Discrete Mathematics*, 4(3):305–322, 1991.
2. Stefano Ferilli, Nicola Di Mauro, Marenglen Biba, and Floriana Esposito. SimilarityGuidedClause Generalization. In AI* IA 2007: *Artificial Intelligence and Human-Oriented Computing*, pages 278-289. Springer, 2007.
3. Zhisheng Huang and Frank van Harmelen. Using semantic distances for reasoning with inconsistent ontologies. In The Semantic Web - ISWC 2008, pages 178–194. Springer, 2008.
4. Ngoc Thanh Nguyen. Methods for Consensus Choice and their Applications in Conflict Resolving in Distributed Systems. Wroclaw University of Technology Press, 2002.
5. Ngoc Thanh Nguyen. Advanced Methods for Inconsistent Knowledge Management (Advanced Information and Knowledge Processing). Springer, 2007.

A METHOD TO RESOLVE INCONSISTENCE OF KNOWLEDGE ON SYNTACTIC LEVEL

Trung Van Nguyen[1], Hanh Huu Hoang[2]

[1]University of Sciences - Hue University
[2]Hue University

Abstract. There are two levels for processing inconsistent knowledge which are based on logic: syntactic level and semantic level. In this paper, we propose a consensus-based method to resolve inconsistent knowledge on the syntactic level, where a knowledge state can be encoded by a conjunction of literals.