TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG



BÁO CÁO BÀI TẬP LỚN MÔN HỆ CƠ SỞ TRI THỨC

Úng dụng Đại số gia tử, Picture Fuzzy Set vào xây dựng hệ thống hỗ trợ chuẩn đoán bệnh

Giảng viên hướng dẫn: PGS. TS. Phạm Văn Hải Nhóm 2

Nguyễn Lưu Nhật 20163041 Đinh Hoàng Nam 20162793

Hà Nội, 2019

Lời cảm ơn

Chúng em xin được gửi lời cảm ơn sâu sắc tới PGS. TS. Phạm Văn Hải, người đã trực tiếp tận tình hướng dẫn, chỉ bảo, và đưa ra những lời khuyên quý báu giúp đỡ chúng em trong quá trình thực hiện đề tài.

Tóm tắt nội dung đề tài

Đề tài thực hiện tìm hiểu lý thuyết Đại số gia tử nhằm chuẩn hóa các thuộc tính chưa thuần nhất. Ngoài ra, đề tài cũng tìm hiểu các lý thuyết về Fuzzy logic, Picture Fuzzy Set, để từ đó xây dựng một hệ thống hỗ trợ chuẩn đoán bệnh đơn giản.

MỤC LỤC

CH	UONG 1.	TỔNG QUAN ĐỀ TÀI	1
1.1	Đặt vấi	n đề	1
1.2	Mục tiế	êu và nhiệm vụ đề tài	2
1.3	Phân cô	ông công việc	2
CH	ƯƠNG 2.	CƠ SỞ LÝ THUYẾT	3
2.1	Đại số	gia tử	3
	2.1.1	Giới thiệu đại số gia tử	3
	2.1.2	Định nghĩa đại số gia tử	4
	2.1.3	Các đại lượng đo trên đại số gia tử	5
	2.1.4	Xây dựng hàm định lượng ngữ nghĩa	7
	2.1.5 ψ _{max}] và t	Định lượng giá trị ngôn ngữ khi chỉ biết một đoạn con co coàn bộ IC(ω)	=
	2.1.6 ψ _{max}] nhu	Định lượng giá trị ngôn ngữ khi chỉ biết một đoạn con c rng chưa xác định được toàn bộ IC(ω)	•
2.2	Picture	Fuzzy Set	10
	2.2.1	Fuzzy Set	10
	2.2.2	Fuzzly Logic	11
	2.2.3	Picture Fuzzy Set	12
CH	UONG 3.	MÔ HÌNH KIẾN TRÚC HỆ THỐNG	13
3.1	Mô hìn	h hệ thống	13
CH	UONG 4.	MÔ TẢ THUẬT TOÁN	14
Bướ	c 1: Biểu ơ	diễn tri thức	14
	Step 1:	Biểu diễn tri thức bằng mạng ngữ nghĩa	14
	Step2:	Lưu trữ tri thức bằng cơ sở dữ liệu quan hệ	15
Bướ	c 2: Chuẩn	ı hóa input	15
Bướ	c 3: Chuẩn	ı đoán bệnh	15
CH	UONG 5.	VÍ DŲ	17
Ví d	lụ 1: Xây c	dựng hàm định lượng ngữ nghĩa cho Đại số gia tử AT	17
Ví d	lụ 2: Chuẩ	n đoán bệnh	19

CHU	ONG 6. KÉT LUẬN	21
6.1	Kết luận	21
6.2	Hướng phát triển trong tương lai	21
TÀI I	LIỆU THAM KHẢO	22

DANH MỤC HÌNH VỄ

Hình 2.1 Tính mờ của một giá trị ngôn ngữ [1]
Hình 2.2 Minh họa Fuzzy Logic [3]11
Hình 3.1 Mô hình kiến trúc hệ thống
Hình 3.2 Biểu diễn tri thức bằng mạng ngữ nghĩa14
Hình 3.3 Lưu trữ tri thức bằng csdl quan hệ
DANH MỤC BẨNG BIỂU
Bảng 5.1 Quan hệ giữa bệnh nhân và các triệu chứng
Bảng 5.2: Quan hệ giữa các triệu chứng và các bệnh
Bảng 5.3: Quan hệ T = R * O

CHƯƠNG 1. TỔNG QUAN ĐỀ TÀI

1.1 Đặt vấn đề

Những năm gần đây, mặc dù nhận được nhiều sự quan tâm và đầu tư nhưng nhìn chung, hệ thống y tế nước ta vẫn còn nhiều điểm bất cập. Đặc biệt là tình trạng quá tải của các bệnh viện lớn khi số lượng bệnh nhân từ khắp cả nước đổ về khám chữa. Trong khi đó, ở các miền vùng sâu, vùng xa, việc tiếp cận với các dịch vụ y tế còn rất hạn chế. Việc lên các bệnh viện trung tâm để thăm khám khi này rất tốn kém về thời gian và công sức. Nhận thức của người dân về bệnh tật cũng còn rất mơ hồ. Điều này thường dẫn đến việc chỉ khi bệnh đã nặng, người bệnh mới bắt đầu quan tâm chữa trị.

Để khắc phục những hạn chế này, chúng ta cần một giải pháp giúp cho người dân sớm có thể sớm phát hiện những căn bệnh của mình, từ đó giúp ích cho việc điều trị, cũng như việc cung cấp những thông tin hữu ích cho việc điều trị những căn bệnh đơn giản, làm giảm tránh áp lực lên các bệnh viện. Trong thời đại công nghệ phát triển như hiện nay, các giải pháp công nghệ thông tin tỏ ra khá khả thi và hữu ích. Tuy vậy, việc xây dựng một hệ thống hỗ trợ trong lĩnh vực healthcare cũng còn gặp phải rất nhiều các khó khăn, trở ngại.

Việc chuẩn đoán bệnh trong y học là một công việc khá phức tạp, đòi hỏi nhiều kiến thức chuyên môn cũng như kinh nghiệm của y bác sĩ. Các kiến thức và kinh nghiệm này lại rất khó để có thể biểu diễn một cách cụ thể, chính xác. Thông thường, các thông tin có được từ các chuyên gia trong lĩnh vực này thường ở dạng ngôn ngữ tự nhiên, các thông tin mờ, chỉ có mức độ chắc chắn nhất định nào đó.

Mặc dù các giá trị ngôn ngữ rất tự nhiên và gần gũi với cách tư duy, suy nghĩ của con người, nhưng những thông tin này thường không mang một giá trị cụ thể, chính xác. Do vậy mà các thông tin này thường khó có thể xử lý trực tiếp bởi các hệ thống máy tính. Điều này làm nảy sinh một nhu cầu về việc chuẩn hóa các giá trị ngôn ngữ về miền các giá trị số kinh điển.

Những năm gần đây, lý thuyết Đại số gia tử đã nhận được rất nhiều sự quan tâm, nghiên cứu và ứng dụng. Đại số gia tử là một công cụ mạnh có thể giúp ta trong việc định lượng các giá trị ngôn ngữ. Chính vì những lý do trên, đề tài này xin thực hiện tìm hiểu việc chuẩn hóa thuộc tính giá trị ngôn ngữ bằng đại số gia tử, tìm hiểu các kiến thức liên quan đến logic mờ, và từ đó xây dựng một hệ thống hỗ trợ chuẩn đoán bệnh đơn giản.

1.2 Mục tiêu và nhiệm vụ đề tài

- 1. Tìm hiểu lý thuyết Đại số gia tử
- 2. Tìm hiểu lý thuyết Picture Fuzzy Set
- 3. Tìm hiểu phương pháp định lượng giá trị ngôn ngữ về giá trị số
- 4. Tìm hiểu cách xây dựng một hệ thống có sử dụng tri thức
- 5. Thực hành xây dựng một hệ thống hỗ trợ chuẩn đoán bệnh đơn giản

1.3 Phân công công việc

Nguyễn Lưu Nhật	- Tìm kiếm tài liệu		
(nhóm trưởng)	- Tìm hiểu lý thuyết Đại số gia tử		
	- Tìm hiểu lý thuyết Picture Fuzzy Set		
	- Tham gia xây dựng giải thuật		
	- Cài đặt chương trình		
	- Viết báo cáo		
Đinh Hoàng Nam	- Tìm hiểu lý thuyết Đại số gia tử		
	- Tìm hiểu lý thuyết Picture Fuzzy Set		
	- Viết giải thuật		
	- Tham gia xây dựng chương trình		
	- Viết báo cáo		

CHƯƠNG 2. CƠ SỞ LÝ THUYẾT

Nội dung trong chương này được tham khảo từ các tài liệu được thầy gửi cho lớp và trên internet.

2.1 Đại số gia tử

Phần trình bày dưới đây được tham khảo từ tài liệu [1]

2.1.1 Giới thiệu đại số gia tử

Xét một tập giá trị ngôn ngữ là miền của biến ngôn ngữ (linguistic domain) của biến chân lý TRUTH gồm các từ sau:

T = dom(TRUTH) = {true, false, very true, very false, more true, more false, approximately true, approximately false, little true, little false, less true, less false, very more true, very more false, very possible true, very possible false, very more true, very more false, ...}.

Khi đó miền ngôn ngữ T = dom(TRUTH) có thể biểu thị như là một cấu trúc đại số $AT = (T, G, H, \leq)$, trong đó:

- T: Là tập cơ sở của AT.
- G: Là tập các từ nguyên thủy (tập các phần tử sinh: true, false).
- H: Là tập các toán tử một ngôi, gọi là các gia tử (các trạng từ nhấn).

Ví dụ: dựa trên ngữ nghĩa ta có: false \leq true, more true \leq very true, very false \leq more false...

Các quy ước:

- (1) Các gia tử trong H là các **toán tử thứ tự**, nghĩa là $(\forall h \in H, h: T \to T), (\forall x \in T) \{hx \le x \ hoặc \ hx \ge x\}.$
- (2) Hai gia tử h, k \in H được gọi là **ngược nhau** nếu ($\forall x \in T$) { $hx \le x$ khi và chỉ khi $kx \ge x$ }.
- (3) Hai gia tử h, k \in H được gọi là **tương thích** nhau nếu $(\forall x \in T) \{hx \leq x \ khi \ và \ chỉ \ khi \ kx \leq x\}.$
- (4) Ta ký hiệu $\mathbf{h} \ge \mathbf{k}$ nếu h, k tương thích nhau và $(\forall x \in T) \{hx \le kx \le x \text{ hoặc } hx \ge kx \ge x\}.$

Khi đó: tập H có thể được phân hoặch thành hai tập H + và H - với các gia tử trong tập H + hay H - là tương thích nhau, mỗi phần tử trong H + cũng ngược với bất kỳ phần tử nào trong H - và ngược lại.

- (5) Giả sử V (Very) là phần tử lớn nhất trong H + thì: phần tử sinh $g \in G$ là **dương** nếu $g \leq Vg$ và là **âm** nếu $g \geq Vg$.
- (6) Một gia tử **h dương** đối với một gia tử **k** nếu: $(\forall x \in T) \{hkx \le kx \le x \text{ hoặc } hkx \ge kx \ge x\}.$
- (7) T được sinh ra từ G bởi các gia tử trong H, như vậy: Mỗi phần tử của T sẽ có dạng biểu diễn là $x = h_n h_{n-1} h \dots h_2 h_1 u$, $u \in G$.
 - (8) Tập tất cả các phần tử được sinh ra từ phần tử x có dạng biểu diễn là H(x).

Nếu G chỉ có đúng 2 từ nguyên thủy mờ: thì một được gọi là phần tử sinh dương ký hiệu là t, một được gọi là phần tử sinh âm ký hiệu là f và ta có f < t.

2.1.2 Định nghĩa đại số gia tử

Một cấu trúc đại số $AT = (T, G, H, \leq)$ với H được phân hoặch thành H+ và H- các gia tử ngược nhau được gọi là một đại số gia tử nếu nó thỏa mãn các tiên đề sau:

- (1) Mỗi gia tử hoặc là dương hoặc là âm đối với bất kỳ một gia tử nào khác.
- (2) Nếu hai khái niệm u và v là độc lập nhau, nghĩa là $u \notin H(v)$ và $v \notin H(u)$, thì $\forall x \in H(u) \{x \notin H(v)\}.$
- (3) Nếu $x \neq hx$ thì $x \notin H(hx)$ và nếu $h \neq k$ và $hx \leq kx$ thì $h'hx \leq k'kx$, với mọi gia tử h, k, h' và k'. Hơn nữa nếu $hx \neq kx$ thì hx và kx là độc lập.
- (4) Nếu $u \notin H(v)$ và $u \le v$ (hoặc $u \ge v$) thì $u \le hv$ (hoặc $u \ge hv$) đối với mọi gia tử h.

Xét đại số gia tử AT có đúng 3 phần tử sinh: dương, âm và một phần tử trung hòa w nằm giữa hai phần tử sinh kia và có tính chất hw = w, với mọi h \in H. Một phần tử y được gọi là phần tử đối nghịch của phần tử x nếu có tồn tại một biểu diễn của x có dạng $x = h_n \dots h_1 g$, w \neq g \in G, sao cho $y = h_n \dots h_1 g'$, với w \neq g \in G và g' \neq g

Đặc biệt phần đối nghịch của w được định nghĩa chính là w. Phần tử đối nghịch của x được ký hiệu là –x với chỉ số nếu cần thiết. Nhìn chung một phần tử có thể có nhiều phần tử đối nghịch.

Nếu mỗi phần tử của T chỉ có duy nhất một phần tử đối nghịch thì AT được gọi là đại số gia tử đối xứng.

2.1.3 Các đại lượng đo trên đại số gia tử

Định nghĩa 1: Hàm đo trên đại số gia tử:

Cho đại số gia tử mở rộng đối xứng $(T, G, H, \leq), f : T \rightarrow [0, 1]$ là một hàm đo trên T nếu thoả mãn:

(1)
$$\forall t \in T$$
: $f(t) \in [0, 1]$, $f(g+) = 1$, $f(g-) = 0$.

(2)
$$\forall x, y \in T$$
, nếu $x < y$ thì $f(x) < f(y)$.

Thông qua hàm đo ta có thể phần nào so sánh được mức độ ngữ nghĩa giữa các phần tử của các đại số gia tử khác nhau.

Ví dụ: từ hai đại số gia tử chiều_cao và cân_nặng thì mức độ chênh lệch giữa "rất cao" và "không cao lắm" phần nào tương ứng với "rất nặng" và "không nặng lắm".

Định nghĩa 2: Hàm định lượng ngữ nghĩa

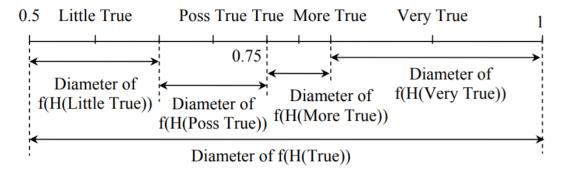
Cho đại số gia tử mở rộng đối xứng AT = (T, G, H, \leq), f: T \rightarrow [0, 1] là một hàm ngữ nghĩa định lượng của AT nếu \forall h, k \in H+ hoặc \forall h, k \in H- và \forall x, y \in T, ta có:

$$\left| \frac{f(hx) - f(x)}{f(kx) - f(x)} \right| = \left| \frac{f(hy) - f(y)}{f(ky) - f(y)} \right|$$

Với đại số gia tử và hàm ngữ nghĩa định lượng, chúng ta có thể định nghĩa một khái niệm rất trừu tượng và khó định nghĩa một cách thoả đáng trong lý thuyết tập mò là tính mò của một khái niệm mò hay của tập mò biểu diễn nó.

Định nghĩa 3: Tính mờ của một giá trị ngôn ngữ

Cho trước một hàm định lượng ngữ nghĩa f của X. Xét bất kỳ $x \in X$, tính mờ của x khi đó được đo bằng đường kính tập $f(H(x)) \subseteq [0, 1]$.



Hình 2.1 Tính mờ của một giá trị ngôn ngữ [1]

Định nghĩa 4: Độ đo tính mờ.

Hàm $fm: T \rightarrow [0, 1]$ được gọi là độ đo tính mờ nếu:

(1)
$$fm(c-) = \theta > 0$$
 $và fm(c+) = 1 - \theta > 0$

Giả sử tập các gia tử $H = H + \cup H -$, trong đó:

$$H -= h_1, h_2, ..., h_p$$
 với $h_1 > h_2 > \cdots > h_p$

$$H += h_{p+1}, h_{p+2}, \dots, h_{p+q}$$
 với $h_{p+1} < h_{p+2} < \dots < h_{p+q}$.

Khi đó:

$$(2)\sum_{i=1}^{p+q} fm(h_i c) = fm(c), c \in \{c^-, c^+\}$$

(3)
$$V \acute{o}im \acute{o}ix, y \in T, h \in H, \frac{fm(hx)}{fm(x)} = \frac{fm(hy)}{fm(y)}$$

Đẳng thức này không phụ thuộc vào các phần tử x, y và do đó ta có thể ký hiệu là $\mu(h)$ và gọi là độ đo tính mờ (fuzziness measure) của gia tử h.

Mệnh đề: Tính chất của fm(x) và $\mu(h)$. Chúng ta có:

$$(1)fm(hx) = \mu(h)fm(x), \forall x \in T.$$

$$(2)\sum_{i=1}^{p+q} fm(h_ic) = fm(c), c \in \{c^-, c^+\}.$$

(3)
$$\sum_{i=1}^{p+q} fm(h_i x) = fm(x).$$

$$(4)\sum_{i=1}^{p}\mu(h_i)=\alpha,\sum_{i=p+1}^{q}\mu(h_i)=\beta,v\acute{o}i\alpha,\beta>0v\grave{a}\alpha+\beta=1.$$

2.1.4 Xây dựng hàm định lượng ngữ nghĩa

Định nghĩa 5: Hàm sgn: $T \rightarrow [-1, 0, 1]$

$$(1) \, sgn(c-) = \, -1 \, v\grave{a} \, sgn(hc-) = \begin{cases} +sgn(c-) \text{ , n\'eu } hc- < c-\\ -sgn(c-) \text{ , n\'eu } hc- \geq c- \end{cases}$$

$$(2) \operatorname{sgn}(c+) = +1 \operatorname{va} \operatorname{sgn}(hc+) = \begin{cases} +\operatorname{sgn}(c+), \operatorname{n\~e}u \ hc+ < c + \\ -\operatorname{sgn}(c+), \operatorname{n\~e}u \ hc+ \ge c + \end{cases}$$

(3)
$$sgn(h'kx) = -sgn(hx)$$
 nếu h' là âm (negative) đối với h và h'hx \neq hx

(4)
$$sgn(h'kx) = +sgn(hx)$$
 nếu h' là dương (positive) đối với h'hx \neq hx

(5)
$$sgn(h'kx) = 0$$
 nếu h'hx = hx

Định nghĩa 6: Hàm định lượng ngữ nghĩa v

Giả sử cho trước độ đo tính mờ của các gia tử $\mu(h)$ và các giá trị độ đo tính mờ của các phần tử sinh fm (c-), fm (c+) và θ là phần tử trung hoà (neutral).

Hàm định lượng ngữ nghĩa v của T được xây dựng như sau với $x=h_{im}\dots h_{i2}h_{i1}c$:

$$(1) v(c-) = \theta - \alpha fm(c-), v(c+) = \theta + \alpha fm(c+)$$

$$(2) fm(x) = fm(h_{im} \dots h_{i2}h_{i1}) = \mu(h_{im}) \dots \mu(h_{i2})\mu(h_{i1})fm(c)$$

(3)
$$v(h_j x) = v(x) + Sgn(h_j x) * [\sum_{i=j}^{p} fm(h_i x) - 0.5 (1 - 1)]$$

$$Sgn(h_jx)Sgn(h_1h_jx)(\beta-\infty))fm(h_jx)$$
 với $j \le p$

(4)
$$v(h_j x) = v(x) + Sgn(h_j x) * [\sum_{i=j}^{p} fm(h_i x) - 0.5 (1 - 1)]$$

$$Sgn(h_ix)Sgn(h_1h_ix)(\beta-\infty))fm(h_ix)$$
 với $j > p$

Phần trình bày dưới đây được tham khảo từ tài liệu [2]

Định nghĩa 7: Hàm IC: $Dom(Ai) \rightarrow [0,1]$

Nếu $LD_{Ai} = \emptyset$ và $D_{Ai} \neq \emptyset$ thì $\forall \omega \in Dom(Ai)$ (với $Dom(Ai) = [\psi_{min}, \psi_{max}]$ là miền trị kinh điển của Ai):

$$IC(\omega) = 1 - \frac{\psi_{max} - \omega}{\psi_{max} - \psi_{min}}$$

Nếu $D_{Ai} \neq \emptyset$, $LD_{Ai} \neq \emptyset$ thì $\forall \omega \in Dom(Ai)$ (với $LD_{Ai} = [\psi_{minLV}, \psi_{maxLV}]$ là miền trị ngôn ngữ của Ai)

$$IC(\omega) = \{\omega * \nu(\psi_{maxLV})\}/\psi_{max}$$

2.1.5 Định lượng giá trị ngôn ngữ khi chỉ biết một đoạn con của $[\psi_{min}, \psi_{max}]$ và toàn bộ $IC(\omega)$

Cho thuộc tính không thuần nhất A_i , lúc này ta có $Dom(A_i) = D_{Ai} \cup LD_{Ai}$ nhưng giá trị biên $[\psi_{min}, \psi_{max}]$ đối với miền trị kinh điển D_{Ai} của $_{Ai}$ không được xác định, mà ta chỉ biết một đoạn con $[\psi_1, \psi_2]$ tương ứng giá trị ngôn ngữ $[\psi_{LV1}, \psi_{LV2}]$ của L_{DAi} và tất cả các giá trị định lượng mờ $IC(\omega)$ của chúng.

Lúc này, do $IC(\omega) = 1 - \frac{\psi_{max} - \omega}{\psi_{max} - \psi_{min}}$ nên tất cả các ω nằm giữa $[\psi_1, \psi_2]$ sẽ đúng với quy tắc này.

Hơn nữa, do độ lớn của các ω sẽ tỷ lệ với bán kính $f(H(x)) \subseteq [0,1]$ tức là $\omega_1 > \omega_2$ khi $IC(\omega_1) > IC(\omega_2)$ và $\frac{\omega^1}{IC(\omega_1)} = \frac{\omega^2}{IC(\omega_2)}$ khi tất cả các $IC(\omega_1)$, $IC(\omega_2)$ về cùng một phía với W. Do vậy, giá trị định lượng cho các giá trị ngôn ngữ này được tính theo giải thuật như sau:

Bước 1: Với ω mà giá trị ngôn ngữ tương ứng trong đoạn [ψ_{LV1} , ψ_{LV2}], ta có:

$$\omega = IC(\omega)(\psi_2 - \psi_1) + \psi_1$$

Bước 2: Với ω mà giá trị ngôn ngữ tương ứng trong đoạn $[\psi_{LV2}, \psi_{maxLV}]$, ta tính tuần tự tăng theo đoạn $\psi_{LV2}..\psi_{maxLV}$ với $\omega_1 = \psi_2 \frac{IC(\omega_2)}{\omega_1}$ và dịch chuyển ψ_{LV2} đến vị trí i vừa tìm được.

Bước 3: Với ω mà giá trị ngôn ngữ tương ứng trong đoạn $[\psi_{\text{minLV}}, \psi_{\text{LV1}}]$, ta tính tuần tự giảm theo đoạn $\psi_{\text{LV1}}..\psi_{\text{minLV}}$ với $\omega_i = \psi \frac{IC(\omega_1)}{IC(\omega_i)}$ và dịch chuyển vị trí ψ_{LV1} lùi về vị trí i vừa tìm được

2.1.6 Định lượng giá trị ngôn ngữ khi chỉ biết một đoạn con của $[\psi_{min}, \psi_{max}]$ nhưng chưa xác định được toàn bộ $IC(\omega)$

Cho thuộc tính không thuần nhất Ai, lúc này ta có $Dom(Ai) = D_{Ai} \cup LD_{Ai}$ nhưng giá trị biên $[\psi_{min}, \psi_{max}]$ đối với miền trị kinh điển D_{Ai} của Ai không được xác định, mà ta cũng chỉ tìm được một đoạn con $[\psi_1, \psi_2]$ của nó tương ứng giá trị ngôn ngữ $[\psi_{LV1}, \psi_{LV2}]$ của LD_{Ai} tức là $v(\psi_{LV1}) = IC(\psi_1)$ và $v(\psi_{LV2}) = IC(\psi_2)$. Lúc này ta phải tìm các giá trị $IC(\omega_i)$ còn lại tức các $IC(\omega_i)$ thỏa $IC(\psi_i) < IC(\psi_1)$ hoặc $IC(\psi_i) > IC(\psi_2)$.

Do $IC(\omega)=1-\frac{\psi_{max}-\omega}{\psi_{max}-\psi_{min}}$ nên tất cả các ω nằm giữa $[\psi_1,\psi_2]$ sẽ đúng với quy tắc này, tức là $IC(\omega)=1-\frac{\psi_2-\omega}{\psi_2-\psi_1}$ với $\omega\in[\psi_1,\psi_2]$. Do vậy có thể xây dựng một ĐSGT để định lượng giá trị cho chúng.

Ta có giải thuật như sau:

Bước 1: Xây dựng 1 ĐSGT trong miền $[\psi_1, \psi_2]$ để tính các $IC(\omega)$ tương ứng cho các giá trị trong đoạn $[\psi_1, \psi_2]$ này.

Bước 2: Tính lại các phân hoạch cho các $IC(\omega)$ như sau :

Nếu $\psi_i < \psi_1$ thì :

- 1. Phân hoạch đoạn $[0, \nu(\psi_1)]$ thành $[0, \nu(\psi_i)]$ và $[\nu(\psi_i), \nu(\psi_1)]$
- 2. $T \sinh f m(h_1) \sim f m(h_1) \times I(\psi_1) \vee h f m(h_1) = f m(h_1) f m(h_1)$

Nếu $\psi_i > \psi_2$ thì :

- 1. Phân hoạch đoạn $[\nu(\psi_2), 1]$ thành $[\nu(\psi_2), \nu(\psi_i)]$ và $[\nu(\psi_i), 1]$
- 2. Tính $fm(hi) \sim fm(h2) \times I(\psi_2) \vee a fm(h_2) = fm(h_2) fm(h_i)$

Tính giá trị $IC(\omega_i)$ và ψ_i tại vị trí i. Gán vị trí i đang có thành vị trí 1 và tiếp tục tính lùi với các giá trị còn lại với $\psi_i < \psi_1$ hay gán vị trí i đang có thành vị trí 2 và tiếp tục tính tiến với các giá trị còn lại với $\psi_i > \psi_2$

Bước 3: Thực hiện định lượng các giá trị ngôn ngữ với cách tính ở mục 1 khi đã biết toàn bộ $IC(\omega)$.

Tính đúng của giải thuật: Do tất cả các phân hoạch trên không vượt ra khỏi đoạn đang xét là $|fm(h_1)|$ hay $|fm(h_2)|$ nên không làm phá vỡ các phân hoạch đang có của đoạn [0,1], do $I(\psi_1)>0$ và $I(\psi_2)$, nên cách phân hoạch trên là phù hợp với phương pháp thuần nhất đã nêu ở mục trên.

2.2 Picture Fuzzy Set

2.2.1 Fuzzy Set

Lý thuyết tập mờ (fuzzy set) là một sự mở rộng của lý thuyết tập hợp cổ điển (classiical set). Trong lý thuyết tập hợp cổ điển, quan hệ giữa một phần tử và một tập hợp chỉ có thể mang giá trị nhị phân: thuộc hoặc không thuộc về tập hợp. Ngược lại, trong lý thuyết tập mờ, mức độ thành viên của một phần tử trong tập mờ được thể hiện bởi một hàm liên thuộc có giá trị thực trong đoạn [0, 1]. Cụ thể:

Một tập mờ A là một cặp (U, m), trong đó, U là một tập hợp cổ điển, còn $m: U \rightarrow [0, 1]$ là hàm liên thuộc. Với mỗi phần tử x thuộc U, giá trị m(x) được gọi là mức độ liên thuộc của phần tử x trong tập mờ (U, m):

- Nếu m(x) = 0, ta nói x không có trong A (no member)
- Nếu m(x) = 1, ta nói x là thành viên toàn phần của A (full member)
- Nếu 0 < m(x) < 1, ta nói x là thành viên mở của A (fuzzy member)

Các phép toán trên tập mờ

Kí hiệu t là một t-norm, và s là s-norm tương ứng, khi đó ta có:

(1)
$$\forall x \in U: \mu_{\neg A}(x) = 1 - \mu_A(x)$$

$$(2) \,\forall x \in U: \mu_{A \cap B}(x) = t\big(\mu_A(x), \mu_B(x)\big)$$

$$(3) \forall x \in U: \mu_{A \cup B}(x) = s(\mu_A(x), \mu_B(x))$$

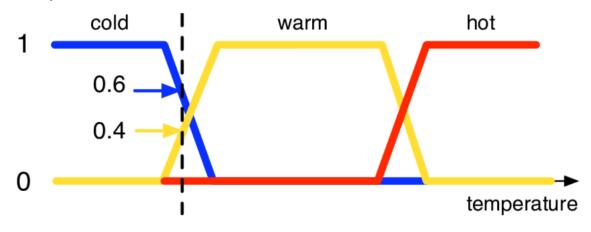
$$(4) \forall x \in U: \mu_{A \setminus B}(x) = \min(\mu_A(x), 1 - \mu_B(x))$$

2.2.2 Fuzzly Logic

Logic mờ (Fuzzy logic) được phát triển từ lý thuyết tập mờ để thực hiện lập luận một cách xấp xỉ thay vì lập luận chính xác theo logic vị từ cổ điển. Logic mờ rất phù hợp với cách tư duy, suy nghĩ của con người, bởi con người thường ra quyết định dựa trên những thông tin không hoàn toàn rõ ràng, hoặc các thông tin không phải ở dạng con số cụ thể (vd: các giá trị ngôn ngữ).

Logic mờ cho phép độ liên thuộc có giá trị trong khoảng đóng 0 và 1, và ở hình thức ngôn từ, nó cho phép sử dụng các khái niệm không chính xác như "hơi hơi", "gần như", "khá là" và "rất". Lôgic mờ có thể được coi là mặt ứng dụng của lý thuyết tập mờ để xử lý các giá trị trong thế giới thực cho các bài toán phức tạp.

Ví dụ:



Hình 2.2 Minh họa Fuzzy Logic [3]

Trong hình, cold (lạnh), warm (ấm), và hot (nóng) là các hàm ánh xạ một thang nhiệt độ. Một điểm trên thang nhiệt độ có 3 "chân giá trị" — mỗi hàm cho một giá trị. Đối với nhiệt độ cụ thể trong hình, 3 chân giá trị này có thể được giải nghĩa là 3 miêu tả sau về nhiệt độ này: "tương đối lạnh", "hơi hơi ấm", và "không nóng".

2.2.3 Picture Fuzzy Set

Phần trình bày dưới đây được tham khảo từ tài liệu [4]

Trong bài trình bày này, U được biểu thị là một tập (khác rỗng) được gọi là Universe of Discourse. Lớp các tập con của U sẽ được biểu thị bằng P(U) và lớp các tập con mờ (fuzzy) của U sẽ được biểu thị bằng F(U).

Định nghĩa 1: Picture fuzzy set:

(Bui và Kreinovick, 2013) Một picture fuzzy set A trên U là một đối tượng có dạng:

$$A = \{(x, \mu_A(x), \eta_A(x), \gamma_A(x)) | x \in U \}$$

trong đó:

 $\mu(x) \in [0,1]$, là "độ tích cực của x trong A"

 $\eta(x) \in [0,1]$, "độ trung lập của x trong A"

 $y(x) \in [0,1]$ "mức độ phủ định của x trong A"

và μ, η và γ thỏa mãn điều kiện sau:

$$\mu_A(x) + \eta_A(X) + \gamma_A(x) \le 1, \forall x \in X$$

Một picture fuzzy set liên kết 3 tập mờ được xác định như sau: $\mu_A : U \to [0,1]$, $\eta_A : U \to [0,1]$ và $\gamma_A : U \to [0,1]$ và có thể biểu diễn bằng một bộ $(\mu_A, \eta_A, \gamma_A)$.

Toán tử trên PFS (U) đã được giới thiệu: $\forall A, B \in PFS(U)$

- $A \subseteq B \ iff \ \mu_A(x) \le \mu_B, \eta_A(x) \le \eta_B(x) \ and \ \gamma_A(x) \ge \gamma_B(x), \forall x \in U$
- $A = B iff A \subseteq B and B \subseteq A$
- $A \cup B =$ { $(x, max(\mu A(x), \mu B(x)), min(\eta A(x), \eta B(x), min(\gamma A(x), \gamma B(x))) | x \in U$ }
- $A \cap B =$ { $(x, min(\mu A(x), \mu B(x)), min(\eta A(x), \eta B(x), max(\gamma A(x), \gamma B(x))) | x \in U$ }

Định nghĩa 2: Picture Fuzzy Relation

U là một tập vũ trụ khác rỗng . Một quan hệ mờ bức tranh (Picture Fuzzy Relation) từ $U{\to}V$ là một bộ Picture Fuzzy trên $U{\times}V$ và biểu thị bởi $R(U{\to}V)$, tức là, là một biểu thức được đưa ra bởi

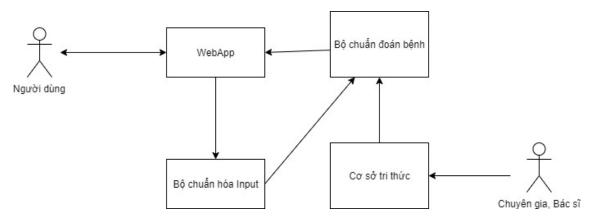
$$R = \{((x, y), \mu_R(x, y), \eta_R(x, y), \gamma_R(x, yy)) | (x, y) \in U \times V\}$$

trong đó:

$$\mu_R, \gamma_R, \eta_R \colon U \times V \to [0,1] \colon \mu_R(x,y) + \eta_R(x,y) + \gamma_R(x,y) \le 1, \forall (x,y) \in U \times V.$$

CHƯƠNG 3. MÔ HÌNH KIẾN TRÚC HỆ THỐNG

3.1 Mô hình hệ thống



Hình 3.1 Mô hình kiến trúc hệ thống

Webapp: Giao diện tương tác với người dùng. Cho phép người dùng nhập các giá trị cho các triệu chứng.

Cơ sở tri thức: Biểu diễn tri thức dưới dạng mạng ngữ nghĩa, và lưu trữ bằng csdl quan hệ.

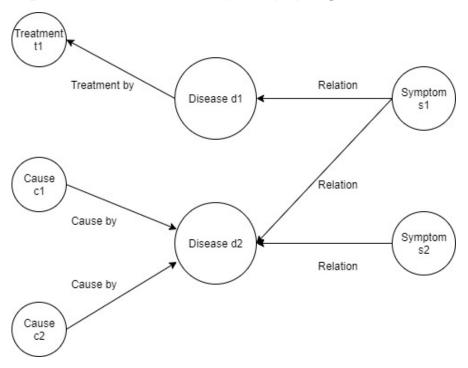
Bộ chuẩn hóa input: Chuyển đổi các giá trị ngôn ngữ sang giá trị số thuộc miền [0, 1].

Bộ chuẩn đoán bệnh: Kết hợp các giá trị triệu chứng mà bệnh nhân nhập vào với tri thức trong cơ sở tri thức để tính toán kết quả và đưa ra kết luận.

CHƯƠNG 4. MÔ TẢ THUẬT TOÁN

Bước 1: Biểu diễn tri thức

Step 1: Biểu diễn tri thức bằng mạng ngữ nghĩa



Hình 4.1 Biểu diễn tri thức bằng mạng ngữ nghĩa

Các loai nút:

1. **Disease**: Nút bênh

2. **Symptom**: Nút triệu chứng

3. Cause: Nút nguyên nhân gây ra bệnh4. Treatment: Nút cách điều tri bênh

Các loại cung:

 Relation: Biểu thị mối quan hệ giữa triệu chứng s và bệnh d dưới dạng một bộ ba số PFS:

$$Q(s, d) = (\mu, \eta, \gamma)$$

Trong đó:

 $\mu\text{:}\ M\text{\'{u}}\text{c}$ độ ủng hộ giả thiết bệnh nhân mắc bệnh d khi có triệu chứng s

γ: Mức độ phủ định giả thiết bệnh nhân mắc bệnh d khi có triệu chứng s

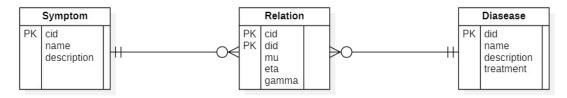
 $\eta \colon M\text{\'e} c$ độ trung lập của triệu chứng s với bệnh d

 $1-(\mu+\eta+\gamma)$: Sự thiếu thông tin về mối quan hệ giữa triệu chứng s và bệnh d

Ví dụ: Bệnh nhân gặp triệu chứng s: "nhìn xa bị mờ", suy ra s ủng hộ giả thiết bệnh nhân bị bệnh "Cận thị", phủ định giả thiết bệnh nhân bị bệnh "Viễn thị", và giữ trung lập với giả thiết bệnh nhân bị "Đau dạ dày".

- 2. Cause by: Biểu thị mối quan hệ: c là nguyên nhân gây ra bệnh d
- 3. Treament by: Biểu thị mối quan hệ: t là các điều trị của bệnh d

Step2: Lưu trữ tri thức bằng cơ sở dữ liệu quan hệ



Hình 4.2 Lưu trữ tri thức bằng csdl quan hệ

Bước 2: Chuẩn hóa input

Khi sử dụng hệ thống, bệnh nhân p sẽ được hỏi về cảm nhận của từng triệu chứng s1, s2, s3.... Với mỗi triệu chứng s, bệnh nhân p sẽ nhập vào các giá trị cho 3 thành phần tương ứng với positive, neutraul, negative của bộ PFS dưới dạng các giá trị ngôn ngữ. Các giá trị này sẽ được chuẩn hóa về miền [0, 1] bằng cách sử dụng phương pháp đại số gia tử. Khi đó, mối quan hệ giữa bệnh nhân p và triệu chứng s được biểu thị thông qua:

$$R(p, s) = (\mu, \eta, \gamma)$$

Trong đó:

 $\mu\text{:}\ M\text{\'{u}}\text{c}\ \text{độ bệnh nhân p có triệu chứng s}$

γ: Mức độ phủ định bệnh nhân p có triệu chứng s

 η : Mức độ trung lập của bệnh nhân p với triệu chứng s

 $1 - (\mu + \eta + \gamma)$: Sự thiếu thông tin về mối quan hệ giữa bệnh nhân p triệu chứng s Ví dụ: Khi bệnh nhân được hỏi về triệu chứng đau đầu:

- Nếu bệnh nhân cảm thấy rất đau đầu, đau thường xuyên \rightarrow giá trị μ sẽ cao
- Nếu bệnh nhân cảm thấy không hề có dấu hiệu đau đầu \rightarrow giá trị γ sẽ cao
- Nếu bệnh nhân không khẳng định, nhưng cũng không phủ định \rightarrow giá trị η sẽ cao
- Nếu bệnh nhân từ chối trả lời \rightarrow giá trị $1 (\mu + \eta + \gamma)$ sẽ cao

Bước 3: Chuẩn đoán bệnh

Sau khi chuẩn hóa các giá trị input của người dùng, ta thu được mối quan hệ của bệnh nhân với các triệu chứng R(p, s), kết hợp với các tri thức được thể hiện thông qua mối quan hệ giữa triệu chứng với bệnh Q(s, d), ta xây dựng mối quan hệ giữa bệnh nhân và bênh như sau:

$$T(p, d) = R(p, s) \circ Q(s, d)$$

Khi này, bộ PFS $T(p, d) = (\mu_T, \eta_T, \gamma_T)$ được tính như sau:

$$\mu_T(p,d) = \max_{s \in S} \min \{ \mu_R(p,s), \mu_Q(s,d) \}$$

$$\eta_T(p,d) = \min_{s \in S} \min \eta_R(p,s), \eta_Q(s,d)$$

$$\gamma_T(p,d) = \min_{s \in S} \max \chi_R(p,s), \gamma_Q(s,d)$$

$$\forall p \in P, d \in D.$$

Cuối cùng, ta tính toán xác suất bệnh nhân p mắc phải bệnh d bằng công thức:

$$S_T(p,d)=\mu_T(p,d)-\gamma_T(p,d)^*\,\pi_T(p,d)$$
 Với $\pi_T(p,d)=1-[\mu_T(p,d)+\eta_T(p,d)+\gamma_T(p,d)]$

Nếu giá trị $S_T(p,d) >= 0.5$, ta kết luận bệnh nhân p có mắc bệnh d, ngược lại, ta kết luận bệnh nhân p không mắc bệnh d.

CHƯƠNG 5. VÍ DỤ

Ví dụ 1: Xây dựng hàm định lượng ngữ nghĩa cho Đại số gia tử AT Xét AT = $\{T, G, H, \leq\}$ với:

- T = {cao, thấp, hơi cao, rất cao, rất rất cao, hơi thấp ...}
- $G = \{cao, th\hat{a}p\}$
- $\blacksquare \quad \mathbf{H} = \{\mathbf{hoi}, \, \mathbf{rat}\}$
- ≤: rất thấp < thấp < ... < cao < rất cao < ...

Như vậy, ta có

- $c-=th\hat{a}p, c+=cao$
- $H = \{hoi\}, H = \{r\acute{a}t\}$

Bước 1: Thiết lập các tham số ban đầu

fm(cao)	fm(thấp)	μ(cao)	$\mu(ext{th\'ap})$	θ
0.5	0.5	0.5	0.5	0.5

Từ đó, ta tính được

$$\alpha = \sum_{h=0}^{\epsilon H^{-}} \mu(h) = 0.5$$

$$\beta = \sum_{h=0}^{\epsilon H+} \mu(h) = 0.5$$

Bước 2: Thực hiện định lượng một số giá trị ngôn ngữ

$$v(cao) = \theta + \alpha. fm(cao) = 0.5 + 0.5 \times 0.5 = 0.75$$

$$v(th\tilde{a}p) = \theta - \alpha . fm(th\tilde{a}p) = 0.5 - 0.5 \times 0.5 = 0.25$$

 $v(r ilde{a} t cao)$ được tính theo công thức:

$$v(h_j x) = v(x) + sign(h_j x) \{ \sum_{i=sign(j)}^{j} fm(h_i x) - \omega(h_j x) fm(h_j x) \}$$

với

$$\omega(h_jx) = \frac{1}{2}[1 + Sign(h_jx)sign(h_ph_jx)(\beta - \alpha)] \in \{\alpha,\beta\} \,,\, \mathbf{j} \in [-\mathbf{q}^\mathbf{p}],\, \mathbf{j} \neq \mathbf{0}.$$

Ở đây:

j	X	v(x)	sign(h _j x)	$fm(h_jx) = \mu(h_jx).fm(x)$	$\omega(h_j x)$
1	cao	0.75	+1	$0.5 \times 0.5 = 0.25$	0.5

Do đó, ta có:
$$v(r \tilde{a}t \ cao) = 0.75 + (+1)[0.25 - 0.5 \times 0.25) = 0.875$$

Tương tự, ta tính được:

$$v(r\tilde{a}t \ th\tilde{a}p) = 0.125, v(h\sigma i \ cao) = 0.625, v(h\sigma i \ th\tilde{a}p) = 0.375$$

Bước 3: Tổng hợp kết quả

	Rất	Thấp		Trung	Ноі	Cao	Rất	
	thấp		thấp	bình	cao		cao	
0	0.125	0.25	0.375	0.5	0.625	0.75	0.875	1

Ví dụ 2: Chuẩn đoán bệnh

Giả sử sau khi thu thập thông tin người dùng nhập vào và chuẩn hóa input, ta được bảng dữ liệu sau:

Bảng 5.1 Quan hệ giữa bệnh nhân và các triệu chứng

D	Bệnh nhân p				
R	μ_{R}	η_R	γ_{R}		
Thân nhiệt cao	0.8	0.1	0.05		
Đau đầu	0.4	0.05	0.1		
Đau bụng	0.05		0.7		
Но	0.6	0.05	0.1		
Đau ngực	0.2	0.4	0.2		

Giả sử tri thức về các triệu chứng và các bệnh được thể hiện như sau:

Bảng 5.2: Quan hệ giữa các triệu chứng và các bệnh

	Cảm cúm				Bệnh		
Q	μ_Q	$\eta_{\it Q}$	γ_Q	μ_Q	$\eta_{\it Q}$	γ_Q	
Thân nhiệt cao	0.8	0.05	0.05	0.2	0.4	0.1	
Đau đầu	0.6	0.1	0.1	0.2	0.5	0.1	
Đau bụng	0.1	0.5	0.2	0.2	0.6	0.05	
Но	0.5	0.1	0.1	0.8	0.05	0.1	
Đau ngực	0.4	0.05	0.1	0.75	0.05	0.05	

Khi đó ta đi tính quan hệ T = R * Q. Quan hệ này thể hiện mối quan hệ giữa bệnh nhân với các bệnh.

Bång 5.3: Quan hệ T = R * Q

	Cảm cúm				Bệnh		
T	min	min	max	min	min	max	
	(μ_R, μ_Q)	(η_R, η_Q)	(γ_R, γ_Q)	(μ_R, μ_Q)	(η_R, η_Q)	(γ_R,γ_Q)	
Thân nhiệt cao	0.8	0.05	0.05	0.2	0.1	0.1	
Đau đầu	0.4	0.05	0.1	0.2	0.05	0.1	
Đau bụng	0.05	0.1	0.7	0.05	0.1	0.7	
Но	0.5	0.05	0.1	0.6	0.05	0.1	
Đau ngực	0.2	0.05	0.1	0.2	0.05	0.1	

 $T(p, Cåm c\'um) = (\mu_{Tmax}, \eta_{Tmin}, \gamma_{Tmin}) = (0.8, 0.05, 0.05)$

T(p, Lao phổi) = $(\mu_{Tmax}, \eta_{Tmin}, \gamma_{Tmin}) = (0.6, 0.05, 0.1)$

Áp dụng công thức $S_T(p,d) = \mu_T(p,d) - \gamma_T(p,d)^* \pi_T(p,d)$, ta tính được:

$$S_T(p, \text{Cåm c\'um}) = 0.8 - 0.05(1 - 0.8 - 0.05 - 0.05) = 0.795$$

$$S_T(p, Lao phổi) = 0.5 - 0.1(1 - 0.6 - 0.05 - 0.1) = 0.575$$

 $\mathring{\mathrm{O}}$ đây, ta quy ước ST >= 0.5 thì kết luận bệnh nhân mắc bệnh, do vậy:

Kết luận: Bệnh nhân p mắc bệnh Cảm cúm với xác suất 0.795, và mắc bệnh Lao phổi với xác suất 0.575

CHƯƠNG 6. KẾT LUẬN

6.1 Kết luận

Mặc dù trong quá trình tìm hiểu lý thuyết cũng như bắt tay vào xây dựng hệ thống còn gặp phải không ít các khó khăn, nhưng về cơ bản việc thực hiện đề tài đã đạt được các mục tiêu đã đề ra.

Thông qua quá trình tìm hiểu và xây dựng, chúng em đã thu được nhiều kiến thức bổ ích mới, học được nhiều kỹ năng trong việc làm việc nhóm, tìm kiếm tài liệu, tổng hợp thông tin, kỹ năng trình bày, báo cáo.

6.2 Hướng phát triển trong tương lai

Trong tương lai, hệ thống cần được hoàn thiện hơn về khai thác và sử dụng tri thức, cũng như các chức năng bổ sung, cập nhật tri thức. Ngoài ra, để có thể gần gũi, tiện ích hơn với người dùng, hệ thống có thể cần được triển khai thêm trên các nền tảng mobile như Android và iOS.

TÀI LIỆU THAM KHẢO

```
N. T. Duy, "Đại số gia tử và suy luân mờ". Khoa Điện tử - Hội thảo khoa học
1 2007.
1
[ N. M. H. N. C. H. Lê Văn Tường Lân, "MỘT PHƯƠNG PHÁP ĐỊNH
2 LƯỢNG GIÁ TRỊ NGÔN NGỮ TRONG TẬP MẪU HUẨN LUYỆN ĐỂ
] XÂY DUNG CÂY QUYẾT ĐỊNH MÒ".
[ "wikipedia.org,"
                                  [Online].
                                                             Available:
3 https://vi.wikipedia.org/wiki/Logic_m%E1%BB%9D.
1
[ Pham Hong Phong, Hieu Trong Dinh, "Research Gate," [Online]. Available: •
4 https://www.researchgate.net/profile/Hieu_Dinh/publication/266084039_So
] me_compositions_of_picture_fuzzy_relations/links/542520b40cf238c6ea73c
 d7a/Some-compositions-of-picture-fuzzy-relations.pdf.
[ "Fuzzy set," [Online]. Available: https://en.wikipedia.org/wiki/Fuzzy_set.
5
1
[ "wikipedia.org,"
                                  [Online].
                                                             Available:
6 https://vi.wikipedia.org/wiki/T%E1%BA%ADp_m%E1%BB%9D.
```