

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**

**Ứng 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**

[CHƯƠNG 1. TỔNG QUAN ĐỀ TÀI 1](#_Toc28094330)

[1.1 Đặt vấn đề 1](#_Toc28094331)

[1.2 Mục tiêu và nhiệm vụ đề tài 2](#_Toc28094332)

[1.3 Phân công công việc 2](#_Toc28094333)

[CHƯƠNG 2. CƠ SỞ LÝ THUYẾT 3](#_Toc28094334)

[2.1 Đại số gia tử 3](#_Toc28094335)

[2.1.1 Giới thiệu đại số gia tử 3](#_Toc28094336)

[2.1.2 Định nghĩa đại số gia tử 4](#_Toc28094337)

[2.1.3 Các đại lượng đo trên đại số gia tử 5](#_Toc28094338)

[2.1.4 Xây dựng hàm định lượng ngữ nghĩa 7](#_Toc28094339)

[2.1.5 Định lượng giá trị ngôn ngữ khi chỉ biết một đoạn con của [ψmin, ψmax] và toàn bộ IC(ω) 8](#_Toc28094340)

[2.1.6 Định lượng giá trị ngôn ngữ khi chỉ biết một đoạn con của [ψmin, ψmax] nhưng chưa xác định được toàn bộ IC(ω) 9](#_Toc28094341)

[2.2 Picture Fuzzy Set 10](#_Toc28094342)

[2.2.1 Fuzzy Set 10](#_Toc28094343)

[2.2.2 Fuzzly Logic 11](#_Toc28094344)

[2.2.3 Picture Fuzzy Set 12](#_Toc28094345)

[CHƯƠNG 3. MÔ HÌNH KIẾN TRÚC HỆ THỐNG 13](#_Toc28094346)

[3.1 Mô hình hệ thống 13](#_Toc28094347)

[CHƯƠNG 4. MÔ TẢ THUẬT TOÁN 14](#_Toc28094348)

[Bước 1: Biểu diễn tri thức 14](#_Toc28094349)

[Step 1: Biểu diễn tri thức bằng mạng ngữ nghĩa 14](#_Toc28094350)

[Step2: Lưu trữ tri thức bằng cơ sở dữ liệu quan hệ 15](#_Toc28094351)

[Bước 2: Chuẩn hóa input 15](#_Toc28094352)

[Bước 3: Chuẩn đoán bệnh 15](#_Toc28094353)

[CHƯƠNG 5. VÍ DỤ 17](#_Toc28094354)

[**Ví dụ 1:** Xây dựng hàm định lượng ngữ nghĩa cho Đại số gia tử AT 17](#_Toc28094355)

[**Ví dụ 2:** Chuẩn đoán bệnh 19](#_Toc28094356)

[CHƯƠNG 6. KẾT LUẬN 21](#_Toc28094357)

[6.1 Kết luận 21](#_Toc28094358)

[6.2 Hướng phát triển trong tương lai 21](#_Toc28094359)

[TÀI LIỆU THAM KHẢO 22](#_Toc28094360)

**DANH MỤC HÌNH VẼ**

[Hình 2.1 Tính mờ của một giá trị ngôn ngữ [1] 5](#_Toc28036959)

[Hình 2.2 Minh họa Fuzzy Logic [3] 11](#_Toc28036960)

[Hình 3.1 Mô hình kiến trúc hệ thống 13](#_Toc28036961)

[Hình 3.2 Biểu diễn tri thức bằng mạng ngữ nghĩa 14](#_Toc28036962)

[Hình 3.3 Lưu trữ tri thức bằng csdl quan hệ 15](#_Toc28036963)

**DANH MỤC BẢNG BIỂU**

[Bảng 5.1 Quan hệ giữa bệnh nhân và các triệu chứng 19](#_Toc28036964)

[Bảng 5.2: Quan hệ giữa các triệu chứng và các bệnh 19](#_Toc28036965)

[Bảng 5.3: Quan hệ T = R \* Q 20](#_Toc28036966)

# TỔNG QUAN ĐỀ TÀI

## 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.

## 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

## Phân công công việc

|  |  |
| --- | --- |
| Nguyễn Lưu Nhật (nhóm trưởng) | - Tìm kiếm tài liệu  - 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 |

# CƠ SỞ LÝ THUYẾT

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

## Đại số gia tử

*Phần trình bày dưới đây được tham khảo từ tài liệu* [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, ≤), 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).
* ≤: Là biểu thị quan hệ thứ tự trên các từ (các khái niệm mờ), nó được “cảm sinh” từ ngữ nghĩa tự nhiên.

Ví dụ: dựa trên ngữ nghĩa ta có: false ≤ true, more true ≤ very true, very false ≤ more false…

Các quy ước:

1. Các gia tử trong H là các **toán tử thứ tự**, nghĩa là

*(∀h ∈ H, h: T → T), (∀x ∈ T) {hx ≤ x hoặc hx ≥ x}.*

1. Hai gia tử h, k ∈ H được gọi là **ngược nhau** nếu

*(∀x ∈ T) {hx ≤ x khi và chỉ khi kx ≥ x}.*

1. Hai gia tử h, k ∈ H được gọi là **tương thích** nhau nếu

*(∀x ∈ T) {hx ≤ x khi và chỉ khi kx ≤ x}.*

1. Ta ký hiệu **h ≥ k** nếu h, k tương thích nhau và

*(∀x ∈ T) {hx ≤ kx ≤ x hoặc hx ≥ kx ≥ x}.*

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

1. Giả sử V (Very) là phần tử lớn nhất trong thì:

phần tử sinh ***g*** *∈ G* là **dương** nếu ***g ≤ Vg*** và là **âm** nếu ***g ≥ Vg***.

1. Một gia tử **h** **dương** đối với một gia tử **k** nếu:

*(∀x ∈ T) {hkx ≤ kx ≤ x hoặc hkx ≥ kx ≥ x}.*

1. T được sinh ra từ G bởi các gia tử trong , như vậy:

Mỗi phần tử của T sẽ có dạng biểu diễn là .

1. 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*.

### Định nghĩa đại số gia tử

Một cấu trúc đại số AT = (T, G, H, ≤) 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 ∉ H(v)* và *v ∉ H(u)*, thì *∀x ∈ H(u) {x ∉ H(v)}.*
3. Nếu *x ≠ hx* thì *x ∉ H*(hx) và nếu *h ≠ k* và *hx ≤ kx* thì *h’hx ≤ k’kx*, với mọi gia tử *h, k, h’* và *k’*. Hơn nữa nếu *hx ≠ kx* thì *hx* và *kx* là độc lập.
4. Nếu *u ∉ H(v)* và *u ≤ v* (hoặc *u ≥ v*) thì *u ≤ hv* (hoặc *u ≥ 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 ∈ 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 , w ≠ g ∈ G, sao cho , với w ≠ g ’∈ G và g’ ≠ 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.

### Các đại lượng đo trên đại số gia tử

* + - 1. Hàm đo trên đại số gia tử:

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

1. *∀ t ∈ T: f(t) ∈ [0, 1], f(g+) = 1, f(g-) = 0.*
2. *∀ x, y ∈ 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”.

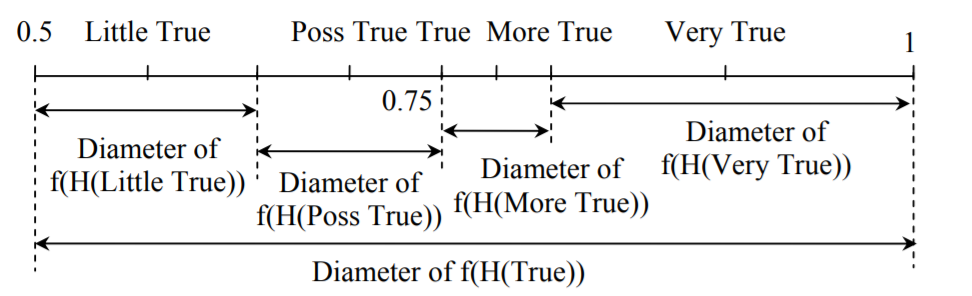
* + - 1. Hàm định lượng ngữ nghĩa

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

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 mờ là tính mờ của một khái niệm mờ hay của tập mờ biểu diễn nó.

* + - 1. Tính mờ của một giá trị ngôn ngữ

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



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

* + - 1. Độ đo tính mờ.

Hàm *fm*: T → [0, 1] được gọi là độ đo tính mờ nếu:

(1) fm(c- ) = θ > 0 và fm(c+ ) = 1 – θ > 0

Giả sử tập các gia tử , trong đó:

với

với .

Khi đó:

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

**Mệnh đề:** Tính chất của *fm(x)* và *μ(h).* Chúng ta có:

### Xây dựng hàm định lượng ngữ nghĩa

* + - 1. Hàm sgn: T→[-1, 0, 1]

1. nếu h’ là âm (negative) đối với h và h’hx hx
2. nếu h’ là dương (positive) đối với h’hx hx
3. nếu h’hx = hx
   * + 1. Hàm định lượng ngữ nghĩa v

Giả sử cho trước độ đo tính mờ của các gia tử *μ(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 :



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

* + - 1. Hàm IC: Dom(Ai) → [0,1]

Nếu LDAi = ∅ và DAi ≠ ∅ thì ∀ω ∈ Dom(Ai) (với Dom(Ai) = [ψmin, ψmax] là miền trị kinh điển của Ai):

Nếu DAi ≠ ∅, LDAi ≠ ∅ thì ∀ω ∈ Dom(Ai) (với LDAi = [ψminLV, ψmaxLV] là miền trị ngôn ngữ của Ai)

*IC(ω) = {ω \* ν (ψmaxLV)}/ψmax*

### Định lượng giá trị ngôn ngữ khi chỉ biết một đoạn con của [ψmin, ψmax] và toàn bộ IC(ω)

Cho thuộc tính không thuần nhất Ai , lúc này ta có Dom(Ai) = DAi ∪ LDAi nhưng giá trị biên [ψmin, ψmax] đối với miền trị kinh điển DAi của Ai không được xác định, mà ta chỉ biết một đoạn con [ψ1, ψ2] tương ứng giá trị ngôn ngữ [ψLV1, ψLV2] của LDAi và tất cả các giá trị định lượng mờ IC(ω) của chúng.

Lúc này, do nên tất cả các ω nằm giữa [ψ1, ψ2] sẽ đúng với quy tắc này.

Hơn nữa, do độ lớn của các ω sẽ tỷ lệ với bán kính f(H(x)) ⊆ [0,1] tức là ω1 > ω2 khi IC(ω1) > IC(ω2) và khi tất cả các IC(ω1), IC(ω2) 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 [ψLV1, ψLV2], ta có:

Bước 2: Với ω mà giá trị ngôn ngữ tương ứng trong đoạn [ψLV2, ψmaxLV], ta tính tuần tự tăng theo đoạn ψLV2..ψmaxLV với và dịch chuyển ψLV2 đến vị trí i vừa tìm được.

Bước 3: Với ω mà giá trị ngôn ngữ tương ứng trong đoạn [ψminLV, ψLV1], ta tính tuần tự giảm theo đoạn ψLV1..ψminLV  với và dịch chuyển vị trí ψLV1 lùi về vị trí i vừa tìm được

### Định lượng giá trị ngôn ngữ khi chỉ biết một đoạn con của [ψmin, ψmax] nhưng chưa xác định được toàn bộ IC(ω)

Cho thuộc tính không thuần nhất Ai, lúc này ta có Dom(Ai) = DAi ∪LDAi nhưng giá trị biên [ψmin, ψmax] đối với miền trị kinh điển DAi của Ai không được xác định, mà ta cũng chỉ tìm được một đoạn con [ψ1, ψ2] của nó tương ứng giá trị ngôn ngữ [ψLV1, ψLV2] của LDAi tức là ν(ψLV1) = IC(ψ1) và ν(ψLV2) = IC(ψ2). Lúc này ta phải tìm các giá trị IC(ωi) còn lại tức các IC(ωi) thỏa IC(ψi) < IC(ψ1) hoặc IC(ψi) > IC(ψ2).

Do nên tất cả các ω nằm giữa [ψ1, ψ2] sẽ đúng với quy tắc này, tức là với . 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 [ψ1, ψ2] để tính các IC(ω) tương ứng cho các giá trị trong đoạn [ψ1, ψ2] này.

Bước 2: Tính lại các phân hoạch cho các IC(ω) như sau :

Nếu ψi < ψ1 thì :

1. Phân hoạch đoạn [0, ν(ψ1)] thành [0, ν(ψi)] và [ν(ψi), ν(ψ1)]
2. Tính *fm*(hi) ~ *fm*(h1) x I(ψ1) và *fm*(h1) = *fm*(h1) - *fm*(hi)

Nếu ψi > ψ2 thì :

1. Phân hoạch đoạn [ν(ψ2), 1] thành [ν(ψ2), ν(ψi)] và [ν(ψi), 1]
2. Tính *fm*(hi) ~ *fm*(h2) x I(ψ2) và *fm*(h2) = *fm*(h2) - *fm*(hi)

Tính giá trị IC(ωi) và ψi 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 ψi < ψ1 hay gán vị trí i đang có thành vị trí 2 và tiếp tục tính tiến với các giá trị còn lại với ψi > ψ2

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

Tính đúng của giải thuật: Do tất cả các phân hoạch trên không vượt ra khỏi đoạn đang xét là |*fm*(h1)| hay |*fm*(h2)| nên không làm phá vỡ các phân hoạch đang có của đoạn [0,1], do I(ψ1)>0 và I(ψ2), nên cách phân hoạch trên là phù hợp với phương pháp thuần nhất đã nêu ở mục trên.

## Picture Fuzzy Set

### 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→ [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 mờ của A (fuzzy member)

Các phép toán trên tập mờ

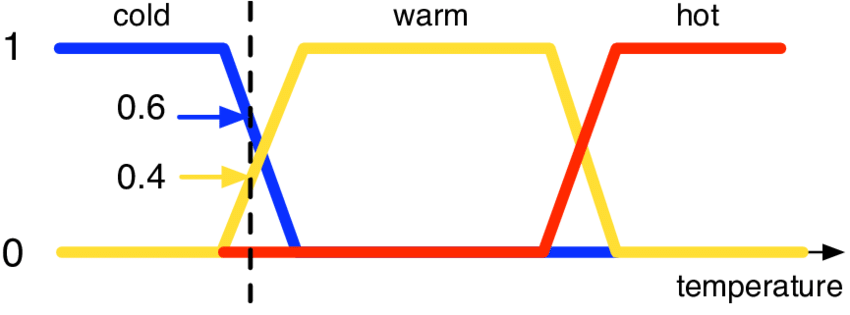
Kí hiệu *t* là một t-norm, và *s* là s-norm tương ứng, khi đó ta có:

### 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".

### 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).

* + - 1. Picture fuzzy set:

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

*A = {(x , μA (x), ηA (x), γA (x))|x ∈U }*

trong đó:

*μ (x)* ∈ [0,1], là “ độ tích cực của x trong A”

*η (x)* ∈ [0,1], “độ trung lập của x trong A”

*γ (x)* ∈ [0,1] “mức độ phủ định của x trong A”

và μ, η và γ thỏa mãn điều kiện sau:

Một picture fuzzy set liên kết 3 tập mờ được xác định như sau: μA :U → [0,1], ηA:U → [0,1] và γA :U → [0,1] và có thể biểu diễn bằng một bộ (μA , ηA , γA).

Toán tử trên PFS (U) đã được giới thiệu: *∀A, B ∈ PFS(U)*

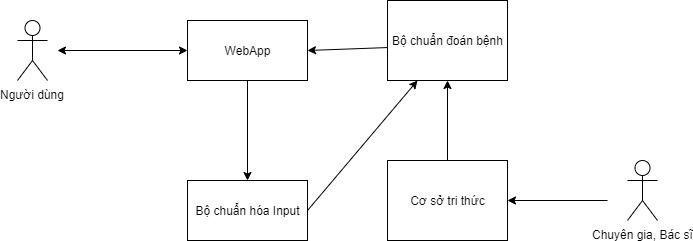
* + - 1. Picture Fuzzy Relation

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

trong đó:

# MÔ HÌNH KIẾN TRÚC HỆ THỐNG

## 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.

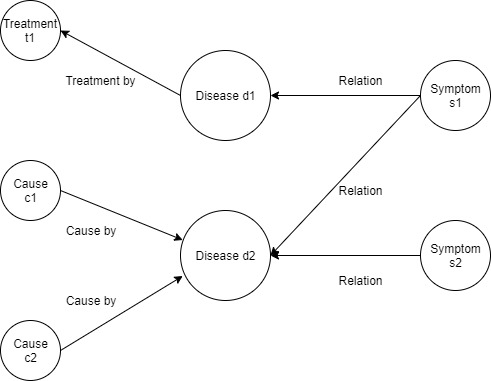
**Bộ chuẩn hóa input**: Chuyển đổi các giá trị ngôn ngữ sang giá trị số thuộc miền [0, 1].

**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.

# 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 loại nút:

1. **Disease**: Nút bệnh
2. **Symptom**: Nút triệu chứng
3. **Cause**: Nút nguyên nhân gây ra bệnh
4. **Treatment**: Nút cách điều trị bệnh

Các loại cung:

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

Q(s, d) = ()

Trong đó:

: Mức độ ủng hộ giả thiết bệnh nhân mắc bệnh d khi có triệu chứng s

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

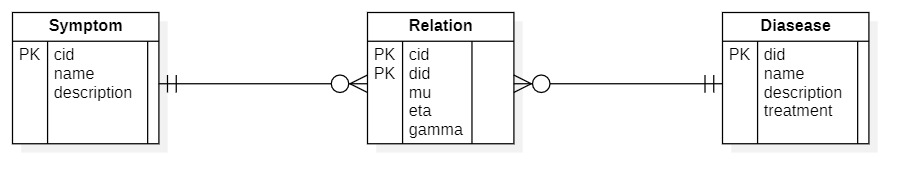
: Mức độ trung lập của triệu chứng s với bệnh d

1 – ( + + ): 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”.

1. **Cause by**: Biểu thị mối quan hệ: c là nguyên nhân gây ra bệnh d
2. **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) = ()

Trong đó:

: Mức độ bệnh nhân p có triệu chứng s

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

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

1 – ( + + ): 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 giá trị sẽ cao
* Nếu bệnh nhân cảm thấy không hề có dấu hiệu đau đầu giá trị sẽ cao
* Nếu bệnh nhân không khẳng định, nhưng cũng không phủ định giá trị sẽ cao
* Nếu bệnh nhân từ chối trả lời giá trị 1 – ( + + ) 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 bệnh như sau:

T(p, d) = R(p, s) Q(s, d)

Khi này, bộ PFS T(p, d) = () được tính như sau:

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

\*

Với

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

# VÍ DỤ

**Ví dụ 1:** Xây dựng hàm định lượng ngữ nghĩa cho Đại số gia tử AT

Xét AT = {T, G, H, ≤} với:

* T = {cao, thấp, hơi cao, rất cao, rất rất cao, hơi thấp …}
* G = {cao, thấp}
* H = {hơi, rất}
* ≤: rất thấp < thấp < … < cao < rất cao < …

Như vậy, ta có

* c- = thấp, c+ = cao
* H- = {hơi}, H+ = {rất}

Bước 1: Thiết lập các tham số ban đầu

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *fm*(cao) | *fm*(thấp) | (cao) | (thấp) |  |
| 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |

Từ đó, ta tính được

= 0.5

= 0.5

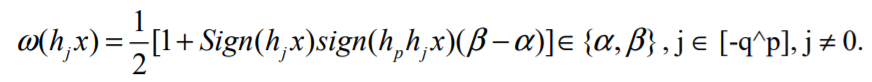
Bước 2: Thực hiện định lượng một số giá trị ngôn ngữ

= 0.5 + 0.5x0.5 = 0.75

= 0.5 – 0.5x0.5 = 0.25

được tính theo công thức:



với

Ở đây:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| j | x |  | sign(hjx) |  |  |
| 1 | cao | 0.75 | +1 | 0.5 x 0.5 = 0.25 | 0.5 |

Do đó, ta có: = 0.75 + (+1)[0.25 – 0.5x0.25) = 0.875

Tương tự, ta tính được:

, ,

Bước 3: Tổng hợp kết quả

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Rất thấp | Thấp | Hơi thấp | Trung bình | Hơi cao | Cao | Rất 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

|  |  |  |  |
| --- | --- | --- | --- |
| R | Bệnh nhân p | | |
|  |  |  |
| Thân nhiệt cao | **0.8** | 0.1 | 0.05 |
| Đau đầu | **0.4** | 0.05 | 0.1 |
| Đau bụng | 0.05 | 0.1 | **0.7** |
| Ho | **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

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Q | Cảm cúm | | | Lao phổi | | | Bệnh.... |
|  |  |  |  |  |  | …… |
| 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 |  |
| Ho | **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.

Bảng 5.3: Quan hệ T = R \* Q

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| T | Cảm cúm | | | Lao phổi | | | Bệnh.... |
| min | min () | max | min | min () | max | …… |
| 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 |  |
| Ho | 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, Cảm cúm) = = (0.8, 0.05, 0.05)

T(p, Lao phổi) = = (0.6, 0.05, 0.1)

Áp dụng công thức \* , ta tính được:

ST(p, Cảm cúm) = 0.8 – 0.05( 1 – 0.8 – 0.05 – 0.05) = 0.795

ST(p, Lao phổi) = 0.5 – 0.1(1 – 0.6 – 0.05 – 0.1) = 0.575

Ở 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

# KẾT LUẬN

## 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.

## 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

|  |  |
| --- | --- |
| [1] | N. T. Duy, "Đại số gia tử và suy luận mờ".*Khoa Điện tử - Hội thảo khoa học 2007 .* |
| [2] | N. M. H. N. C. H. Lê Văn Tường Lân, "MỘT PHƯƠNG PHÁP ĐỊNH LƯỢNG GIÁ TRỊ NGÔN NGỮ TRONG TẬP MẪU HUẤN LUYỆN ĐỂ XÂY DỰNG CÂY QUYẾT ĐỊNH MỜ". |
| [3] | "wikipedia.org," [Online]. Available: https://vi.wikipedia.org/wiki/Logic\_m%E1%BB%9D. |
| [4] | Pham Hong Phong, Hieu Trong Dinh, "Research Gate," [Online]. Available: • https://www.researchgate.net/profile/Hieu\_Dinh/publication/266084039\_Some\_compositions\_of\_picture\_fuzzy\_relations/links/542520b40cf238c6ea73cd7a/Some-compositions-of-picture-fuzzy-relations.pdf. |
| [5] | "Fuzzy set," [Online]. Available: https://en.wikipedia.org/wiki/Fuzzy\_set. |
| [6] | "wikipedia.org," [Online]. Available: https://vi.wikipedia.org/wiki/T%E1%BA%ADp\_m%E1%BB%9D. |