



TRI THỨC VÀ MỘT SỐ PHƯƠNG PHÁP BIỂU DIỄN TRI THỨC

- Tổng quan
- Một số phương pháp biểu diễn tri thức
- Xử lý tri thức không chắc chắn

- TS. Phạm Hạ Thủy
- Khoa CNTT- Đại học Nông Nghiệp HN



Tổng quan

- Khái niệm về tri thức
- Phân loại tri thức
- Biểu diễn tri thức

Khái niệm về tri thức

- Tri thức được định nghĩa trong Oxford Dictionaries:
 - facts, information, and skills acquired by a person through experience or education; the theoretical or practical understanding of a subject.
 - awareness or familiarity gained by experience of a fact or situation.
- Định nghĩa trong Bách khoa toàn thư Việt Nam
 - Tri thức là kết quả của các quá trình nhận thức của con người về đối tượng được nhận thức, làm tái hiện trong tư tưởng con người những thuộc tính, những mối quan hệ, những quy luật vận động, phát triển của đối tượng và được diễn đạt bằng ngôn ngữ tự nhiên hay hệ thống kí hiệu khác.

Khái niệm về biểu diễn tri thức

- Biểu diễn tri thức(knowledge representation): các phương pháp diễn tả và tổ chức tri thức trong máy tính cho các hệ thống tin có tính chất trí tuệ để máy có thể tiến hành các phép lập luận tự động.
- Trong trí tuệ nhân tạo, các phương pháp BDTT thường được sử dụng như logic vị từ, mạng ngữ nghĩa, biểu diễn khung, luật dẫn.



Phân loại tri thức (knowledge classification)

- procedural knowledge-Descriptive knowledge)
- Explicit knowledge -Implicit knowledge
- Certain knowledge – Uncertain knowledge

Tri thức thủ tục – Tri thức mô tả

- **Tri thức thủ tục (procedural knowledge)** là tri thức mô tả cách giải quyết một vấn đề, quy trình xử lý các công việc, lịch trình tiến hành các thao tác ... Các dạng của tri thức thủ tục thường dùng là các luật, chiến lược, lịch trình...
- *VD: - Các bước giải một phương trình bậc 2; cách làm bánh chưng,...*
- **Tri thức mô tả (Descriptive knowledge)** là một khẳng định về một sự kiện, hiện tượng hay một khái niệm nào đó trong một hoàn cảnh không gian hoặc thời gian nhất định.
- *Ví dụ: khẳng định về hiện tượng: "Mặt trời lặn ở phương Tây". Khái niệm về: "tam giác đều: là tam giác có ba góc bằng nhau"; Hà nội là thủ đô của Việt Nam; Tùng là kỹ sư của nhà máy; Nhiệt độ ngày 20/8/2013 là 33°....*

Tri thức ẩn- tri thức hiện

- Tri thức hiện(Explicit knowledge)
 - là những tri thức được giải thích và mã hóa dưới dạng văn bản, tài liệu, âm thanh, phim, ảnh,... thông qua ngôn ngữ Đây là những tri thức đã được thể hiện ra ngoài và dễ dàng chuyển giao, thường được tiếp nhận qua hệ thống giáo dục và đào tạo chính quy.
- Tri thức ẩn(implicit knowledge)
 - là những tri thức thu được từ sự trải nghiệm thực tế, dạng tri thức này thường ẩn trong mỗi cá nhân và rất khó “mã hóa” và chuyển giao, thường bao gồm: niềm tin, giá trị, kinh nghiệm, bí quyết, kỹ năng...

Tri thức chắc chắn- không chắc chắn

- **Tri thức chắc chắn** (certain knowledge): là những tri thức chắc chắn đúng.
- VD:-khẳng định:tổng các góc trong một tam giác bằng 180° .
 - -Nếu hỏng hệ thống điện thì xe máy không khởi động được.
- Tri thức không chắc chắn(uncertain knowledge): là những khẳng định, luật suy diễn không chắc chắn đúng.
 - VD:-Khẳng định:Nếu bệnh nhân bị sốt cao và ho thì bệnh nhân bị viêm phổi.
 - -Nếu xe máy không khởi động được thì xe máy bị hỏng bộ điện.

Các hình thức chia sẻ tri thức

- Tri thức có thể được chuyển giao và chia sẻ. Có thể chia các hình thức chia sẻ tri thức thành bốn dạng chính:
- Ẩn - Ẩn: Khi người chia sẻ và người tiếp nhận giao tiếp trực tiếp với nhau (ví dụ: học nghề, giao tiếp, giảng bài...) thì việc tiếp nhận này là từ tri thức ẩn thành tri thức ẩn. Tri thức từ người này không qua trung gian mà chuyển ngay thành tri thức của người kia.
- Ẩn - Hiện: Một người mã hóa tri thức của mình ra thành văn bản hay các hình thức hiện hữu khác thì đó lại là quá trình tri thức từ ẩn (trong đầu người đó) trở thành hiện (văn bản, tài liệu, v.v.).
- Hiện - Hiện: Tập hợp các tri thức hiện đã có để tạo ra tri thức hiện khác. Quá trình này được thể hiện qua việc sao lưu, chuyển giao hay tổng hợp dữ liệu.
- Hiện - Ẩn: Tri thức từ dạng hiện trở thành dạng ẩn. Điển hình quá trình này là việc đọc sách. Học sinh đọc sách (tri thức hiện) và rút ra được các bài học, tri thức cho mình (ẩn).

Một số phương pháp biểu diễn tri thức

- Bảng văn bản
- Bảng logic vị từ
- Bảng mạng ngữ nghĩa
- Bảng frame (khung)

Các phương pháp biểu diễn tri thức

- **Biểu diễn bằng văn bản, băng ghi âm, băng ghi hình**

-Mô tả tri thức bằng cách viết ra trong một văn bản, hoặc băng lời ghi vào băng có thể kết hợp hình ảnh minh họa

VD: sách giáo khoa, băng ghi cách dạy nấu ăn...

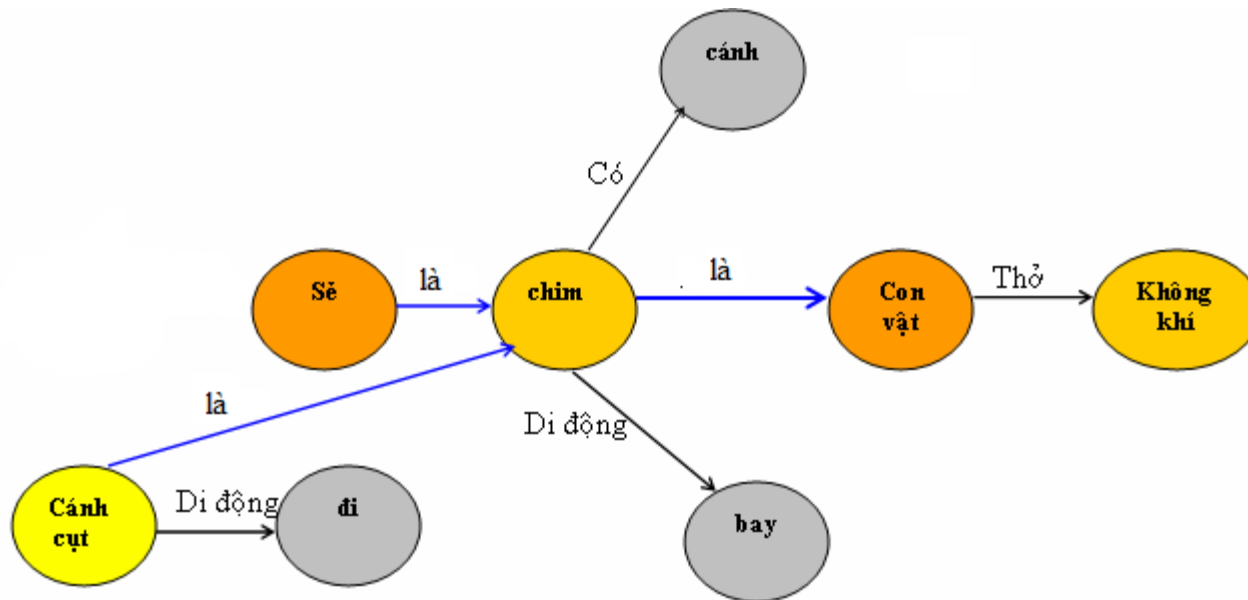
Để xử lý được tri thức loại này cần có các công cụ phân tích câu, hoặc đoạn văn bản.

Biểu diễn tri thức bằng logic mệnh đề, logic vị từ

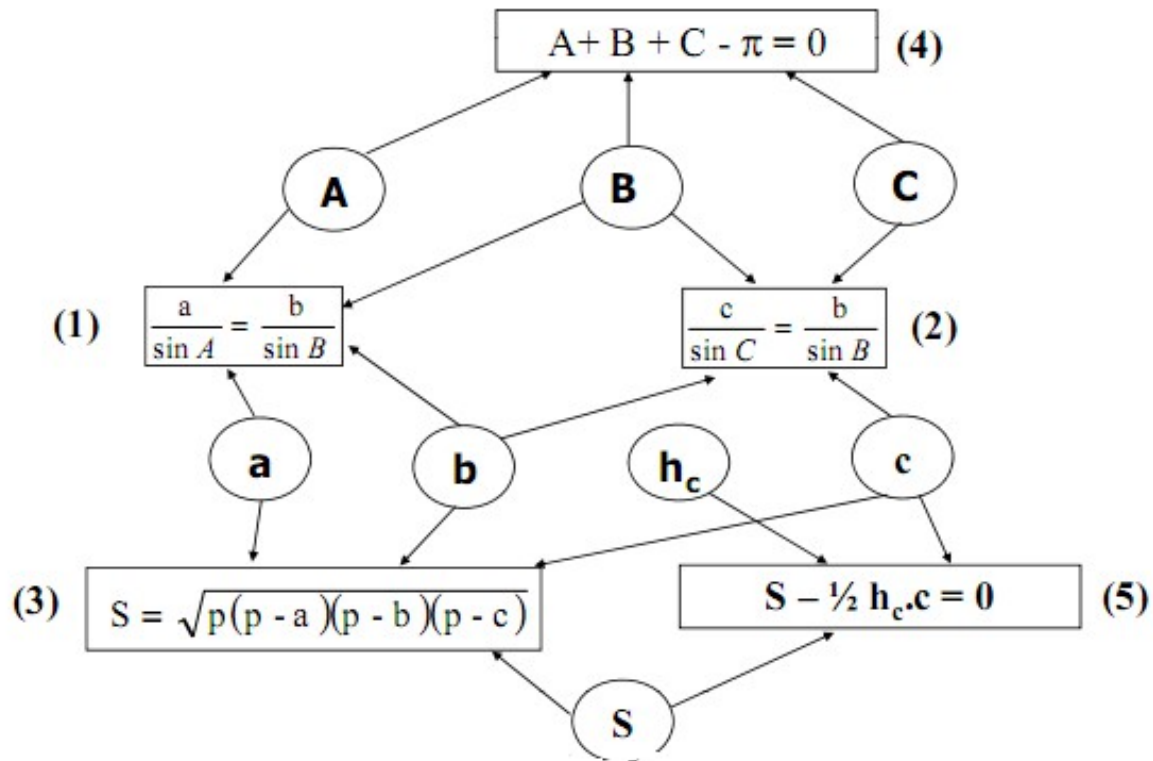
- Thông qua các qui tắc biểu diễn và xử lý mệnh đề, vị từ, các luật sản xuất, luật suy diễn
 - VD: Định nghĩa các vị từ $LH(x,y)$: x lớn hơn y ,
 $CTG(a,b,c)$: a, b, c là cạnh của một tam giác...
 - Luật sản xuất: $If A \text{ then } B$ ($A \rightarrow B$)
 - Các câu Horn chuẩn là các luật có dạng:
Head:- Body tương đương với Body \rightarrow Head
với Head là một vị từ còn Body là một tích của các vị từ. Vd: $IsDiv(X,6) : IsDiv(X,3), IsDiv(X,2)$.
(có nghĩa Nếu X chia hết cho 3 và chia hết cho 2 thì X chia hết cho 6)
 - Các luật suy diễn: luật modus ponens
 - Nếu $A \rightarrow B$ và A đúng thì suy ra B đúng

Biểu diễn bằng mạng ngữ nghĩa

- Phương pháp sử dụng các nút để chỉ các đối tượng và các mũi tên để chỉ mối quan hệ giữa các nút. Ví dụ:



VD: Mạng ngữ nghĩa dùng để giải bài toán: cho 2 góc A, B và cạnh a, tìm đường cao h_c .



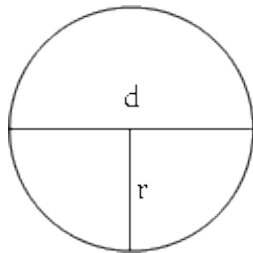
Biểu diễn bằng khung (Frame)

- Mỗi một đối tượng tri thức được biểu diễn thông qua một khung gồm các slot (mô tả các thuộc tính của đối tượng) và các facet mô tả các giá trị, hoặc các thủ tục tính các giá trị.

Frame : **CIRCLE**

(hình tròn)

- r : radius;
- s : area;
- p : perimeter;
- d : diameter;
- $d = 2 \times r$;
- $s = \pi \times r^2$;
- $p = 2 \times \pi \times r$;



Frame **RECTANGLE**

(hình chữ nhật)

$b1$: side;

$b2$: side;

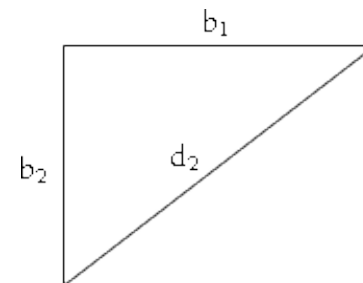
s : area;

p : perimeter;

$s = b1 \times b2$;

$p = 2 \times (b1 + b2)$;

$d^2 = b1^2 + b2^2$;



Xử lý tri thức không chắc chắn

- Tri thức không chắc chắn(uncertain knowledge): là những khẳng định, luật suy diễn không chắc chắn đúng
- Ví dụ:- nhiệt độ ngày mai là 32°
 - Nếu xe máy không khởi động được thì bugi bị hỏng
- Tuy không chắc chắn nhưng những tri thức không chắc chắn quan trọng và cần thiết trong đời sống
- Cần thiết phải có phương pháp biểu diễn và đánh giá

Một số phương pháp biểu diễn tri thức không chắc chắn

- Biểu diễn bằng xác suất
- Biểu diễn bằng hệ số chắc chắn

Biểu diễn tri thức không chắc chắn bằng xác suất

- Xác suất của một sự kiện

$$P(A) = \frac{\text{Số trường hợp thuận lợi cho } A}{\text{Tổng số trường hợp có thể xảy ra}}$$

- Ví dụ: gọi A là sự kiện xuất hiện mặt chẵn khi gieo một con xúc sắc 6 mặt, Khi đó $P(A)=1/2$
- Trong thống kê: Xác suất của một sự kiện A được đo bằng tỉ số số lần xảy ra và tổng số lần quan sát
- VD: khi thống kê 1000 người dân thì thấy có 1 người bị mắc bệnh lao phổi. Nếu gọi A là sự kiện một người bị mắc bệnh lao phổi. Khi đó $P(A) = 1/1000 = 0.001=0.1\%$

Xác suất có điều kiện

- $P(A/B)$: XS để sự kiện A xảy ra với điều kiện sự kiện B đã xảy ra.
- (Đôi khi người ta gọi A là giả thuyết, B là bằng chứng. $P(A/B)$ đo mức độ đúng của giả thuyết A khi có bằng chứng B)
- VD: Xét luật: A B
B- bệnh nhân bị sốt
A – bệnh nhân bị cúm
Khi đó ta có $P(A/B) = 0.45$
(Trong 100 trường hợp BN bị sốt thì có 45% là bị cúm)
- Xét luật ngược lại : B A
Ta có $P(B/A) = 0.85$
(trong 100 trường hợp BN bị cúm thì có 85% bị sốt)

Công thức Bayes

$$P(A/B) = \frac{P(B/A) * P(A)}{P(B)}$$

- Trường hợp A,B độc lập thì $P(A/B)=P(A)$
- Công thức tổng quát:

$$P(A_i/B) = \frac{P(B/A_i) * P(A_i)}{\sum_{k=1}^n P(B/A_k) * P(A_k)}$$

-
- Ví dụ: Một bệnh nhân có triệu chứng sốt và ho (B). Các giả thuyết là A1: bệnh nhân bị viêm phổi; A2: bệnh nhân bị cúm. Xác định xem khả năng nào lớn hơn.
 - Giả sử ta có: $P(B/A1) = 0.8$; $P(B/A2) = 0.2$. Đồng thời ta có $P(A1) = 0.005$ và $P(A2)=0.001$.
 - Khi đó: $P(A1/B) = 0.8*0.005/(0.8*0.005+0.2*0.001) = 0.95238$
 - $P(A2/B) = 0.2*0.001/(0.8*0.005+0.2*0.001) = 0.04762$
 - Như vậy có thể khẳng định rằng 95% là bệnh nhân viêm phổi.

Biểu diễn tri thức không chắc chắn với hệ số chắc chắn

- Hệ số chắc chắn (Certainty Factors- CF)
- $CF(\text{Fact}) \in [-1, 1]$
- Sự kiện (Fact) càng gần -1 thì độ tin tưởng sự kiện đó là đúng càng nhỏ
- Sự kiện (Fact) càng gần 1 thì độ tin tưởng sự kiện đó là đúng càng lớn.
- Xét luật $R: A \rightarrow B$
- Gọi A là điều kiện, B là kết luận của luật
- Ta có hệ số chắc chắn của ĐK, của luật và của kết luận.

Quy tắc tính HSCC của điều kiện

○ **Tính HSCC của điều kiện:**

- $CF(\text{ĐK1 And ĐK2}) = \text{Min}[CF(\text{ĐK1}), CF(\text{ĐK2})]$
- $CF(\text{ĐK1 Or ĐK2}) = \text{Max}[CF(\text{ĐK1}), CF(\text{ĐK2})]$
- Ví dụ : $CF(\text{bệnh nhân bị sốt}) = 0.9$

$$CF(\text{bệnh nhân bị hắt hơi}) = 0.6$$

$$\Rightarrow CF(\text{bệnh nhân bị sốt And bệnh nhân bị hắt hơi}) = 0.6$$

$$\Rightarrow CF(\text{bệnh nhân bị sốt Or bệnh nhân bị hắt hơi}) = 0.9$$

Quy tắc tính HSCC của kết luận

- Xét luật R: ĐK KL

Khi đó: $CF(KL) = CF(R) * CF(ĐK)$

Khi có nhiều luật suy ra cùng kết luận:

R1: ĐK1 KL với $CF_1(KL)$

R2: ĐK2 KL với $CF_2(KL)$

$$CF_{\text{COMBINE}}(CF_1, CF_2, \dots) = \begin{cases} CF_1 + CF_2 (1 - CF_1) & \text{both} > 0 \\ \frac{CF_1 + CF_2}{1 - \min(|CF_1|, |CF_2|)} & \text{one} < 0 \\ CF_1 + CF_2 (1 + CF_1) & \text{both} < 0 \end{cases}$$

Tính HSCC trường hợp một kết luận suy từ nhiều luật

Ví dụ :

- $CF(\text{bệnh nhân bị hắt hơi}) = 0.8$
- $CF(\text{bệnh nhân bị sốt}) = 1$
- $CF(\text{If bệnh nhân bị hắt hơi Then bệnh nhân bị cúm}) = 0.5$
- $CF(\text{If bệnh nhân bị sốt Then bệnh nhân bị cúm}) = 0.6$
- $\Rightarrow CF_1(\text{bệnh nhân bị cúm}) = 0.4$
- $\Rightarrow CF_2(\text{bệnh nhân bị cúm}) = 0.6$
- $\Rightarrow CF(\text{bệnh nhân bị cúm}) = 0.4 + 0.6(1 - 0.4) = 0.76$

Tài liệu tham khảo

- Wiig, K. (1993). *Knowledge management foundations*. Arlington, TX: Schema Press.
- Malhotra, Y. (2000). *Knowledge management and virtual organizations*. London: Idea Group Publishing.
- Wiig, K. M. (1997). “*Roles of knowledge-based systems in support of knowledge management*”. In J. Liebowitz & L. C. Wilcox, *Knowledge management and its integrative elements* (pp. 69-87). New York: CRC Press.
- Nguyễn Thanh Thủy(1999). *Trí tuệ nhân tạo: Các phương pháp giải quyết vấn đề và xử lý tri thức*. Nhà xuất bản Giáo dục.
- Nguyễn Thiện Thành (2006). Bài giảng trí tuệ nhân tạo và hệ chuyên gia, Đại học Bách khoa Thành phố HCM
- <https://db69bfee-a-3a390da7-s-sites.googlegroups.com/a /ptvuong.com/cac-he-co-so-tri-thuc>
- http://vi.wikipedia.org/wiki/Tri_th%E1%BB%A9c
- <http://www.knowledge-management-tools.net/>
- <http://oxforddictionaries.com/definition/english/>
- <http://bachkhoatoanthu.vass.gov.vn/>