

# Công nghệ tri thức và ứng dụng - CS2101

Tuesday, September 8, 2020 6:09 PM

## Knowledge engineering and applications - CS2101

### Buổi 1 (08/09/2020)

#### Nội dung

1. Nhập môn KE và ứng dụng
2. KBS và ứng dụng (knowledge-based systems)
3. Một số phương pháp và kỹ thuật cơ bản
4. Case study (mẫu ứng dụng cụ thể)

dvnhon@gmail.com  
0908107799

1. Nhập môn công nghệ tri thức
  - Data
  - Information
  - Knowledge:

#### Khái niệm AI và tri thức KE?

Giới thiệu một số ứng dụng của KE trong lĩnh vực thực tế:

1. Hệ giải bài tập HHP cho học sinh THPT:  
Input: Đề bài tập hình học phẳng  
Output: bài giải

Rõ ràng là hệ thống phải có cơ sở tri thức HHP và module thực hiện việc suy luận tìm lời giải cho bài toán được nhập vào.

Bài tập: Mỗi bạn cho một ví dụ (ứng dụng thực tế) về hệ cơ sở tri thức.

Domain hẹp hơn Lĩnh vực

### Buổi 2: 15/09/2020

- Khái niệm (Concepts)
- Quan hệ (Relations)
- Luật ()

Công nghệ tri thức và ứng dụng là gì?

5 Công đoạn của Knowledge Engineering:  
+ Thu thập tri thức (Knowledge ecquisition)  
+ Validation tri thức (Knowledge validation)  
+ Biểu diễn tri thức (Knowledge representation)  
+ Suy luận (Inferencing)  
+ (Explanation and justification)

Một hệ thống để gọi là hệ cơ sở tri thức cần có

- Lưu trữ tri thức của một lĩnh vực, phạm vi, domain, ..
- Có khả năng giải được vấn đề dựa trên tri thức

Các hệ thống ES, IPS, DSS, ....

#### Bài tập tuần:

Làm và nộp trên dropox (hạn 1 tuần 21/09/2020)

- Mỗi bạn cho một ví dụ (ứng dụng thực tế) v
- Miền tri thức cụ thể (Ứng dụng đang đề cập)
- User và yêu cầu ứng dụng (Đối tượng sử dụng)
- Phân tích để thấy rằng ứng dụng này là cần

VD:

Đề tài: ....

Tổng quan

- Giới thiệu
- Động cơ nghiên cứu
- Mục tiêu luận văn
- Phạm vi nghiên cứu
- Ý nghĩa khoa học và thực tiễn

### Buổi 3: 21/09/2020

#### Bài tập 2:

General gồm 3 phần chính:

- + User interface (giao diện hệ thống)
  - Giao diện cho người dùng (user) -> (user interface) -> Người sử dụng cơ sở tri thức
  - Giao diện cho người có chuyên môn (Expert interface) -> (Expert interface) -> Người quản lý cơ sở tri thức
- + inference engine (module -> thực hiện quá trình suy luận hay suy diễn)
- + knowledgebase (lưu trữ (miền tri thức của ứng dụng), trái tim của hệ thống)

## Các thành phần khác

- + Working Memory
- + Explanation componet
- + KMS
- + Knowledge Editor
- + Knowledge Acquisition

- Ví dụ:  
Input: đề bài tập HHP (Hình học phẳng)  
Output: Bài giải
- User: học sinh
- Knowledge engineer: GV (dạy về HHP)
- KB: cơ sở tri thức chứa (lưu) kiến thức về HHP trong chương trình toán THPT
- IE:

## 5. Quy trình xây dựng hệ thống

Các bước (giai đoạn)  
Xây dựng KBS

## 5.1 Xác định miền tri thức

- + Xác định vấn đề, bài toán, yêu cầu cần giải quyết
- + Thu thập tri thức.
- + Thu thập các đề bài toán cụ thể, cùng với các bài toán đã biết.

## 5.2 Thiết kế cơ sở tri thức (knowledge base)

- + Đưa ra cách tổ chức lưu trữ tri thức trên máy tính một cách phù hợp cho việc cập nhật, xử lý và giải quyết vấn đề của hệ thống, đặc biệt là phục vụ cho hoạt động suy diễn của inference Engine

## ==&gt; Vấn đề kỹ thuật chủ yếu:

- **Biểu diễn tri thức (Knowledge Representation Methods)**
- 

## 5.3 Thiết kế bộ suy diễn (Inference Engine)

Thiết kế việc suy diễn để giải vấn đề dựa trên KB, thuật giải thể hiện phương pháp suy diễn được gọi là "Thuật giải suy diễn" (reasoning Algrithm)

## 5.4 Thiết kế user interface

(chưa xét interface cho knowledge engineer

## 5.5 Coding and Testing

## Vấn đề kỹ thuật

- Về thu thập tri thức:
  - + Nguồn thu thập tri thức?
    1. Tài liệu, sách, văn bản, ...
    2. Người chuyên gia. Note: lựa chọn chuyên gia dựa vào trình độ hay đẳng cấp CHUYÊN GIA, thời gian có thể tham gia, sự nhiệt tình tích cực, hiểu biết về IT (sơ cấp) và tin vào sự thành công của dự án, ...
  - + Ai thu thập? (KE và CTV)
  - + Thu thập như thế nào (Thủ công / tự động / bán tự động), bằng cách nào?
- Về thu thập đề bài và bài giải tương ứng:
  - + ...
- Viết ra / ghi chép cho kết quả thu thập
  - + Phân loại
  - + Hệ thống hóa
  - + Chọn lọc, sà lọc
- 

Cập nhật slide KBS architecture

Không có mũi tên từ Inference engine -> Knowledge base của hệ thống chỉ có chuyên gia mới có quyền

Buổi 4: 29/09/2020

## Case study 1:

Xây dựng ứng dụng (phần mềm máy tính) giải bài toán tính toán trên một tam giác.

Phạm vi: ta chỉ xét 12 yếu tố của tam giác là a, b, c, A, B, C, ha, hb, hc, S, p, R

Dạng bài toán: Trên tam giác ABC, cho trước giá trị của một số yếu tố (giả thiết), yêu cầu tính giá trị của một vài yếu tố mục tiêu.

## Mẫu 1:

GT: a = 5, b = 4, A =  $\pi/2$   
Goal: S, R

## Mẫu 2:

GT: a = 5, b = 5, R = 2.5

## Case study 2:

Tìm các phản ứng hóa học để điều chế các chất cần có từ các chất đã có

## Case study 3:

Dựa trên kiến thức về điện một chiều, cho một mạch điện đơn giản trên đó gồm các điện trở mắc (kết hợp) theo 2 cách "song song" và "nối tiếp", cùng với một số yếu tố được cho trước giá trị yêu cầu tính một vài yếu tố được quan tâm (mục tiêu) trên mạch điện

Tri thức và các kỹ thuật cơ bản

Vấn đề BDDT (Biểu diễn tri thức)

- Có: (1) tri thức của một miền tri thức nhất định đang được diễn đạt ở dạng tự nhiên  
(2) mục tiêu hay nhu cầu xử lý giải quyết vấn đề

Kết quả mong muốn: mô hình biểu diễn cho tri thức trên (1) thỏa: (i) tổ chức lưu trữ xử lý được trên máy, (ii) đáp ứng cho việc suy luận hay suy diễn giải quyết vấn đề (2).

## 5 lớp công cụ

Hiện nay, mô hình biểu diễn tri thức thường dựa trên các công cụ (từ trừu tượng đến ít trừu tượng hơn)

- Các cấu trúc dữ liệu cơ bản và trừu tượng đã biết
  - Dãy, danh sách, tập hợp
  - Cây
  - Ngăn xếp, hàng đợi
  - Đồ thị
- Các mô hình và cấu trúc toán học:
  - Tập hợp, quan hệ, hàm hay ánh xạ.
  - Toán rời rạc và đồ thị
  - Logic toán học
    - Logic mệnh đề
    - Logic vị từ
    - Logic mờ
    - Logic thời gian
    - V.v...
  - Xác suất thống kê
  - Đại số tuyến tính
  - Giải tích
- Các mô hình biểu diễn tri thức cơ bản -> Nâng cao
  - + Logic vị từ
  - + Mạng ngữ nghĩa
  - + Hệ luật dẫn
  - + frames, classes.
  - + những mô hình hay phương pháp mới hơn như:
    - Mạng tính toán, mạng các đối tượng tính toán
    - Đồ thị khái niệm, đồ thị keyphrase
- Các ngôn ngữ đặc tả:
- Các **ontology**

Buổi 5 (06/10/2020)

- Lý thuyết: Các phương pháp BDDT cơ bản và việc vận dụng
- Ví dụ - bài tập (cục bộ)
- Case study - Bài tập chương 1

Bài tập 1:

Case study điện một chiều

Khái niệm về công nghệ tri thức:

- Có thể được xem là một nhánh nghiên cứu của trí tuệ nhân tạo, phân tích tri thức lĩnh vực và chuyển nó thành những mô hình tính toán đưa vào máy tính để phục vụ những nhu cầu cần thiết.
- Hệ cơ sở tri thức gồm 2 thành phần cơ bản chính: Cơ sở tri thức và bộ suy diễn dựa trên cơ sở tri thức đó.

### Các phương pháp biểu diễn tri thức cơ bản

- Logic vị từ
    - + Mô hình: (Predicates, Clauses)
- Predicates là tập hợp các vị từ, mỗi vị từ biểu diễn cho phát biểu nói về một tính chất của đối tượng hay một quan hệ giữa các đối tượng, mỗi vị từ xác định bởi tên vị từ và các kiểu tham biến

Ví dụ: giới(x:sinhvien)

Ví dụ: vg(f: vector, P: plane).

Clauses là tập hợp các biểu thức vị từ gồm 2 dạng:

Fact: <vị từ>(các đối tượng)

Rule: (được viết dưới 2 dạng)

- fact :- fact2, fact2, ..., factk.
- "if fact1, fact2, ..., factk then factct"

Predicates

Ví dụ:

Cho

$$\forall k \in \mathbb{N}, \exists n < N, p(n) \bigwedge (n > k)$$

$$\forall a, b, c < N : us(a, b) \wedge us(b, c)$$

Viết theo fact

P(5).

$$\forall a, \forall b, \forall c, us(a, b) \wedge us(b, c) \rightarrow us(a, c)$$

Viết thành:

$$us(A, B) \wedge us(B, C) \rightarrow us(A, C).$$

Final là:  $us(A, C) :- us(A, B), us(B, C).$

```

nguyen_to(integer)
us(integer, integer)
...

```

#### Clauses

```

Nguyễn tố (5)
Us(5, 10)
Us(A,C) :- us(A,B), us(B,C)

```

#### Hệ luật dẫn

- Khái niệm luật diễn:  
+ Luật có dạng "I ... then ... ?  
Ví dụ:
  - Định lý pitago
  - Định lý ...

#### Mô hình hệ luật dẫn: (Facts, Rules)

Facts là tập các sự kiện hay tác vụ trong phạm vi tri thức.  
Rules là tập các luật dẫn, mỗi r thuộc Rules có dạng  
R:  $gt(r) \Rightarrow kl(r)$

#### Sự kiện (fact) ?

NX: sự kiện trong thực tế thường là có cấu trúc như: (1) phát biểu về tính chất của 1 phần tử, (2) phát biểu về quan hệ (2 ngôi hay nhiều ngôi) giữa các phần tử.

Ví dụ: một phần kiến thức về một tam giác trong hình học

(1) Các yếu tố của tam giác, mỗi sự kiện là một phát biểu nói lên tính xác định của yếu tố -> ký hiệu cho các sự kiện:

a, b, c, A, B, C, S, p, R, ha, hb, hc, ...

(2) Các luật nói lên liên hệ "dẫn xuất" giữa các ....

Facts = {a, b, c, A, B, C, S, p, R, ha, hb, hc, ...}

```

Rules = {
    r1: {A, B} -> {C}, và C = pi - A - B;
    r2: {A, C} -> {B}, và B = pi - A - C;
    v.v..
}

```

Tổ chức lưu trữ: được xác lập cụ thể dựa trên dạng facts. Thường là ta sử dụng các cấu trúc dữ liệu đã biết như struct, frames, classes, ..

Lưu trên đĩa: sử dụng 2 tập tin dạng "text có cấu trúc": Facts.txt và Rules.txt

Cấu trúc file Fact.txt

#### Begin

```

A: cạnh a của tam giác
B: ...
V.v ...

```

#### End

#### Cấu trúc file Rules.txt

#### Begin

```

{A, B} => {C}, C = pi - A - B: luật về góc tam giác.
V.v. ...

```

#### End

#### Vấn đề suy diễn (suy luận) trên hệ luật dẫn:

Cho trước hệ luật dẫn K = (Facts, Rules). Giả sử có một tập sự kiện GT đã xác định, ta xét một tập sự kiện mục tiêu

KL. Hỏi có thể suy ra KL từ GT dựa trên thi thức K hay không?

Ký hiệu bài toán: GT->KL

#### Bài tập Case Study

#### Thuật giải suy diễn tiến:

Các biến trọng thuật giải:

- Tập sự kiện đã biết: Known
- Danh sách các luật được áp dụng: Solution

```

B1: Solution = []
Known = GT

```

B2: While (KL chưa nằm trong Known) do

2.1: tìm luật r để áp dụng trên Known nhằm sinh ra sự kiện mới:

$Gt^* \subseteq Known$ , và  $kl(r)$  nằm trong Known.

2.2 if (không có r) then

Dừng: không tìm được lời giải

2.3 Thêm r vào Solution; trên  $kl(r)$  vào Known;

End While

B3: Tìm được lời giải sử dụng danh sách luật Solution

#### Suy diễn ngược

....

Class trong BDDT nó có nghĩa rộng hơn trong OOP

<< class >> sẽ được hiểu theo nghĩa trừu tượng hơn:

+ cấu trúc thành phần: các thuộc tính, các quan hệ trên các thuộc tính, các tính chất, các luật nội tại;

+ các phương thức, các hành vi trừu tượng khác như suy diễn giải quyết vấn đề, giao tiếp, ...

+ có cả khái niệm về các thuộc tính đặc trưng, slot.

- Các thuộc tính đặc trưng thường có giá trị default.

Buổi 6: 13/10/2020

Nhắc lại:

Bài tập:

Deadline đ/k: 20/10/2020

Deadline nộp B/T: 15/11/2020

Mỗi nhóm 2 - 3 người, mỗi nhóm gửi mail tới dvnhon@gmail.com để đăng ký và cho thông tin

+ danh sách học viên

+ mô tả đề bài: (1) tựa đề, (2) mục tiêu

+ bài nộp (upload lên thư mục của nhóm lên trên dropox của Thầy share)

File báo cáo, kèm theo những cái cần thiết

1

KBS:

+ IPS

+ ES

+ ITS

+ ...

Buổi 7: 20/10/202

Hệ thống các khái niệm (concept) (hay các lớp đối tượng) gồm các khái niệm cùng với các liên cơ bản như:

+ IS\_A

+ trình tự xây dựng (hình thành) khái niệm các khái niệm tiền đề, các khái niệm được định nghĩa dựa trên các khái niệm có trước

Sự kiện (fact, event) và luật(rule)

Ví dụ 1: Kiến thức hình học giải tích 2 chiều, 3 chiều

- Khái niệm:
- Hệ trục toạ

Buổi 8 (27/10/2020)

Ví dụ 1: Tri thức về hình học, diện 1 chiều, hóa vô cơ

- Phạm vi tri thức
- Chức năng ứng dụng cho đối tượng cụ thể

Ontology (Bản thể học)

Ví dụ 2: Kiến thức về nhập môn tin học, nhận môn lập trình

- Phạm vi của tri thức
- Chức năng ứng dụng cho đối tượng cụ thể:

1. Tìm kiếm theo khái niệm, quan hệ, luật.

Buổi 9 (03/11/2020)

What is an Ontology

An Ontology is a

Formal specification

WordNet

Domain and application

Cấu trúc cơ bản của một mô hình ontology

Basic Ontology model = (C, R, Rules)

Ontology model = (C, R, Rules) + (Cac thanh phan tri thuc khac)

Trong do, cac thanh phan phai duoc dinh nghia hay dac ta mot cach ro rang, cu the.

Ví dụ; Trong kiến thức về CSDL (lĩnh vực CNTT), khái niệm "table" có thể được biểu diễn /

Trong đó

```
C = {diem, duong, doan}
diem = class {
    X, y: real;
};

duong = class {
    A, B : diem;
}

doan = class {
    A, B : diem;

    len: real;
    rules { len ^ 2 = (A.x - B.x)^2 + (A.y - B.y)^2};
    Constrain = { len >= 0 }
}
```

ngày 24/11/2020

Nhập môn KE và ứng dụng  
KBS  
Một số phương pháp kỹ thuật cơ bản  
Case study

lịch  
24/11/2020

01/11/2020

08/12/2020  
Ôn tập và giải đáp đề thi cuối kỳ

Đồ án  
chú ý

1. Yêu cầu chung:
  - Nội dung: tìm hiểu công nghệ và làm một ứng dụng nhỏ có CSTT và giải quyết yêu cầu/ vấn đề dựa trên Cơ sở tri thức.
  - Hình thức: theo chuẩn chung của luận văn
2. Tiêu chuẩn về kỹ thuật và ứng dụng: phù hợp với chuyên đề cao học.

## 1. Nhập môn KE và ứng dụng

## 2. KBS (Knowledge - Base Systems) và ứng dụng

### 2.1. Cơ sở lý thuyết về KBS

- Khái niệm về KBS
- Kiến trúc hệ thống
- Quy trình xây dựng hệ thống

### 2.2 Một ứng dụng (cụ thể)

## 3. Một số phương pháp và kỹ thuật cơ bản

### 3.1 Các phương pháp biểu diễn tri thức (cơ bản và nâng cao)

- Vấn đề Biểu diễn tri thức
- Các phương pháp cơ bản
- Ontology
  - Khái niệm
  - Mô hình
  - Ngôn ngữ đặc tả
  - Vấn đề và giải pháp

### 3.2 Phương pháp suy diễn

- \* Suy diễn tiến
- \* Suy diễn lùi
- \* Suy diễn tiến/l

## 4. Case study

### 1 hệ thống

các giai đoạn xây dựng hệ cơ sở tri thức

nội dung của các giai đoạn xây dựng hệ cơ sở tri thức

### 2 kỹ thuật

trình bày phương pháp hệ luật dẫn và phương pháp hệ suy diễn tiến

ứng dụng trong việc thiết kế ứng dụng sau đây

Các bạn hãy phân tích đánh giá phương pháp hệ luật dẫn có ưu nhược điểm gì?

Nêu lên khái niệm và một số thành phần chính trong ontology cho biểu diễn tri thức