

## I- Giới thiệu bài toán

- Hiện nay, khoa học công nghệ phát triển rất nhanh chóng làm thay đổi rất nhiều trong đời sống hàng ngày của chúng ta. Ngày nay, ta thấy rằng chiếc điện thoại di động dường như là vật bất ly thân đối với mỗi người ở mọi tầng lớp. Bởi nó đáp ứng rất nhiều nhu cầu của con người trong cuộc sống hiện đại ngày nay từ nhu cầu thiết yếu trao đổi thông tin tới nhu cầu giải trí, tới phục vụ cho các mục đích chuyên môn khác.
- Những chiếc điện thoại ngày nay không những chỉ giúp con người trao đổi thông tin mà nó còn cung cấp rất nhiều các tiện ích khác: nghe nhạc, chụp hình, truy cập internet.....
- Nhưng hiện tại có rất nhiều hãng và sản phẩm điện thoại di động ra đời với rất nhiều các kiểu dáng màu sắc tính năng, giá cả..... khiến cho việc lựa chọn một chiếc điện thoại phù hợp với mục đích sử dụng, công việc, khả năng tài chính, ... trở nên khó khăn hơn.
- Hệ tư vấn mua điện thoại di động sẽ hỗ trợ cho người sử dụng chọn được chiếc điện thoại phù hợp với cá nhân người sử dụng.

## II- Cơ sở lý thuyết

### 1. Biểu diễn tri thức

- Biểu diễn tri thức (knowledge representation) là một lĩnh vực nghiên cứu quan trọng của Trí tuệ nhân tạo. Nhằm phát triển các phương pháp, cách thức biểu diễn tri thức và các công cụ hỗ trợ việc biểu diễn tri thức.
- Có rất nhiều phương pháp biểu diễn tri thức, và phải phụ thuộc vào tri thức cụ thể của từng bài toán ta mới chọn được một phương pháp biểu diễn hợp lý nhất.
  - Biểu diễn tri thức bằng luật sản xuất (Production rules)
  - Biểu diễn tri thức bằng khung (Frames)
  - Biểu diễn tri thức sử dụng mạng ngữ nghĩa (Semantic networks)
  - Biểu diễn tri thức bằng ontology
  - Các mô hình xác suất
- .....  
 Khi biểu diễn tri thức cho một bài toán, việc biểu diễn này được coi là tốt khi chúng thỏa mãn được 4 yếu tố sau
  - Tính hoàn chỉnh (Completeness)
  - Tính ngắn gọn (Conciseness)
  - Tính hiệu quả về tính toán (Computational efficiency)
  - Tính rõ ràng, dễ hiểu (Transparency)
- Với bài toán cụ thể đang xét- bài toán tư vấn mua điện thoại di động, việc biểu diễn tri thức sử dụng luật là một phương pháp biểu diễn hợp lý.

### 1.1. Biểu diễn tri thức dựa trên luật

- Phương pháp biểu diễn tri thức bằng luật sinh được phát minh bởi Newell và Simon trong lúc hai ông đang cố gắng xây dựng một hệ giải bài toán tổng quát. Đây là một kiểu biểu diễn tri thức có cấu trúc, và được sử dụng phổ biến nhất trong các hệ cơ sở tri thức.
- Ý tưởng cơ bản là tri thức có thể được cấu trúc bằng một cặp điều kiện – hành động : "NẾU điều kiện xảy ra THÌ hành động sẽ được thi hành".
- Trong biểu diễn tri thức sử dụng luật, ta cần xác định:
  - Tập các sự kiện F(Facts)  
 $\{A_1, A_2, A_3, \dots, A_n, B_1, B_2, B_3, \dots, B_m\}$
  - Tập các luật R(rules) áp dụng trên các sự kiện có dạng sau:  
 IF  $A_1$  AND  $A_2$  AND ... AND  $A_n$  THEN B  
 (trong đó  $A_1, A_2, \dots, A_n, B$ ) là các sự kiện trong F.  
 $A_1, A_2, A_3, \dots, A_n$  – được gọi là điều kiện (Conditions)  
 B: là kết luận(conclusion).
- Trong biểu diễn tri thức dựa trên luật, cơ chế suy luận trên các luật sản xuất bao gồm hai cơ chế : suy diễn tiến, và suy diễn lùi.
  - Suy diễn tiến: là quá trình suy luận xuất phát từ một số sự kiện ban đầu, xác định các sự kiện có thể sinh ra từ sự kiện này.
  - Suy diễn lùi: là quá trình suy luận ngược xuất phát từ một số sự kiện ban đầu, ta tìm kiếm các sự kiện đã "sinh" ra sự kiện này.
- Vấn đề tối ưu luật: Tập các luật trong một cơ sở tri thức rất có khả năng thừa, trùng lặp hoặc nguy hiểm hơn là mâu thuẫn nhau. Dĩ nhiên là hệ thống có thể đổ lỗi cho người dùng về việc đưa vào hệ thống những tri thức như vậy. Tuy việc tối ưu một cơ sở tri thức về mặt tổng quát là một thao tác khó (vì giữa các tri thức thường có quan hệ không tường minh), nhưng trong giới hạn cơ sở tri thức dưới dạng luật, ta vẫn có một số thuật toán đơn giản để loại bỏ các vấn đề này.
- *Ưu điểm của phương pháp biểu diễn tri thức dựa trên luật :*

Biểu diễn tri thức bằng luật đặc biệt hữu hiệu trong những tình huống hệ thống cần đưa ra những hành động dựa vào những sự kiện có thể quan sát được. Nó có những ưu điểm chính yếu sau đây :

- Các luật rất dễ hiểu nên có thể dễ dàng dùng để trao đổi với người dùng (vì nó là một trong những dạng tự nhiên của ngôn ngữ).
- Có thể dễ dàng xây dựng được cơ chế suy luận và giải thích từ các luật.
- Việc hiệu chỉnh và bảo trì hệ thống là tương đối dễ dàng.
- Có thể cải tiến dễ dàng để tích hợp các luật mờ.

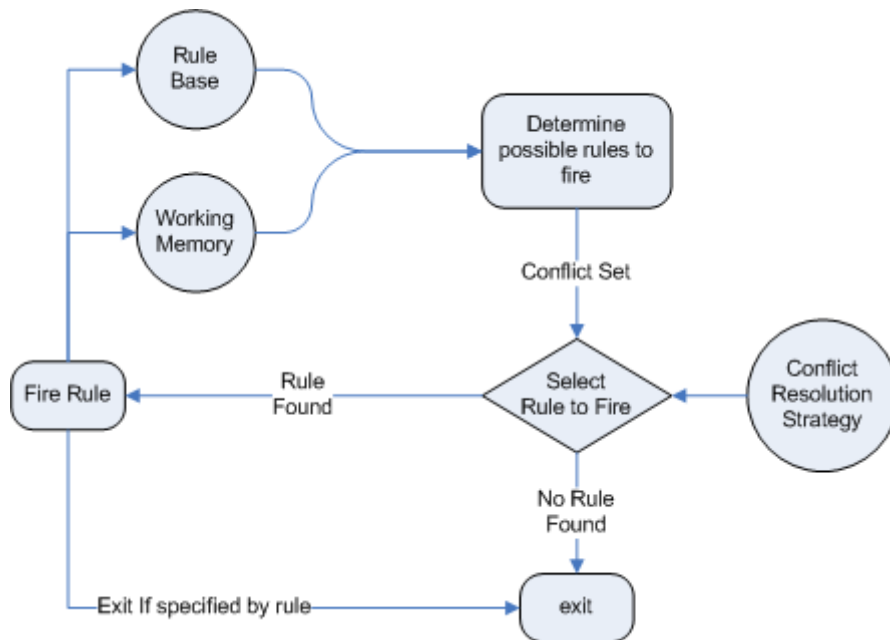
*-Nhược điểm của phương pháp biểu diễn tri thức dựa trên luật.*

- Các tri thức phức tạp đôi lúc đòi hỏi quá nhiều (hàng ngàn) luật sinh. Điều này sẽ làm nảy sinh nhiều vấn đề liên quan đến tốc độ lẫn quản trị hệ thống.
- Thống kê cho thấy, người xây dựng hệ thống trí tuệ nhân tạo thích sử dụng luật sinh hơn tất cả phương pháp khác (dễ hiểu, dễ cài đặt) nên họ thường tìm mọi cách để biểu diễn tri thức bằng luật sinh cho dù có phương pháp khác thích hợp hơn! Đây là nhược điểm mang tính chủ quan của con người.

- Cơ sở tri thức luật sinh lớn sẽ làm giới hạn khả năng tìm kiếm của chương trình điều khiển. Nhiều hệ thống gặp khó khăn trong việc đánh giá các hệ dựa trên luật sinh cũng như gặp khó khăn khi suy luận trên luật sinh.

## 2. Suy diễn tiến

- Là một trong hai cơ chế cơ chế suy luận trên các luật sản xuất trong biểu diễn tri thức dựa trên luật.
- Bài toán cần chứng minh: Với một tập các mệnh đề giả thiết (cơ sở tri thức) KB, cần suy ra mệnh đề kết luận Q. Suy diễn tiến là suy diễn dựa trên dữ liệu nên rất phù hợp với các bài toán đưa ra quyết định, nhận dạng đối tượng.....
- Ý tưởng của suy diễn tiến: Lặp lại 2 bước sau cho đến khi suy ra được kết luận:
  - Áp dụng các luật có mệnh đề giả thiết được thỏa mãn trong KB
  - Bổ sung kết luận của các luật đó vào trong KB.
- Sơ đồ thuật toán.



- Những điểm cần chú ý khi cài đặt giải thuật suy diễn tiến.
  - Tránh các vòng lặp bằng cách kiểm tra xem các mệnh đề mới đã có trong danh sách các mệnh đề cần chứng minh chưa? Nếu rồi thì không bổ sung lại nữa!
  - Tránh việc chứng minh lặp lại đối với một mệnh đề. Mệnh đề này có thể đã được chứng minh là đúng ở trước đó hoặc đã được chứng minh ở trước đó là không thể thỏa mãn được trong KB.
- Ưu điểm của suy diễn tiến
  - Ưu điểm chính của suy diễn tiến là làm việc tốt khi bài toán về bản chất đi thu thập thông tin rồi thấy điều cần suy diễn.
  - Suy diễn tiến cho ra khối lượng lớn các thông tin từ một số thông tin ban đầu. Nó sinh ra nhiều thông tin mới.

- Suy diễn tiến là tiếp cận lý tưởng đối với loại bài toán cần giải quyết các nhiệm vụ như lập kế hoạch, điều hành điều khiển và diễn dịch.
- *Nhược điểm của suy diễn tiến*
  - Một nhược điểm chính của hệ thống suy diễn tiến là không cảm nhận được rằng chỉ một vài thông tin là quan trọng. Hệ thống hỏi các câu hỏi có thể hỏi mà không biết rằng chỉ một ít câu đã đi đến kết luận được.
  - Hệ thống có thể hỏi cả câu không liên quan. Có thể các câu trả lời cũng quan trọng, nhưng làm người dùng lúng túng khi phải trả lời các câu không dính đến chủ đề

### III- Giải quyết bài toán

#### 1. Thu thập tri thức

- Thu thập tri thức là bước quan trọng mở đầu cho mỗi bài toán, đặc biệt đối với bài toán tư vấn ra quyết định.
- Các tri thức cụ thể của bài toán tư vấn mua điện thoại di động.
  - Dựa trên các dữ liệu về điện thoại di động bao gồm: kiểu dáng, màu sắc, hãng sản xuất, giá tiền, tính năng .....
  - Dữ liệu về người sử dụng: tuổi tác, nghề nghiệp, giới tính, sở thích.....
  - Thông tin: mục đích sử dụng, giá tiền, màu sắc phù hợp
  - Tri thức: các dữ liệu về điện thoại phù hợp với các dữ liệu của người sử dụng.  
Ví dụ: người sử dụng là nữ giới thì thường thích màu sắc màu hồng.  
Người sử dụng là nam giới thì thường thích màu đen.  
.....
- Cách thức thu thập thông tin:
  - Dựa trên tìm kiếm thông tin: trên internet dữ liệu về điện thoại di động (kiểu dáng, màu sắc, hãng sản xuất, giá tiền, tính năng..)
  - Dựa trên hỏi đáp: về dữ liệu người sử dụng(tuổi tác, nghề nghiệp, giới tính, sở thích...)
  - Dựa trên thống kê để đưa ra các tri thức.

#### 2. Biểu diễn tri thức

- Trong bài toán cụ thể tư vấn mua điện thoại di động phương pháp biểu diễn tri thức bằng luật là một phương pháp biểu diễn hợp lý.
- Tập các sự kiện F bao gồm các nhóm sự kiện sau:
  - Giá tiền :

GT1	< 2tr
GT2	2tr-5tr
GT3	5tr-10tr
GT4	>10tr

- Màu sắc

MS_ID	Description
M1	Đen
M2	Trắng
M3	Hồng
M4	Xám Bạc

- Kiểu dáng

KD_ID	Description
K1	Thanh dài_có bàn phím
K2	Cảm ứng+ thanh dài
K3	Nấp gập
K4	Trượt

- Hãng sản xuất

HSX_ID	Ten Hang
H1	SamSung
H2	Nokia
H3	BlackBerry
H4	Apple
H5	LG
H6	Sony

- Tính năng

T_ID	Description
T1	2 Sim 2 sóng
T2	Camera, nghe nhạc
T3	Truy cập internet
T4	Hệ Điều Hành
T5	None

- Giới tính người sử dụng.

GT_ID	Value
G1	Nữ
G2	Nam

- Nghề nghiệp

NN_ID	Values
N1	Học sinh/Sinh Viên
N2	Công nhân/ Viên Chức
N3	Lao động tự do
N4	Doanh Nhân



- Sở thích

	ST	Description
▶	ST1	Công Nghệ
	ST2	Giải Trí
	ST3	Tin tức
	ST4	Du lịch

- Điện thoại

D_ID	Name	Description
D1	Nokia 1280	439.000
D10	Nokia N9 64GB	9.489.000
D100	LG KF350	4.339.000
D101	Nokia C2-03	2.169.000
D102	Nokia C2-03 trắng	2.169.000
D103	Nokia Asha 300	2.499.000
D104	Nokia 5233	2.799.000
D105	Nokia C5-06	3.000.000
D106	Nokia Asha 306	1.879.000
D107	Nokia Asha 305	1.679.000
D108	Nokia Asha 311	2.399.000
D109	Nokia E7	6.459.000
D11	Nokia lumia 800	8.949.000
D110	Samsung F250	1.689.000
D111	Nokia 8900 Arte	1.500.000
D112	Nokia 8800 sap...	1.799.000
D113	Samsung C3520	1.260.000
D114	Samsung SGH-J...	1.234.000
D115	Samsung GT-B9...	1.800.000
D12	Nokia lumia 710	5.589.000
D13	Nokia C3-01.5	6.399.000
D14	Nokia 808 pure...	12.249.0...
D15	Nokia lumia 900	10.939.0...
D16	Samsung E1200	449.000

D78	LG C320	678.000
D79	LG Optimous L5	4.679.000
D8	Nokia N311	2.649.000
D80	LG GM360i	2.690.000
D81	LG GT350i	2.199.000
D82	LG GS102	490.000
D83	LG C100	899.000
D84	LG T300	1.249.000
D85	Sony Ericsson T...	2.259.000
D86	Sony Ericsson T...	2.390.000
D87	Sony Ericsson V...	3.690.000
D88	Sony Ericsson X...	2.690.000
D89	Sony Xperia go	7.490.000
D9	Nokia N311	2.649.000
D90	LG Xperia Miro ...	5.290.000
D91	BlackBerry Curv...	3.990.000
D92	BlackBerry Bold ...	8.690.000
D93	BlackBerry Torc...	11.690.0...
D94	New iPad Wifi-4...	13.990.0...
D95	New iPad Wifi-4...	13.690.0...
D96	Samsung Galax...	10.999.0...
D97	Samsung Galax...	15.439.0...
D98	Samsung GT-B9...	4.980.000
D99	Samsung C3520	3.979.000

Điện thoại bao gồm các sự kiện D1,D2.....D115 .

- Tập luật trong bài toán bao gồm 154 luật được kèm theo trong file Rule.txt

*Cú pháp:*

*Ví dụ:* GT4,H4,K2,M2,T4:D94

tương đương với  $GT4 \wedge H4 \wedge K2 \wedge M2 \wedge T4 \rightarrow D94$

*Trong đó*

Mệnh đề điều kiện là : GT4,H4,K2,M2,T4

Mệnh đề kết luận là : D94

### 3. Suy diễn tiến

- Đối với bài toán tư vấn mua điện thoại di động (dạng bài toán đưa ra quyết định) nên phương pháp suy diễn tiến là phương pháp phù hợp.
- Trong nội dung bài tập lớn này, chúng em đã cài đặt thuật toán suy diễn tiến với đầy đủ ý tưởng ban đầu của thuật toán và có thể thêm một số cải tiến để giúp tránh các nhược điểm của thuật toán

*Lặp lại 2 bước sau cho đến khi suy ra được kết luận:*

- Áp dụng các luật có mệnh đề giả thiết được thỏa mãn trong KB
  - Bổ sung kết luận của các luật đó vào trong KB.
- Trong trường hợp xảy ra xung đột luật tức là tại một bước suy diễn có nhiều hơn một bước có thể sử dụng được.

*Giải quyết xung đột luật đồng thời tránh các vòng lặp và tránh việc chứng minh lặp lại đối với một mệnh đề.*

- Không áp dụng các luật sinh ra các kết quả (các sự kiện) đã có trong bộ nhớ làm việc.
- Không áp dụng lại một luật nếu nó vẫn sinh ra cùng một tập các sự kiện (giống như lần áp dụng trước của nó.)

**Ví dụ:**

- Tập giả thiết ban đầu: Người sử dụng: có giới tính là nữ, sở thích công nghệ, nghề nghiệp sinh viên.

KB: G1, NN1, ST1

- Thực hiện suy diễn tiến với tập luật bao gồm
  - Rule1 : G1:M2 (*Giới tính nữ -> màu sắc ưa thích là màu hồng*)
  - Rule2: NN1:GT1 (*nghề nghiệp sinh viên -> mức giá <2triệu*)
  - Rule3: NN1:GT2 (*nghề nghiệp sinh viên -> mức giá 2triệu-5triệu*)
  - Rule4: ST1:H1 (*sở thích công nghệ-> hãng sản xuất 1- Samsung*)
  - Rule5: ST1:H2 (*sở thích công nghệ-> hãng sản xuất 2- Nokia*)
  - Rule6: G1:KD3 (*Giới tính nữ -> kiểu dáng ưa thích là nắp gập*)
  - Rule7: G1:KD2 (*Giới tính nữ -> kiểu dáng ưa thích là cảm ứng*)
  - Rule8: NN1:H2 (*nghề nghiệp sinh viên -> hãng sản xuất là nokia*)

.....

Đến một bước suy diễn nào đó ta thấy nếu ta đã áp dụng luật suy diễn số 5 thì luật suy diễn số 8 sẽ không được áp dụng nữa (vì không áp dụng các luật sinh ra các kết quả đã có trong bộ nhớ). Hoặc ngược lại.

## IV- Cài đặt, demo chương trình

### 1. Cài đặt

- Chương trình được viết bằng : C#
- Chạy bằng visual studio 2010
- Luật được lưu dưới dạng file text (rule.txt) trong file đính kèm.
- Ngoài ra còn có 1 file text(dich.txt) lưu tập các đích có thể có.
- Dữ liệu được lưu trữ dưới dạng các bảng trong database dienthoai (SQL2008)

Bao gồm các bảng

- *Dienthoai*(DT\_ID, GT\_ID,HSX\_ID,MS\_ID,KD\_ID,T\_ID,*Name*,*Description*)
- *GiaDT*(GT\_ID,*Description*)
- *HangSX*(HSX\_ID,*TenHang*)
- *KieuDang*(KD\_ID, *Description*)
- *MauSac* (MS\_ID,*Description*)
- *TinhNang*(T\_ID, *Description*)
- *GioiTinh*(G\_ID,*Description*)
- *NgheNghiep*(N\_ID,*Description*)
- *SoThich*(ST\_ID,*Description*)

## 2. Demo chương trình

- Các chức năng chính của chương trình : search và tư vấn.  
Form giao diện chính.



- Chức năng search cho phép người sử dụng : lựa chọn các giá trị phù hợp về giá tiền, hãng sản xuất, kiểu dáng, màu sắc ưa thích. Thực hiện tìm kiếm và trả về kết



quả

**Tìm kiếm sản phẩm**

Giá tiền: 2tr-5tr | Kiểu dáng: Cảm ứng+ thanh dài

Hãng Sản Xuất: SamSung | Màu sắc: Đen

Search

Kết quả tìm kiếm

	Name	Description
▶	Samsung Torino ...	3.679.000
	Samsung Galaxy Y	2.899.000
	Samsung Galaxy ...	3.890.000
	Samsung Monte ...	2.469.000

- Chức năng tư vấn: cho phép người sử dụng nhập vào thông tin cá nhân bao gồm Giới tính, nghề nghiệp, sở thích. Trả về kết quả tư vấn.

**Phiên tư vấn mua điện thoại di động**

Nhập vào tên bạn: Nguyễn Thị Thúy | Xin chào: Nguyễn Thị Thúy

Giới tính:  Nam  Nữ

Nghề Nghiệp: Học sinh/Sinh Viên | Sở Thích: Du lịch

Tu Van

	Name	Price
▶	Nokia C2-01	1.629.00
*		

## V- Kết luận đánh giá

### 1. Các vấn đề gặp phải

- Trong quá trình xây dựng chương trình, thu thập và biểu diễn tri thức là hai vấn đề xảy ra nhiều sai sót và lỗi.
  - Trong quá trình thu thập : vì số lượng thông tin hỏi đáp không nhiều (134 hỏi đáp) nên kết quả khi thống kê chưa thật chính xác.
  - Khi biểu diễn tri thức: nhiều mâu thuẫn luật xảy ra.  
Ví dụ: Nam sinh viên thường thích công nghệ G2:ST1  
Công nghệ được đa số người được phỏng vấn chọn là hãng Apple, Samsung.  
Sinh viên thường có mức giá khi mua GT1, GT2 (<5tr)  
Hãng Apple thì không cung cấp điện thoại nào có giá dưới 5 triệu.  
.....
- Trong quá trình cài đặt chương trình.
  - Các luật đều được biểu diễn dạng thành từng dòng trong file txt. Khi tách luật ra thành hai vế trái và phải thường bị thừa kí tự kết thúc dòng.
  - Kết quả khi suy diễn là chuỗi string các đích có thể cách nhau bằng dấu ‘ , ’  
Cần tách đích và hiển thị kết quả và thường bị lỗi ở kí tự ‘ , ’ sau cùng.

### 2. Hướng phát triển của đề tài

- Xây dựng mở rộng thêm tri thức, cho phép học và bổ sung thêm các luật vào cơ sở tri thức, cải tiến các tri thức, nâng cao độ chính xác
- Thử nghiệm phương thức biểu diễn tri thức bằng thông tin không chắc chắn dựa trên lý thuyết xác suất.
- Mở rộng tư vấn mua các sản phẩm công nghệ khác: máy tính.....

### 3. Đánh giá

- Chương trình còn nhiều hạn chế về kết quả tư vấn vì số lượng tập luật, lượng dữ liệu còn hạn chế.
- Giao diện còn đơn giản
- Chưa hiển thị hình ảnh của điện thoại.