

# HỆ HỖ TRỢ QUYẾT ĐỊNH

## Bài 15: Hệ chuyên gia

Lê Hải Hà

# Nội dung

- ➊ Các khái niệm cơ bản
- ➋ Các ứng dụng của hệ chuyên gia
- ➌ Cấu trúc của hệ chuyên gia
- ➍ Kỹ nghệ tri thức
- ➎ Sự phát triển của hệ chuyên gia
- ➏ Các lĩnh vực/vấn đề được giải quyết bởi ES
- ➐ Lợi ích, giới hạn và các nhân tố thành công của ES

# Hệ chuyên gia (ES)



- Là một chương trình máy tính mà cố bắt chước các quy trình lập luận và tri thức của chuyên gia khi giải quyết những vấn đề cụ thể nào đó
- **Công nghệ AI được ứng dụng phổ biến**
  - Nâng cao năng suất
  - Tăng cường lực lượng lao động
- Làm việc tốt với các vấn đề trong lĩnh vực hẹp
- Các hệ chuyên gia không thay thế các chuyên gia, nhưng
  - Làm cho tri thức và kinh nghiệm của họ được phổ biến rộng rãi hơn, và do đó
  - Cho phép những người không là chuyên gia làm việc tốt hơn

# Các khái niệm quan trọng của hệ chuyên gia



- **Chuyên gia**

Người đã phát triển tới một mức cao của sự thành thạo trong việc đưa ra các phán quyết trong một lĩnh vực cụ thể

- **Chuyên môn**

Tập các khả năng mà thông qua nó con người hoàn thành công việc, bao gồm

- ✓ Tri thức chuyên ngành rộng,
- ✓ Các quy tắc thực tế giúp đơn giản và cải tiến cách tiếp cận giải quyết vấn đề,
- ✓ Siêu kiến thức và siêu nhận thức, và
- ✓ Các dạng hành vi được biên tập để có khả năng kinh tế lớn với hiệu suất lành nghề



# Các khái niệm quan trọng của hệ chuyên gia

- Chuyên gia
  - Mức độ chuyên môn
  - Tỷ lệ không chuyên gia/chuyên gia thường là 100/1
- Chuyển giao chuyên môn
  - Từ chuyên gia sang máy tính và tới những người không phải chuyên gia thông qua việc thu thập, thể hiện, suy luận và chuyển giao
- Suy luận
  - Tri thức = Thực tế + thủ tục (các luật)
  - Lập luận/suy nghĩ được thực hiện bởi máy tính
- Các luật (IF ... THEN ...)
- Khả năng giải thích (Tại sao? Như thế nào?)

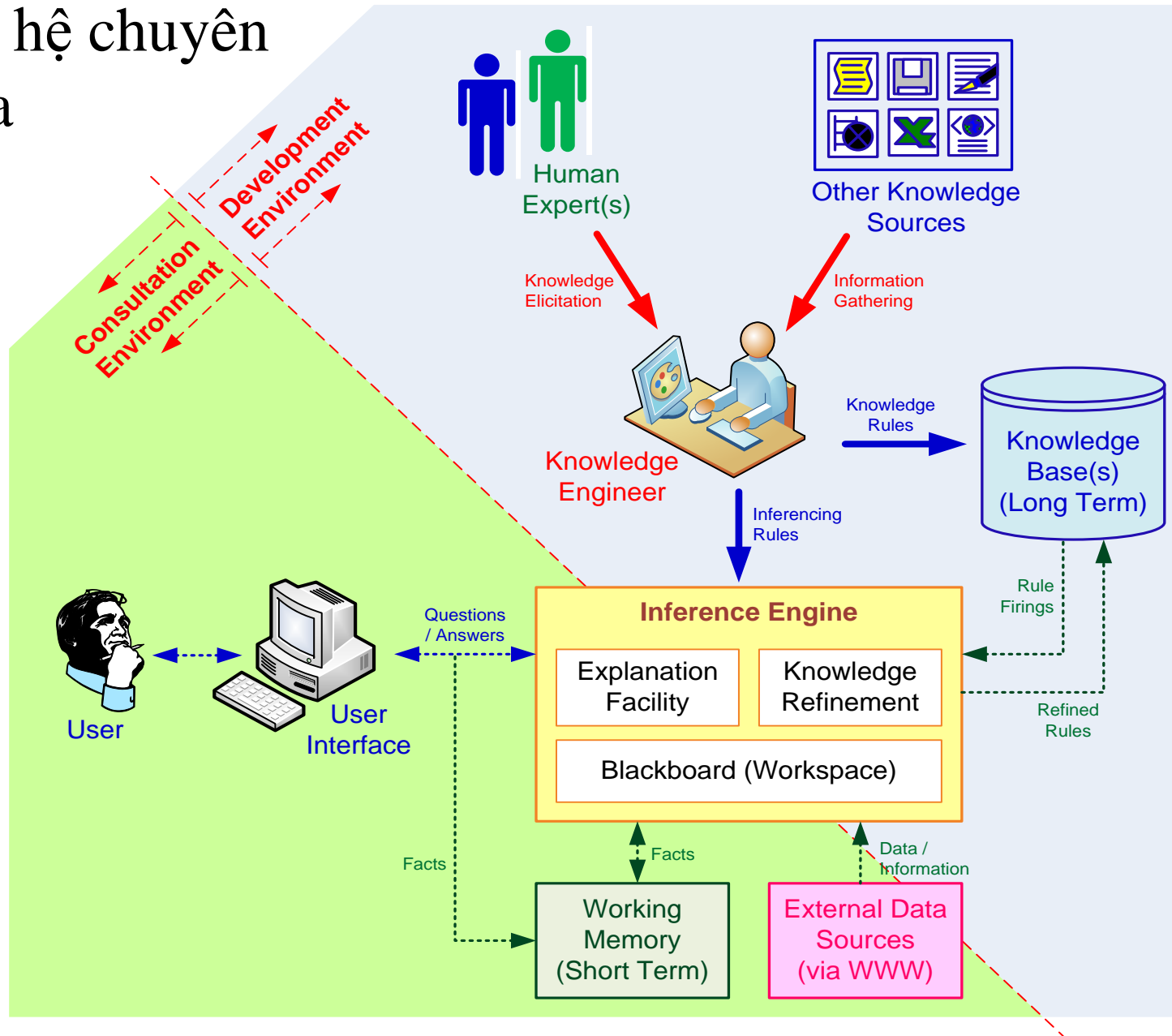


# Các ứng dụng của hệ chuyên gia

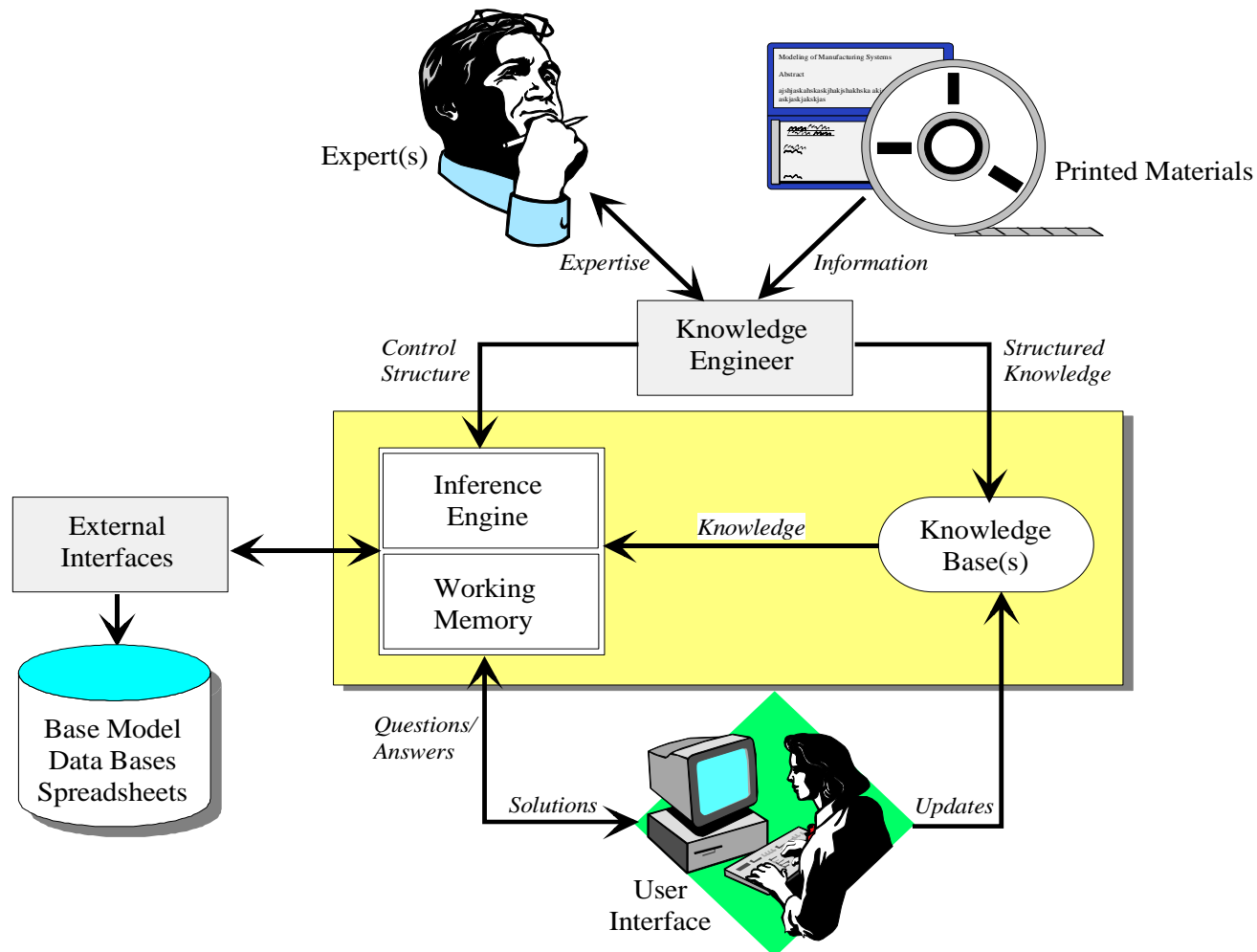
- DENDRAL
  - Ứng dụng tri thức (lập luận dựa trên các luật)
  - Suy đoán cấu trúc phân tử của các hợp chất
- MYCIN
  - Hệ chuyên gia dựa trên luật
  - Được sử dụng để chẩn đoán và điều trị các bệnh nhiễm trùng do vi khuẩn
- XCON
  - Hệ chuyên gia dựa trên luật
  - Được sử dụng để xác định cấu hình các hệ thống tin tối ưu
- **Các ứng dụng mới:** Credit analysis, Marketing, Finance, Manufacturing, Human resources, Science and Engineering, Education, ...

# Cấu trúc của hệ chuyên gia

1. Môi trường phát triển
2. Môi trường tư vấn (chạy)



# Kiến trúc khái niệm của một hệ chuyên gia chuẩn tắc





# Thành phần con người trong ES

- Chuyên gia
  - Có tri thức, phán quyết, kinh nghiệm và các phương pháp đặc biệt để đưa ra lời khuyên và giải quyết vấn đề
- Kỹ nghệ tri thức
  - Giúp các chuyên gia cấu trúc miền vấn đề bằng cách diễn dịch và tích hợp các câu trả lời câu hỏi của con người, rút ra điều tương tự, đặt ra các phản ví dụ và khai sáng các khó khăn về khái niệm
- Người dùng
- Người khác
  - System Analyst, Builder, Support Staff, ...

# Cấu trúc ES

- 3 thành phần chính của ES là:
  - Cơ sở tri thức
  - Máy suy diễn
  - Giao diện người dùng
- ES có thể cũng bao gồm:
  - Hệ thống con thu thập tri thức
  - Bảng đen (workplace)
  - Hệ thống con giải thích (justifier)
  - Hệ thống làm tốt/sàng lọc tri thức

# Cấu trúc ES

- **Thu thập tri thức (KA)**

Trích và hình thành tri thức được dẫn xuất từ một loạt nguồn, đặc biệt từ các chuyên gia (khám phá)

- **Cơ sở tri thức**

Tập các sự kiện, luật, và thủ tục được tổ chức vào các lược đồ. Tập hợp các thông tin và tri thức về một lĩnh vực liên quan cụ thể

- **Bảng đen (*bộ nhớ làm việc*)**

Vùng bộ nhớ làm việc được dành riêng cho việc mô tả vấn đề hiện tại và để ghi lại các kết quả trung gian trong một hệ chuyên gia

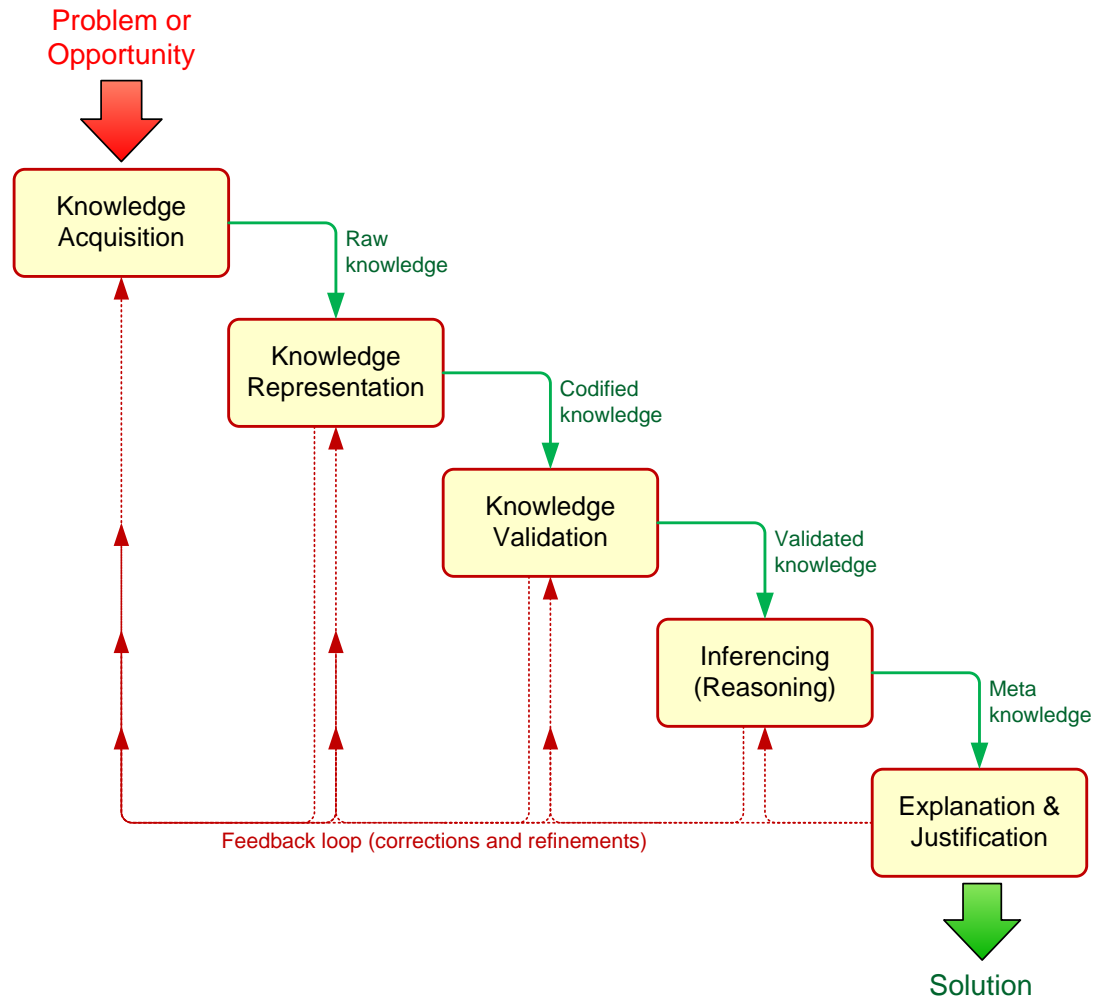
- **Hệ thống con giải thích (*justifier*)**

Thành phần của một hệ chuyên gia mà có thể giải thích suy luận và điều chỉnh kết luận chuyên gia

# Kỹ nghệ tri thức (KE)

- Tập hợp các hoạt động chuyên sâu, bao gồm thu thập tri thức từ các chuyên gia con người (và các nguồn tri thức khác) và chuyển đổi tri thức này vào kho lưu trữ (gọi là cơ sở tri thức)
- Mục đích chính của kỹ nghệ tri thức là
  - để giúp các chuyên gia trình bày rõ ràng cách họ làm, những gì họ làm, và
  - để tài liệu hóa tri thức này ở dạng có thể sử dụng lại
- Định nghĩa hẹp và rộng của kỹ nghệ tri thức?

# Quy trình kỹ nghệ tri thức



# Các nhóm tri thức chính trong ES

- Tri thức khai báo/Declarative Knowledge
  - Mô tả thể hiện của tri thức mà liên quan tới một đối tượng cụ thể.
  - Nôm – được thể hiện ở một phát biểu thực tế
  - Quan trọng trong giai đoạn đầu của việc thu thập tri thức
- Tri thức thủ tục/Procedural Knowledge
  - Cân nhắc cách mọi vật làm việc dưới các tình huống khác nhau
  - Bao gồm tuần tự các bước và tập các chỉ thị
- Siêu tri thức/Metaknowledge
  - Tri thức về tri thức

# ES làm việc thế nào: Cơ chế suy luận

- **Thể hiện và tổ chức tri thức**
  - Tri thức chuyên gia phải được thể hiện ở dạng máy tính có thể hiểu và được tổ chức hoàn hảo trong cơ sở tri thức
  - Các cách khác nhau thể hiện tri thức con người bao gồm:
    - Các quy tắc tạo ra (\*)
    - Các mạng ngữ nghĩa
    - Các phát biểu logic

# Một số dạng quy tắc/luật

- Tiền đề IF, kết luận THEN
  - Nếu thu nhập của bạn cao, thì cơ hội được kiểm toán của IRS cũng cao
- Kết luận, tiền đề IF
  - Cơ hội được kiểm toán của bạn cao, nếu thu nhập của bạn cao
- Bao gồm ELSE
  - Nếu thu nhập của bạn cao, hoặc các khoản khấu trừ của bạn bất thường, thì cơ hội bạn được kiểm toán bởi IRS là cao, ngược lại cơ hội được kiểm toán là thấp
- Nhiều luật phức tạp
  - Nếu xếp hạng tín dụng cao và lương hơn \$30,000, hoặc tài sản nhiều hơn \$75,000, và lịch sử thanh toán không “tồi”, thì chấp thuận khoản vay lên tới \$10,000, và liệt kê khoản mục vay trong nhóm "B.”



# Tri thức và các luật suy luận

- Hai dạng quy tắc chung trong AI:
  - Các quy tắc/luật tri thức và các quy tắc suy diễn
- Các quy tắc tri thức (Knowledge rules/declarative rules), phát biểu tất cả các sự kiện thực tế và các quan hệ về vấn đề
- Các quy tắc suy diễn (Inference rules/procedural rules), tư vấn về cách giải quyết vấn đề, khi một số sự kiện thực tế nhất định được biết
- Các quy tắc suy diễn chứa các quy tắc về quy tắc (siêu quy tắc - metarules)
- Các quy tắc tri thức được lưu trong cơ sở tri thức
- Các quy tắc suy diễn trở thành phần của máy suy diễn
- **Thí dụ:**
  - Nếu không biết dữ liệu cần thiết thì hãy hỏi người dùng
  - Nếu nhiều hơn một quy tắc được áp dụng, thì kích hoạt quy tắc với độ ưu tiên cao trước

# ES làm việc thế nào: Cơ chế suy luận

**Suy diễn/Inference** là tiến trình xâu chuỗi nhiều quy tắc cùng nhau dựa trên dữ liệu đã có

- **Xâu chuỗi hướng thẳng/Forward chaining**

Tìm kiếm hướng dữ liệu (data-driven) trong hệ thống dựa trên luật

Nếu tiền đề phù hợp với tính hướng, thì tiến trình cố khẳng định kết luận

- **Xâu chuỗi ngược/Backward chaining**

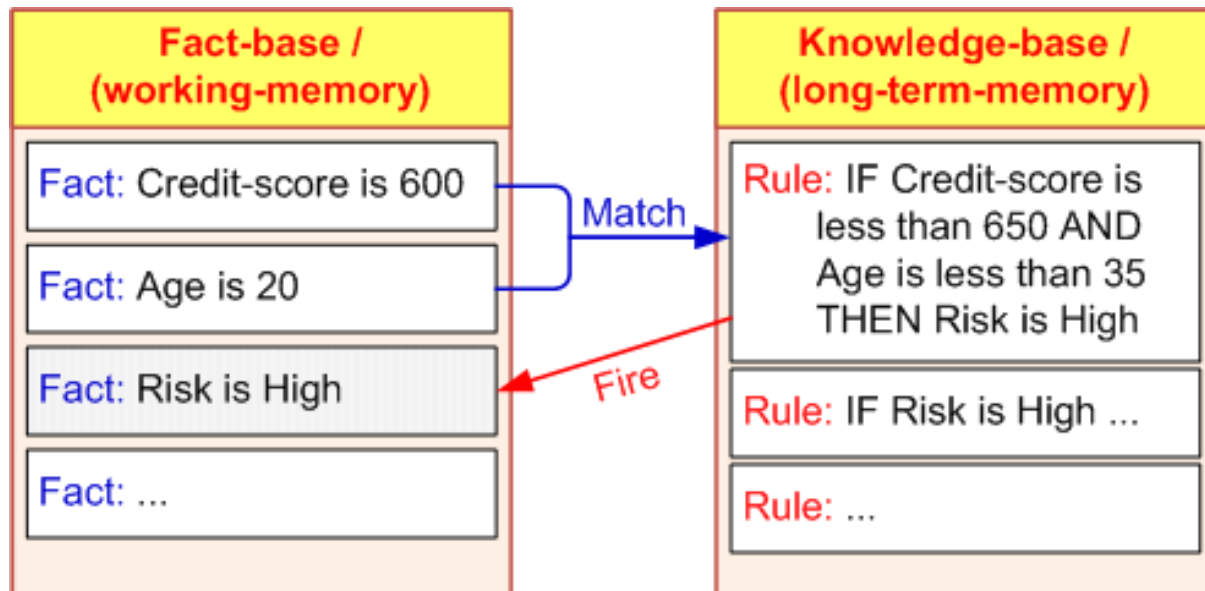
Tìm kiếm hướng đích (goal-driven) trong hệ thống dựa trên luật

Nó bắt đầu từ mệnh đề hành động của một quy tắc và làm việc theo hướng ngược thông qua chuỗi các quy tắc với nỗ lực tìm ra tập các mệnh đề điều kiện có thể xác minh được

# Suy luận với các luật:

## Chuỗi hướng thẳng và quay lui

- Thực hiện một quy tắc
  - Khi tất cả các giả thiết của một quy tắc (phần if) được thỏa mãn, quy tắc được thực hiện (fired)
  - Máy suy diễn kiểm tra mọi quy tắc trong cơ sở tri thức theo hướng thẳng và quay lui để tìm các luật được thực hiện
  - Tiếp tục tới khi không có luật nào có thể thực hiện, hoặc tới khi đạt được mục đích



# Chuỗi quay lui

- **Hướng mục đích/Goal-driven:** Bắt đầu với một kết luận tiềm năng (giả thiết), sau đó tìm kiếm bằng chứng hỗ trợ (hoặc mâu thuẫn) với nó
- Thường bao gồm việc hình thành và kiểm định các giả thiết (hay giả thiết con)

## Knowledge Base

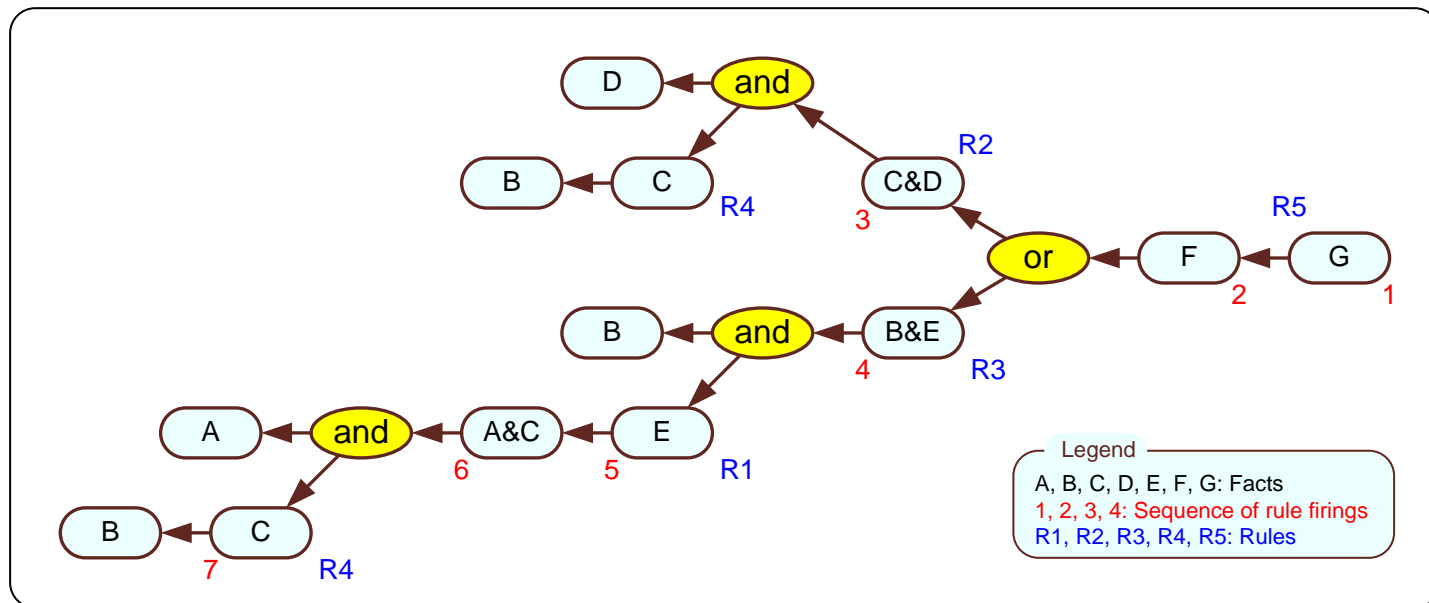
Rule 1: A & C -> E

Rule 2: D & C -> F

Rule 3: B & E -> F (invest in growth stocks)

Rule 4: B -> C

Rule 5: F -> G (invest in IBM)



# Chuỗi hướng thẳng

- **Hướng dữ liệu/Data-driven:** Bắt đầu với thông tin cung cấp khi nó sẵn dùng, sau đó cố rút ra các kết luận
- Sử dụng cái nào?
  - Nếu tất cả các sự kiện thực tế sẵn dùng trước chuỗi hướng thẳng
  - Các vấn đề chuẩn đoán – chuỗi quay lui

## Knowledge Base

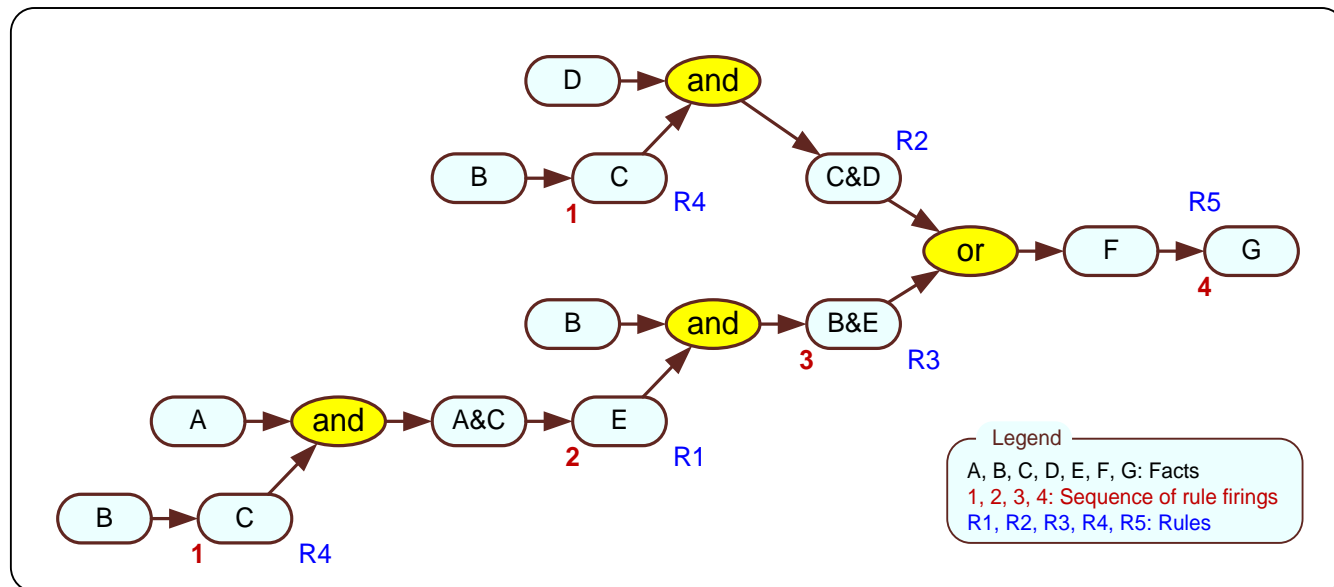
Rule 1: A & C -> E

Rule 2: D & C -> F

Rule 3: B & E -> F (invest in growth stocks)

Rule 4: B -> C

Rule 5: F -> G (invest in IBM)



# Các vấn đề suy luận

- Cách ta chọn giữa chuỗi quay lui và hướng thẳng
  - Theo cách chuyên gia giải quyết vấn đề
    - Nếu trước hết chuyên gia thu thập dữ liệu và suy luận từ chúng  
=> Chuỗi hướng thẳng
    - Nếu chuyên gia bắt đầu với một giải pháp giả thuyết và sau đó cố gắng để tìm các sự kiện thực tế để chứng tỏ nó => Chuỗi quay lui

- Cách xử lý các quy tắc xung đột

IF A & B THEN C

IF X THEN C

1. Thiết lập mục đích và dừng thực hiện các quy tắc khi mục đích đã đạt được
2. Thực hiện luật với độ ưu tiên cao nhất
3. Thực hiện luật cụ thể nhất
4. Thực hiện luật mà sử dụng dữ liệu được nhập vào gần nhất

# Suy luận với sự không chắc chắn

## Lý thuyết chắc chắn (Các nhân tố chắc chắn)

- Các nhân tố và niềm tin chắc chắn
- Không chắc chắn thể hiện mức độ tin cậy
- Thể hiện số đo của độ tin cậy
- Thao tác mức độ tin cậy trong khi sử dụng các hệ dựa trên tri thức
- Nhân tố chắc chắn/Certainty Factors (CF) thể hiện sự tin cậy vào sự kiện dựa trên bằng chứng (hoặc đánh giá của chuyên gia)
  - 1.0 hay 100 = tin tưởng tuyệt đối (hoàn toàn tin cậy)
  - 0 = chắc chắn giả dối
- CF không là xác suất
- CF không có tổng là 100

# Suy luận với sự không chắc chắn

## Kết hợp các nhân tố chắc chắn

- Sự kết hợp một vài nhân tố chắc chắn trong một luật ở đó sử dụng kết hợp các phép toán logic **AND** và **OR**

- **AND**

IF **inflation is high**, CF = 50 percent, (A), AND

**unemployment rate** is above 7, CF = 70 percent, (B), AND

**bond prices decline**, CF = 100 percent, (C)

THEN **stock prices decline**

$$CF(A, B, \text{ and } C) = \text{Minimum}[CF(A), CF(B), CF(C)]$$

=>

- The CF for “stock prices to decline” = 50 percent
- The chain is as strong as its weakest link



# Suy luận với sự không chắc chắn

## Kết hợp các nhân tố chắc chắn

- OR

IF inflation is low, CF = 70 percent, (A), OR

bond prices are high, CF = 85 percent, (B)

THEN stock prices will be high

$$CF(A, B) = \text{Maximum}[CF(A), CF(B)]$$

=>

– The CF for “stock prices to be high” = 85 percent

– Notice that in OR only one IF premise needs to be true

# Suy luận với sự không chắc chắn

## Kết hợp các nhân tố chắc chắn

- Kết hợp của 2 hay nhiều luật
  - Thí dụ:
    - R1: IF the inflation rate is less than 5 percent,  
THEN stock market prices go up (CF = 0.7)
    - R2: IF unemployment level is less than 7 percent,  
THEN stock market prices go up (CF = 0.6)
  - Inflation rate = 4 percent and the unemployment level = 6.5 percent
  - Combined Effect
    - $CF(R1,R2) = CF(R1) + CF(R2)[1 - CF(R1)]$ ; or
    - $CF(R1,R2) = CF(R1) + CF(R2) - CF(R1) \times CF(R2)$

# Suy luận với sự không chắc chắn

## Kết hợp các nhân tố chắc chắn

- Thí dụ tiếp

Given  $\underline{CF(R1)} = 0.7$  AND  $\underline{CF(R2)} = 0.6$ , then:

$$CF(R1, R2) = 0.7 + 0.6(1 - 0.7) = 0.7 + 0.6(0.3) = 0.88$$

- Expert System tells us that there is an 88 percent chance that stock prices will increase
- For a third rule to be added

$$CF(R1, R2, R3) = CF(R1, R2) + CF(R3) [1 - CF(R1, R2)]$$

R3: IF bond price increases THEN stock prices go up (CF = 0.85)

Assuming all rules are true in their IF part, the chance that stock prices will go up is

$$CF(R1, R2, R3) = 0.88 + 0.85 (1 - 0.88) = 0.982$$

# Suy luận với sự không chắc chắn

## Các nhân tố chắc chắn – Thí dụ

- **Các quy tắc**

R1: IF blood test result is yes

THEN the disease is malaria (CF 0.8)

R2: IF living in malaria zone

THEN the disease is malaria (CF 0.5)

R3: IF bit by a flying bug

THEN the disease is malaria (CF 0.3)

- **Câu hỏi**

What is the CF for having malaria (as its calculated by ES), if

1. The first two rules are considered to be true ?
2. All three rules are considered to be true?

# Suy luận với sự không chắc chắn

## Các nhân tố chắc chắn – Thí dụ

- Câu hỏi

What is the CF for having malaria (as its calculated by ES), if

1. The first two rules are considered to be true ?

2. All three rules are considered to be true?

- Câu trả lời 1

$$\begin{aligned} 1. \text{CF}(\text{R1}, \text{R2}) &= \text{CF}(\text{R1}) + \text{CF}(\text{R2}) * (1 - \text{CF}(\text{R1})) \\ &= 0.8 + 0.5 * (1 - 0.8) = 0.8 - 0.1 = 0.9 \end{aligned}$$

- Câu trả lời 2

$$\begin{aligned} 2. \text{CF}(\text{R1}, \text{R2}, \text{R3}) &= \text{CF}(\text{R1}, \text{R2}) + \text{CF}(\text{R3}) - (\text{CF}(\text{R1}, \text{R2}) * \text{CF}(\text{R3})) \\ &= 0.9 + 0.3 - (0.9 * 0.3) = 1.2 - 0.27 = 0.93 \end{aligned}$$

# Giải thích như siêu tri thức

- Giải thích
  - Các chuyên gia con người biện minh và giải thích các hành động của họ ... nên ES cũng vậy
  - **Giải thích**: nỗ lực bởi ES để làm sáng tỏ lý do, khuyến nghị, các hành động khác (như đặt câu hỏi)
  - Cơ sở giải thích = minh chứng (Explanation facility = Justifier)
- Giải thích mục đích...
  - Làm cho hệ thống dễ hiểu hơn
  - Khám phá những thiếu sót của cơ sở tri thức (gỡ lỗi)
  - Giải thích các tình huống không lường trước
  - Đáp ứng các yêu cầu tâm lý xã hội của người dùng
  - Làm sáng tỏ các giả thiết bên dưới của các hoạt động của hệ thống
  - Tiến hành phân tích độ nhạy

# Hai giải thích cơ bản

- **Giải thích tại sao** – Tại sao thực tế xảy ra?
- **Giải thích thế nào** – Để xác định cách đạt được một kết luận hay một khuyến nghị nào đó
  - Một số hệ thống đơn giản – chỉ ở kết luận cuối cùng
  - Đại đa số các hệ thống phức tạp cung cấp chuỗi luật được sử dụng để có được kết luận
- Giải thích là cần thiết trong ES
- Được sử dụng để luyện và đánh giá

# ES làm việc thế nào: Cơ chế suy luận

- Quy trình phát triển ES
  - Quy trình phát triển ES chuẩn tắc gồm:
    - Thu thập tri thức
    - Thể hiện tri thức
    - Lựa chọn các công cụ phát triển
    - Lập nguyên mẫu hệ thống
    - Đánh giá
    - Cải tiến/bảo trì



# Sự phát triển của ES

- Xác định bản chất và phạm vi của vấn đề
  - ES dựa trên luật/quy tắc phù hợp với vấn đề mà bản chất là định tính, tri thức tường minh và chuyên gia sẵn có để giải quyết vấn đề hiệu quả và cung cấp tri thức của họ
- Xác định các chuyên gia thích hợp
  - Một chuyên gia thích hợp phải có hiểu biết thấu đáo về:
    - Tri thức giải quyết vấn đề
    - Vai trò của ES và công nghệ hỗ trợ quyết định
    - Kỹ năng giao tiếp tốt

# Sự phát triển của ES

- Thu thập tri thức

- **Kỹ sư tri thức**

Một chuyên gia AI chịu trách nhiệm về khía cạnh kỹ thuật trong việc phát triển hệ chuyên gia. Kỹ sư tri thức làm việc chặt chẽ với chuyên gia lĩnh vực để thu thập tri thức của chuyên gia

- **Kỹ nghệ tri thức (KE)**

Nguyên tắc kỹ thuật dùng để tích hợp tri thức vào hệ thống máy tính để giải quyết các vấn đề phức tạp mà thông thường đòi hỏi mức chuyên môn con người cao

# Sự phát triển của ES

- Lựa chọn các công cụ xây dựng
  - Môi trường phát triển mục đích chung
  - **Expert system shell** (thí dụ: ExSys hay Corvid)...  
Chương trình máy tính tạo điều kiện dễ dàng cho việc thể hiện một hệ chuyên gia cụ thể
  - **Lựa chọn công cụ phát triển ES**
    - Cân nhắc các lợi ích chi phí
    - Cân nhắc chức năng và sự mềm dẻo của công cụ
    - Cân nhắc sự tương thích của công cụ với hạ tầng thông tin đang có
    - Cân nhắc sự tin cậy và sự hỗ trợ của đối tác

# Sự phát triển của ES

- Coding (thể hiện) hệ thống
  - Mỗi quan tâm chính ở giai đoạn này là liệu việc viết mã (hay thể hiện) có được quản lý phù hợp để tránh lỗi...
- Đánh giá hệ chuyên gia
  - Đánh giá/Evaluation
  - Xác minh/Verification
  - Thẩm định/Validation

# Sự phát triển của ES - Validation and Verification of the ES

- Evaluation

- Đánh giá giá trị tổng thể của hệ chuyên gia
- Phân tích xem hệ thống có thể sử dụng được, hiệu quả và tiết kiệm chi phí

- Validation

- Xem xét hiệu suất của hệ thống (so sánh với chuyên gia)
- Đã xây dựng “đúng” hệ thống (mức độ chính xác có thể chấp nhận được không?)

- Verification

- Hệ thống đã được xây dựng “đúng” không?
- Hệ thống có được triển khai đúng với đặc tả kỹ thuật không?

# Các lĩnh vực/vấn đề được giải quyết bởi ES

- Interpretation systems
- Prediction systems
- Diagnostic systems
- Repair systems
- Design systems
- Planning systems
- Monitoring systems
- Debugging systems
- Instruction systems
- Control systems, ...



# Lợi ích của ES

- Nắm bắt kiến thức chuyên môn khan hiếm
- Tăng năng suất và chất lượng
- Giảm thời gian ra quyết định
- Giảm thời gian ngừng hoạt động thông qua chuẩn đoán
- Vận hành thiết bị dễ dàng
- Loại bỏ thiết bị đắt tiền
- Khả năng giải quyết các vấn đề phức tạp
- Chuyển giao tri thức tới các vị trí xa
- Tích hợp ý kiến của một số chuyên gia
- Có thể làm việc với thông tin không chắc chắn
- ... hơn nữa ...

# Các vấn đề và giới hạn của ES

- Không phải lúc nào cũng có sẵn tri thức
- Chuyên môn có thể khó rút ra từ con người
  - Sợ chia sẻ chuyên môn
  - Tăng xung đột khi làm việc với nhiều chuyên gia
- ES làm việc tốt chỉ trong miền nhỏ của tri thức
- Từ vựng của chuyên gia thường có tính kỹ thuật cao
- Các kỹ sư tri thức hiếm và đắt đỏ
- Thiếu sự tin cậy của người dùng cuối
- ES thỉnh thoảng sinh ra các khuyến nghị không chính xác
- ... nhiều nữa ...



# Các nhân tố ES thành công

- Các nhân tố quan trọng nhất

- Có người quản lý tốt
- Có sự tham gia và đào tạo người dùng
- Xác minh sự quan trọng của vấn đề
- Quản lý dự án tốt

- Cộng thêm

- Mức độ tri thức phải đủ cao
- Phải có (ít nhất) một chuyên gia cộng tác
- Vấn đề chủ yếu là định tính
- Vấn đề đủ hẹp trong phạm vi
- ES shell phải có chất lượng cao, với giao diện người dùng thân thiện, và lưu trữ và sử dụng tri thức một cách tự nhiên

# Tuổi thọ của ES thương mại

- Chỉ 1/3 tồn tại hơn 5 năm
- Nói chung ES thất bại do các vấn đề quản lý
  - Thiếu sự chấp nhận hệ thống của người dùng
  - Không có khả năng giữ chân các nhà phát triển
  - Các vấn đề trong việc chuyển giao từ phát triển sang bảo trì (thiếu sự sàng lọc)
  - Thay đổi các ưu tiên của tổ chức
- Quản lý phù hợp việc phát triển và triển khai ES có thể giải quyết hầu hết vấn đề trên