Học viện Công nghệ Bưu chính Viễn thông
Khoa Công nghệ thông tin 1

Nhập môn trí tuệ nhân tạo

Mạng Bayes

Ngô Xuân Bách

Nội dung

- Định nghĩa và cách xây dựng mạng Bayes
- Suy diễn với mạng Bayes

Vấn đề biểu diễn xác suất

- Bài toán suy diễn:
  - Cho bằng chứng $E_1, E_2, \dots, E_n$
  - Cần xác định yêu cầu Q bằng cách tính $P(Q|E_1, E_2, \dots, E_n)$
- Nếu có tất cả các xác suất đồng thời
  - Có thể tính xác suất điều kiện trên
- Bảng xác suất đồng thời có kích thước tăng theo hàm mũ của số biến
  - Quá lớn trên thực tế

Cần có cách biểu diễn và suy diễn thực tế hơn

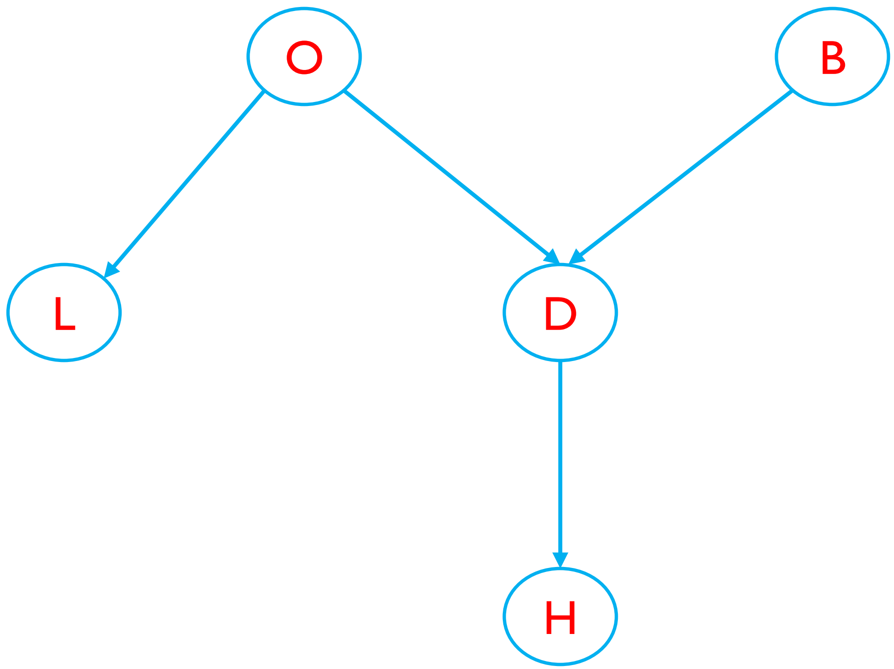
Ví dụ (1/2)

- Bài toán: Một người đi làm về, cần đoán trong nhà có người không?
- Biết rằng:
  - Nếu người nhà đi vắng thì thường (nhưng không luôn luôn) bật đèn ngoài sân
  - Khi không có người ở nhà thì thường buộc chó ở bên ngoài
  - Nếu chó bị ốm cũng bị buộc ở bên ngoài
  - Nếu chó ở ngoài thì có thể nghe tiếng sủa

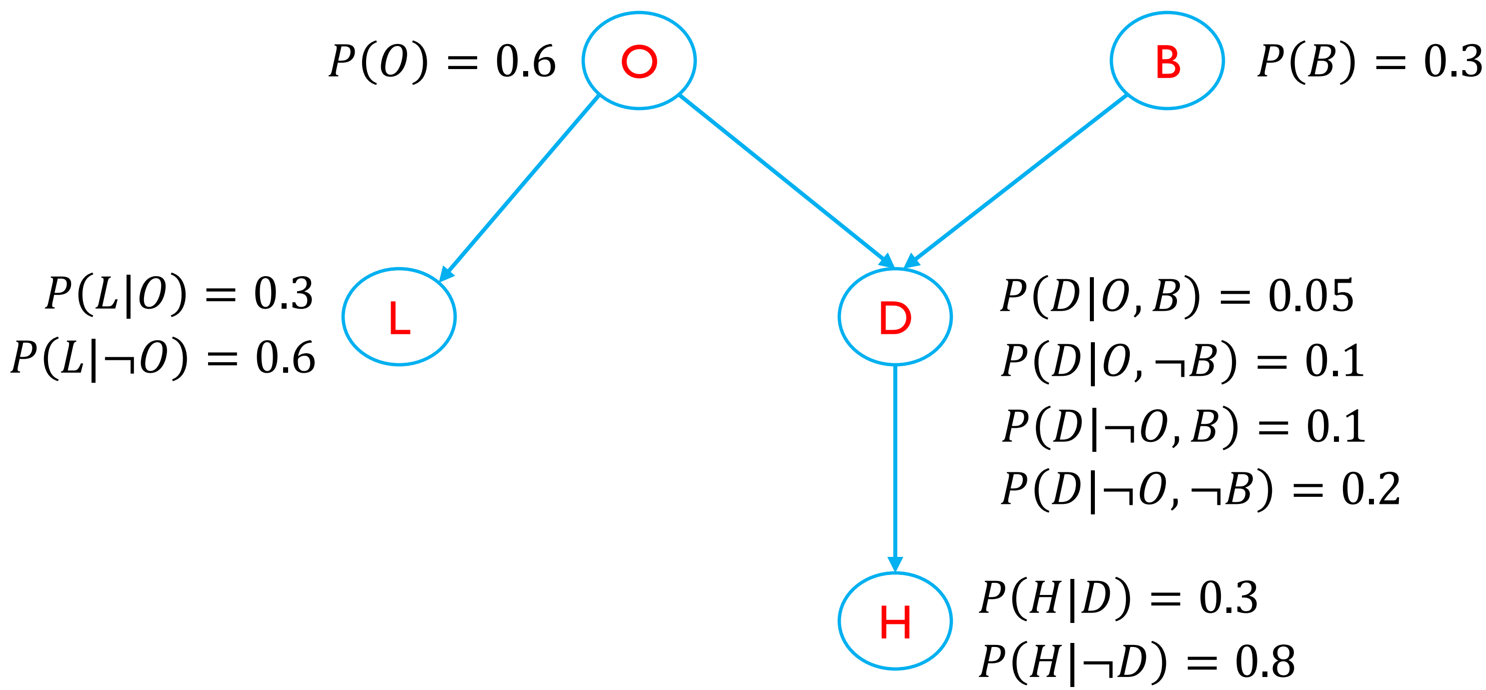
Ví dụ (2/2)

- Xác định 5 biến ngẫu nhiên sau
  - O : nhà không có người
  - L : đèn ngoài sân sáng
  - D : chó buộc ở ngoài
  - B : chó bị ốm (đau bụng)
  - H : nghe thấy tiếng sủa

Quan hệ giữa các nút



Mạng Bayes

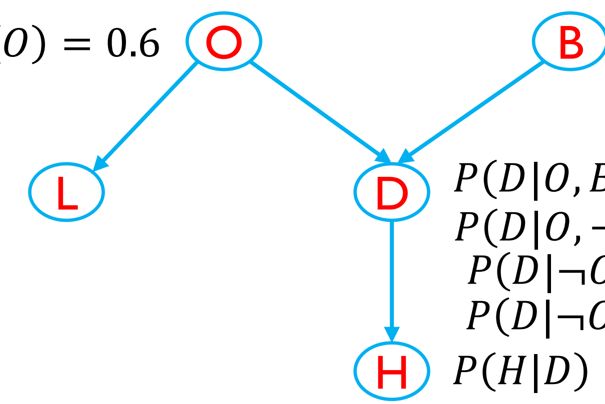

$P(O) = 0.6$
O
B
$P(B) = 0.3$
$P(L|O) = 0.3$
$P(L|\neg O) = 0.6$
L
D
$P(D|O,B) = 0.05$
$P(D|O,\neg B) = 0.1$
$P(D|\neg O,B) = 0.1$
$P(D|\neg O,\neg B) = 0.2$
H
$P(H|D) = 0.3$
$P(H|\neg D) = 0.8$

Định nghĩa mạng Bayes

- Mạng Bayes bao gồm 2 phần
  - Phần thứ nhất là đồ thị có hướng, không chu trình, trong đó mỗi nút ứng với một biến ngẫu nhiêu, mỗi cạnh (có hướng) biểu diễn cho quan hệ giữa nút gốc và nút đích
  - Phần thứ hai là bảng xác suất điều kiện chứa xác suất điều kiện của nút con khi biết tổ hợp giá trị của nút bố mẹ

Tính độc lập xác suất của mạng Bayes

- Mạng Bayes cho phép biểu diễn ngắn gọn toàn bộ các xác suất đồng thời
  - Việc rút gọn nhờ sử dụng tính độc lập xác suất trong mạng
- Độc lập xác suất
  - Mỗi nút V độc lập với tất cả các nút không phải là hậu duệ của V, nếu biết giá trị các nút bố mẹ của V
  - Ví dụ: H độc lập có điều kiện với L, O, B nếu biết D



$P(O) = 0.6$

$P(B) = 0.3$

$P(L|O) = 0.3$

$P(L|\neg O) = 0.6$

$P(D|O, B) = 0.05$

$P(D|O, \neg B) = 0.1$

$P(D|\neg O, B) = 0.1$

$P(D|\neg O, \neg B) = 0.2$

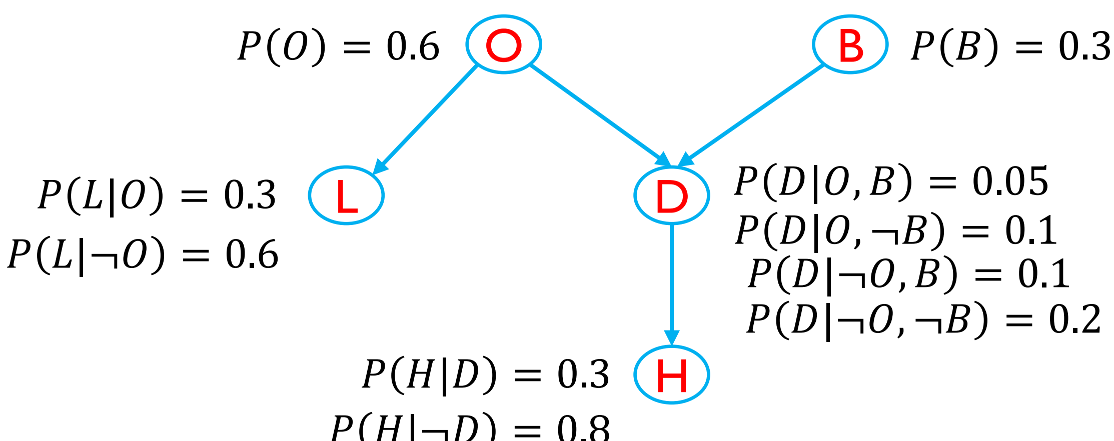
$P(H|D) = 0.3$

$P(H|\neg D) = 0.8$

Tính xác suất đồng thời cho mạng Bayes

- Ví dụ:

$$
\begin{aligned}
P(H, \neg L, D, \neg O, B) \\
&= P(H \mid \neg L, D, \neg O, B)\, P(\neg L, D, \neg O, B) \\
&= P(H|D)\, P(\neg L, D, \neg O, B) \\
&= P(H|D)\, P(\neg L \mid D, \neg O, B)\, P(D, \neg O, B) \\
&= P(H|D)\, P(\neg L|\neg O)\, P(D, \neg O, B) \\
&= P(H|D)\, P(\neg L|\neg O)\, P(D \mid \neg O, B)\, P(\neg O, B) \\
&= P(H|D)\, P(\neg L|\neg O)\, P(D|\neg O, B)\, P(\neg O)\, P(B) \\
&= (0.3)(1\ - 0.6)(0.1)(1\ - 0.6)(0.3)
\end{aligned}
$$



Tính xác suất đồng thời (tổng quát)

$$P(X_1 = x_1, \ldots, X_n = x_n) = \prod_{i=1}^{n} P(X_i = x_i \mid parents(X_i))$$

Xây dựng mạng Bayes

- Có 2 cách xây dựng
  - Xây dựng bằng tay (do người xây dựng)
    - Dựa trên hiểu biết của người về bài toán đang xét
    - Việc xây dựng mạng gồm 2 bước: xác định cấu trúc đồ thị và điền giá trị cho bảng xác suất điều kiện
  - Học máy từ dữ liệu: trong trường hợp có nhiều dữ liệu về tổ hợp giá trị các biến
    - Phân bố xác suất do mạng thể hiện phù hợp nhất với tần suất xuất hiện các giá trị trong tập dữ liệu

Xây dựng mạng Bayes (bằng tay)

1. Xác định tập các biến ngẫu nhiên liên quan
2. Chọn thứ tự cho các biến

   Ví dụ: $X_1, X_2, \dots, X_n$

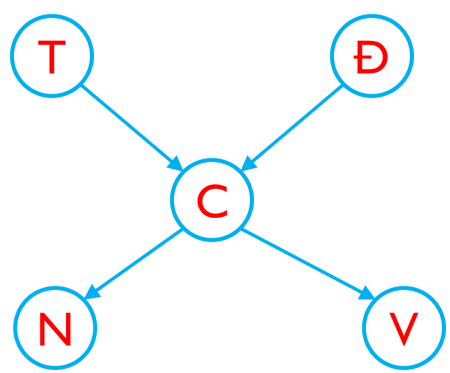
3. **for** $i = 1$ **to** n **do**
   a. thêm một nút cho X_i
   b. chọn $parents(X_i)$ là tập nhỏ nhất các nút đã có sao cho X_i độc lập có điều kiện với tất cả các nút trước đó nếu biết $parents(X_i)$
   c. thêm một cung có hướng từ mỗi nút $parents(X_i)$ tới X_i
   d. thêm các giá trị xác suất điều kiện $P(X_i|parents(X_i))$ hoặc $P(X_i)$ nếu $parents(X_i) = \emptyset$

Ví dụ (1/2)

- Một người vừa lắp hệ thống báo động chống trộm ở nhà
- Hệ thống sẽ phát tiếng động khi có trộm
- Tuy nhiên, hệ thống có thể báo động (sai) nếu có chấn động do động đất
- Trong trường hợp nghe thấy hệ thống báo động, hai người hàng xóm tên làm Nam và Việt sẽ gọi điện cho chủ nhà
- Do nhiều nguyên nhân khác nhau, Nam và Việt có thể thông báo sai, chẳng hạn do ồn nên không nghe thấy chuông báo động hoặc ngược lại, nhầm âm thanh khác là tiếng chuông

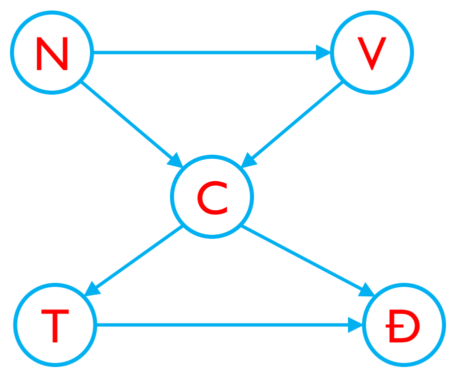
Ví dụ (2/2)

- **Bước 1**: lựa chọn biến: sử dụng 5 biến sau
  - T (có trộm), Đ (động đất), C (chuông báo động), N (Nam gọi điện), V (Việt gọi điện)
- **Bước 2**: các biến được sắp xếp theo thứ tự T, Đ, C, N, V
- **Bước 3**: thực hiện như các bước ở hình vẽ, ta xây dựng được mạng thể hiện trên hình sau (để đơn giản, trên hình vẽ chỉ thể hiện cấu trúc và không có bảng xác suất điều kiện)



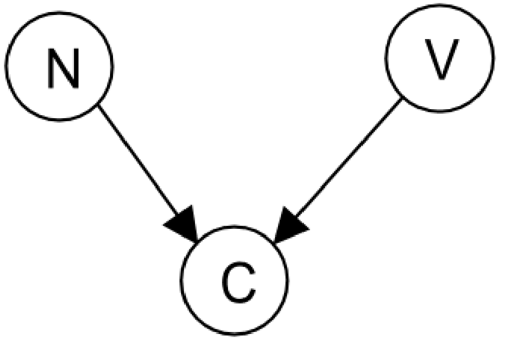
Ảnh hưởng của việc sắp xếp các nút

- Việc xây dựng mạng Bayes trong thực tế không đơn giản
  - Việc chọn thứ tự các nút đúng để từ đây chọn được tập nút cha có kích thước nhỏ là khó khăn
- Giả sử các biến được xếp theo thứ tự khác: $N, V, C, T, Đ$



Tính độc lập xác suất tổng quát: Khái niệm d-phân cách (1/5)

- Nếu không biết giá trị của nút C
  - Theo tính chất mạng Bayes N và V độc lập (không điều kiện)
- Nếu đã biết giá trị của nút C
  - N và V còn độc lập với nhau nữa hay không?



Các kiến thức đã học không cho phép trả lời câu hỏi này !!!

Tính độc lập xác suất tổng quát: Khái niệm d-phân cách (2/5)

- Khái niệm *d-phân cách* trả lời câu hỏi về tính độc lập của tập các nút X với tập nút Y khi biết tập nút E trên một mạng Bayes
  - Các nút X và các nút Y được gọi là bị *d-phân cách* bởi các nút E nếu X và Y là độc lập xác suất với nhau khi biết E
  - Các nút X và các nút Y là *d-kết nối* với nhau nếu chúng không bị *d-phân cách*

- Để xác định tính *d-phân cách* của tập X và Y, trước tiên ta cần xác định tính *d-phân cách* giữa hai nút đơn x thuộc X và y thuộc Y
  - Hai tập nút sẽ độc lập với nhau nếu mỗi nút trong tập này độc lập với tất cả các nút trong tập kia

Tính độc lập xác suất tổng quát: Khái niệm d-phân cách (3/5)

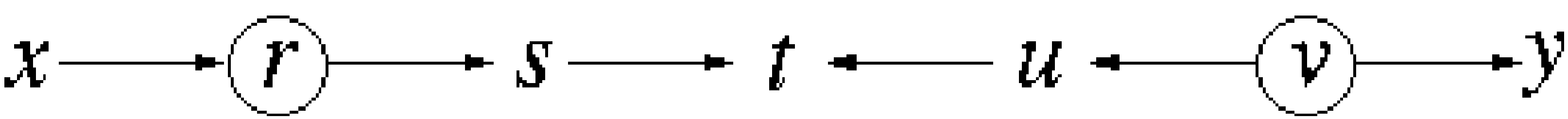
- **Quy tắc 1**: nút x và y là d*-kết nối* nếu tồn tại đường đi không bị *phong tỏa* giữa hai nút. Ngược lại, nếu không tồn tại đường đi như vậy thì x và y là d*-phân cách*.
  - Đường đi là một chuỗi các cung nằm liền nhau, không tính tới hướng của các cung đó
  - Đường đi không bị *phong tỏa* là đường đi mà trên đó không có hai cung liền kề hướng vào nhau
  - Nút có hai cung hướng vào như vậy gọi là nút *xung đột*

$$x \longrightarrow r \longrightarrow s \longrightarrow t \longleftarrow u \longleftarrow v \longrightarrow y$$

  - Tính *kết nối* và *phân cách* xác định theo **Quy tắc 1** là không điều kiện và do vậy tính độc lập xác suất được xác định theo **Quy tắc 1** là độc lập không điều kiện

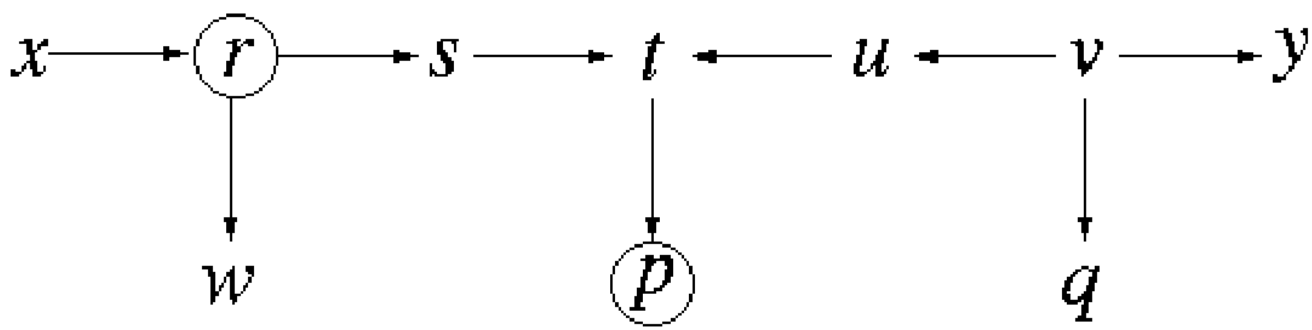
Tính độc lập xác suất tổng quát: Khái niệm d-phân cách (4/5)

- **Quy tắc 2**: nút x và y là *d-kết nối có điều kiện* khi biết tập nút E nếu tồn tại đường đi không bị phong tỏa (không chứa nút xung đột) và không đi qua bất cứ nút nào thuộc E. Ngược lại, nếu không tồn tại đường đi như vậy thì ta nói rằng x và y là *d-phân cách* bởi E. Nói cách khác, mọi đường đi giữa x và y (nếu có) đều bị E phong tỏa.
  - Khi biết giá trị một số nút (tập nút E), tính chất độc lập hay phụ thuộc giữa các nút còn lại có thể thay đổi
  - Tính độc lập hay phụ thuộc trong trường hợp này được gọi là *d-phân cách* có điều kiện theo tập biến E



Tính độc lập xác suất tổng quát: Khái niệm *d*-phân cách (5/5)

- **Quy tắc 3**: nếu một nút xung đột là thành viên của tập E, hoặc có hậu duệ thuộc tập E, thì nút đó không còn phong tỏa các đường đi qua nó nữa
  - Giả sử ta biết một sự kiện được gây ra bởi hai hay nhiều nguyên nhân, nếu ta đã biết một nguyên nhân là đúng thì xác suất những nguyên nhân còn lại giảm đi, nếu ta biết một nguyên nhân là sai thì xác suất những nguyên nhân còn lại tăng lên



s và y là *d-kết nối*, x và u là *d-phân cách*

Nội dung

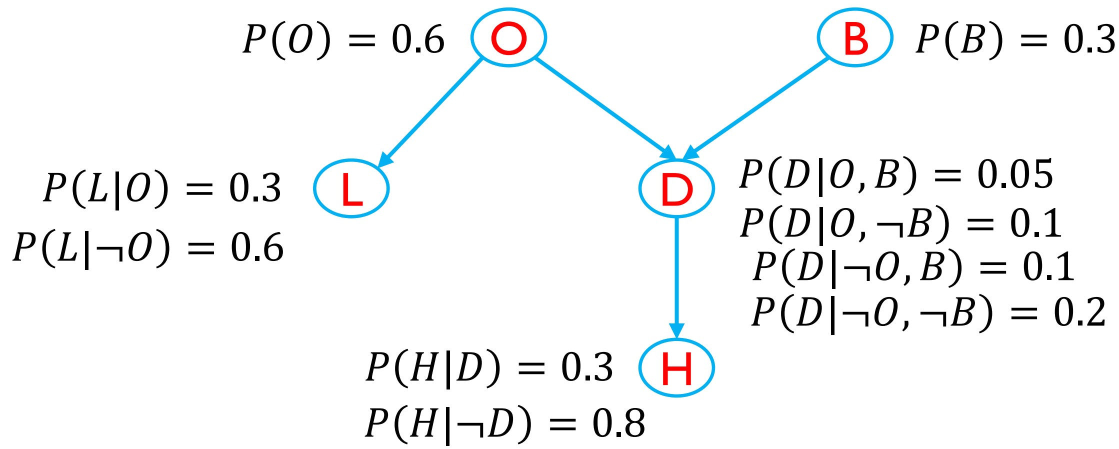
- Định nghĩa và cách xây dựng mạng Bayes
- Suy diễn với mạng Bayes

Những nội dung đã học (nhắc lại)

- Cách xây dựng mạng bayes (bằng tay)
- Mạng bayes cho phép rút gọn việc biểu diễn
  - Không cần lưu toàn bộ bảng xác suất đồng thời
- Có thể tính xác suất đồng thời mọi tổ hợp giá trị các biến
- Do vậy, có thể tính mọi xác suất hậu nghiệm cần cho suy diễn

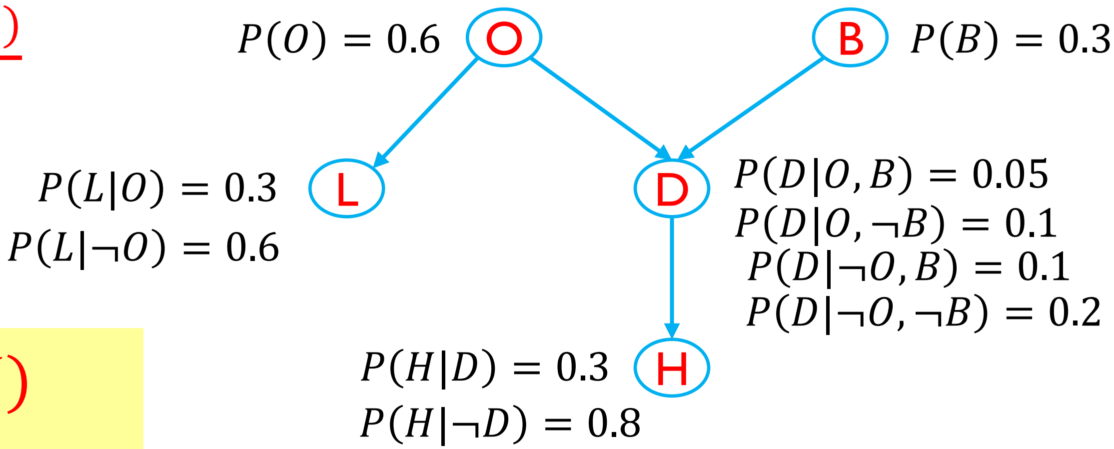
Ví dụ tính xác suất hậu nghiệm

- Cần tính $P(\,L \mid B, \neg H)$



Ví dụ tính xác suất hậu nghiệm

- $P(L|B,\neg H) = \frac{P(L,B,\neg H)}{P(B,\neg H)}$



Bước 1: tính $P(L,B,\neg H)$

Bước 2: tính $P(\neg L,B,\neg H)$

Bước 3: tính

$$\frac{P(L,B,\neg H)}{P(L,B,\neg H)+P(\neg L,B,\neg H)}$$

Xác suất đồng thời tính như trong bài trước

Trường hợp chung

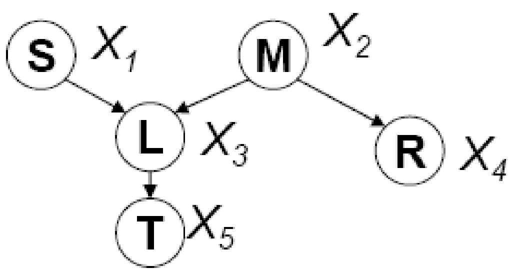
$$P(E_1|E_2) = \frac{P(E_1 \wedge E_2)}{P(E_2)} = \frac{Tổng\ xác\ suất\ đồng\ thời\ chứa\ E_1 và\ E_2}{Tổng\ xác\ suất\ đồng\ thời\ chứa\ E_2}$$

- Vấn đề
  - Đòi hỏi liệt kê các xác suất đồng thời có chứa E_1, E_2
  - Số lượng xác suất đồng thời như vậy tăng theo hàm mũ của số biến ⇒ không thực tế

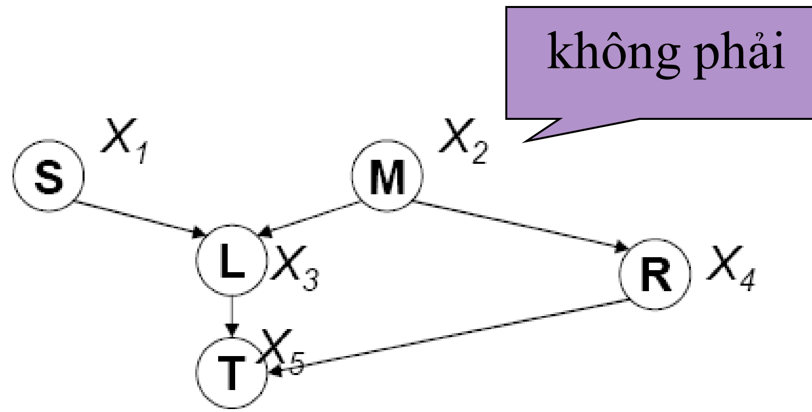
- **Suy diễn nói chung trên mạng bayes là bài toán NP đầy đủ ☹ ☹ ☹**

Suy diễn trên thực tế

- Suy diễn cho trường hợp riêng
  - Khi mạng có dạng liên kết đơn (poly tree): giữa 2 nút bất kỳ không có quá 1 đường đi



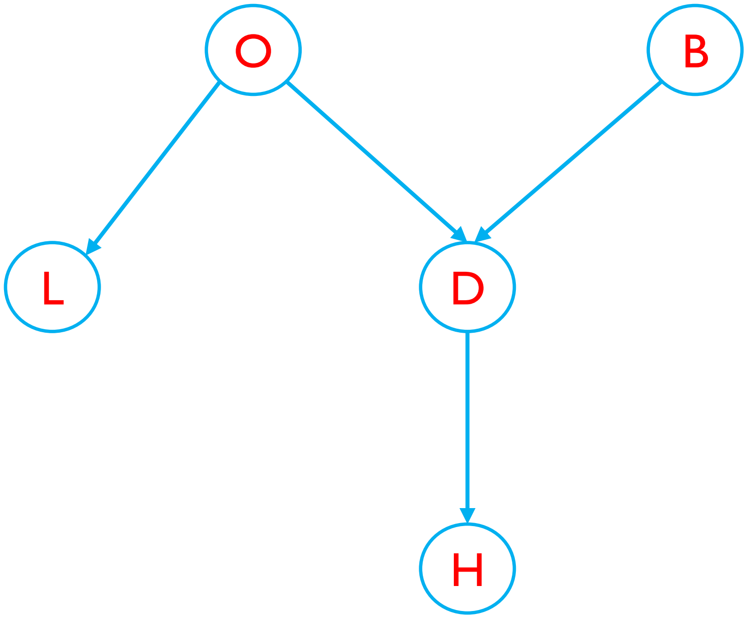
A poly tree



  - Tồn tại thuật toán với độ phức tạp tuyến tính cho poly tree
- Suy diễn xấp xỉ bằng cách lấy mẫu

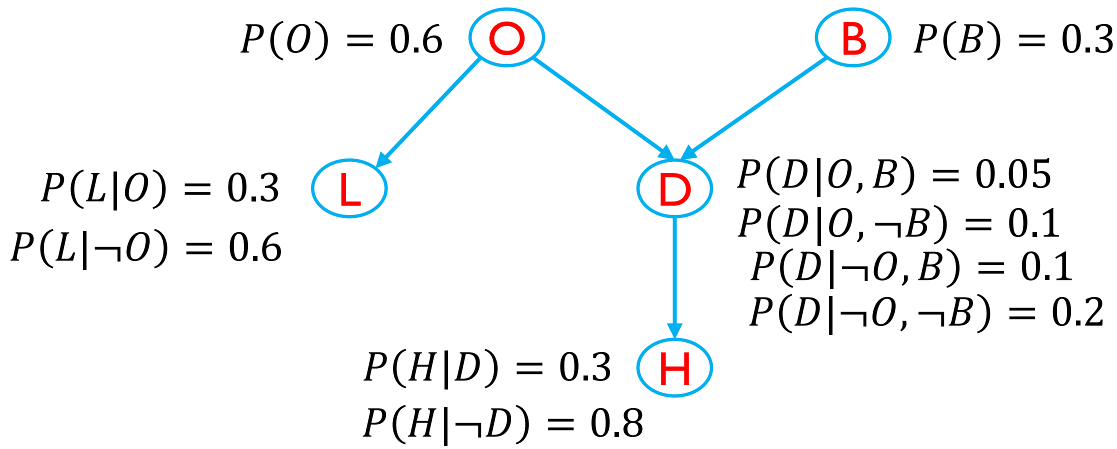
Suy diễn cho trường hợp riêng đơn giản

- Trường hợp đơn giản nhất
  - Khi bằng chứng E và kết quả Q có duy nhất một liên kết trực tiếp với nhau
  - Phân biệt 2 trường hợp:
    - Suy diễn nhân quả (trên xuống): cần tính $P(Q|E)$ khi E là nút cha của Q
    - Suy diễn chẩn đoán (dưới lên): cần tính $P(Q|E)$ khi E là nút con của Q



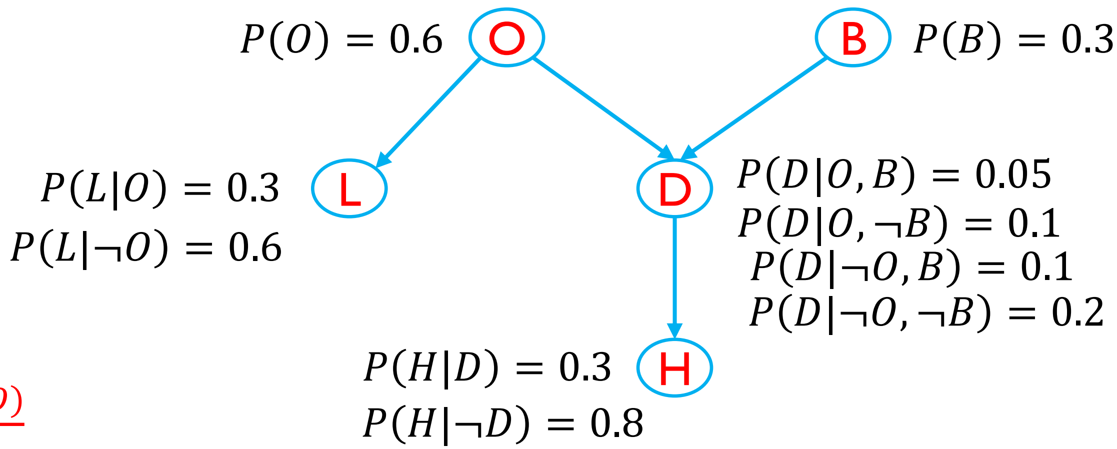
Suy diễn nhân quả (1/3)

- Ví dụ: tính $P(D|B)$



Suy diễn nhân quả (2/3)

- Ví dụ: tính $P(D|B)$



$$P(D|B) = \frac{P(D,B)}{P(B)}$$

$$= \frac{P(D,B,O) + P(D,B,\neg O)}{P(B)}$$

$$= \frac{P(D|B,O)P(B,O) + P(D|\,B,\neg O)P(B,\neg O)}{P(B)}$$

$$= \frac{P(D|B,O)P(B)P(O) + P(D|\,B,\neg O)P(B)P(\neg O)}{P(B)}$$

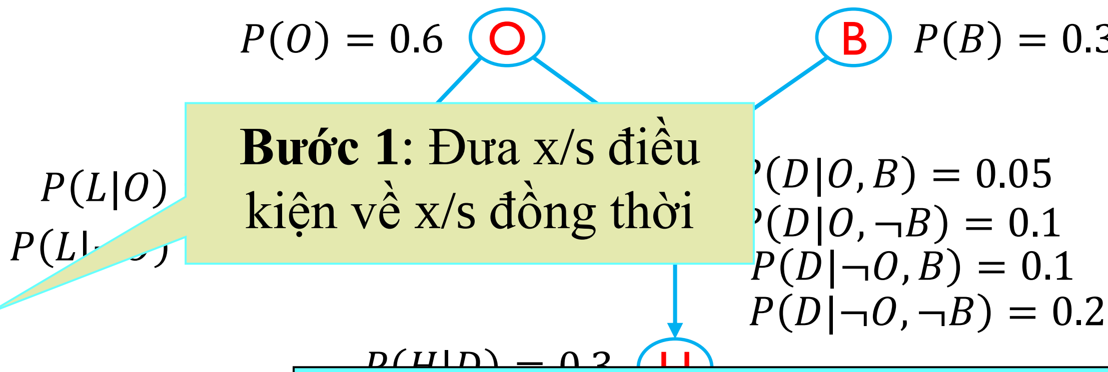
$$= P(D|B,O)P(O) + P(D|\,B,\neg O)P(\neg O)$$

$$= (0.05)(0.6) + (0.1)(1 - 0.6)$$

$$= 0.07$$

Suy diễn nhân quả (3/3)

- Ví dụ: tính $P(D|B)$



$$P(D|B) = \frac{P(D,B)}{P(B)}$$

$$= \frac{P(D,B,O) + P(D,B,\neg O)}{P(B)}$$

Bước 1: Đưa x/s điều kiện về x/s đồng thời

$$= \frac{P(D|B,O)P(B,O) + P(D|\, B,\neg O)P(B,\neg O)}{P(B)}$$

$$= \frac{P(D|B,O)P(B)P(O) + P(D|\, B,\neg O)P(B)P(\neg O)}{P(B)}$$

$$= P(D|B,O)P(O) + P(D|\, B,\neg O)P(\neg O)$$

Bước 2: Sử dụng tính độc lập x/s trên mạng Bayes, viết lại các x/s đồng thời dưới dạng x/s điều kiện của nút con khi biết các giá trị bố mẹ

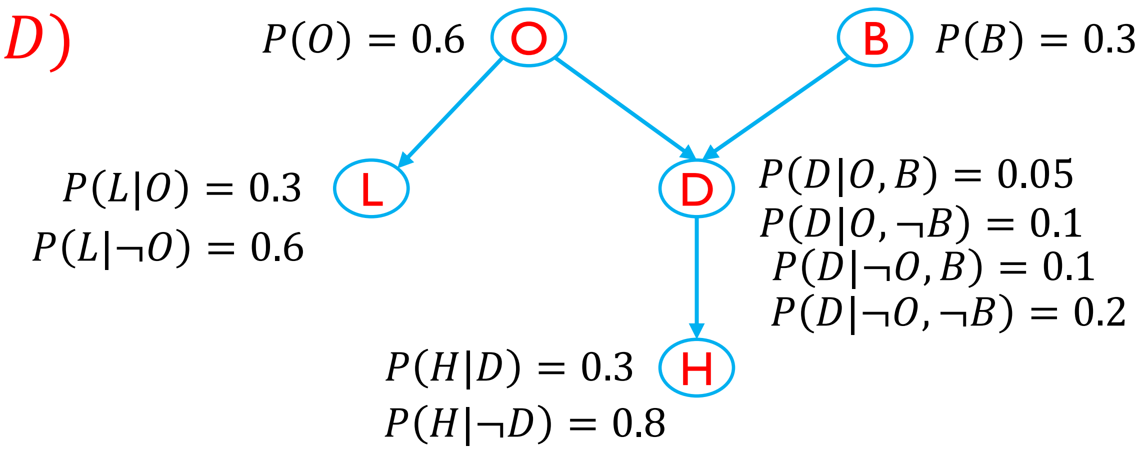
$$= (0.05)(0.6) + (0.1)(1 - 0.6)$$

$$= 0.07$$

Bước 3: Sử dụng các giá trị x/s từ bảng x/s điều kiện để tính

Suy diễn chuẩn đoán (1/5)

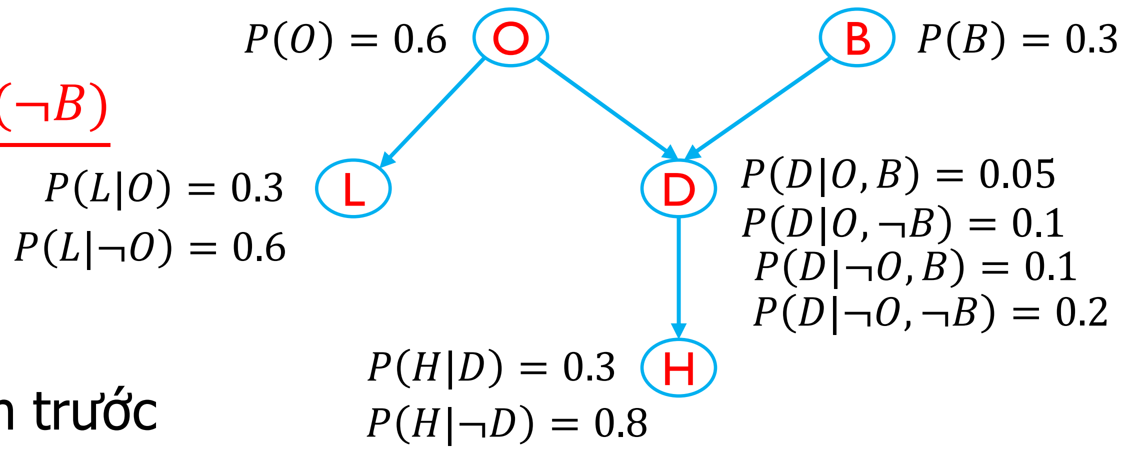
- Ví dụ: tính $P(\neg B|\neg D)$



Suy diễn chuẩn đoán (2/5)

- Theo Bayes

$$P(\neg B|\neg D) = \frac{P(\neg D|\neg B)P(\neg B)}{P(\neg D)}$$

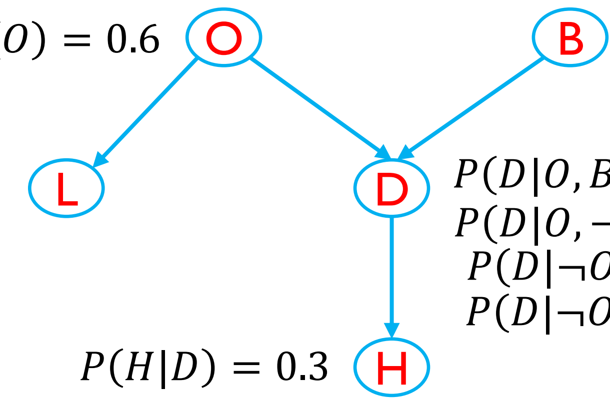


tính $P(\neg D|\neg B)$ như phần trước

Suy diễn chuẩn đoán (3/5)

- Theo Bayes

$$P(\neg B|\neg D) = \frac{P(\neg D|\neg B)P(\neg B)}{P(\neg D)}$$



$P(O) = 0.6$

$P(B) = 0.3$

$P(L|O) = 0.3$

$P(L|\neg O) = 0.6$

$P(D|O,B) = 0.05$

$P(D|O,\neg B) = 0.1$

$P(D|\neg O,B) = 0.1$

$P(D|\neg O,\neg B) = 0.2$

$P(H|D) = 0.3$

$P(H|\neg D) = 0.8$

tính $P(\neg D|\neg B)$ như phần trước

$$\begin{aligned} P(\neg D|\neg B) &= P(\neg D|O,\neg B)P(O) + P(\neg D|\neg O,\neg B)P(\neg O) \\ &= (0.9)(0.6) + (0.8)(0.4) \\ &= 0.86 \end{aligned}$$

$$P(\neg B|\neg D) = \frac{(0.86)(0.7)}{P(\neg D)} = \frac{0.602}{P(\neg D)}$$

Để tính $P(\neg D)$, ta sẽ tính $P(B|\neg D)$

Suy diễn chuẩn đoán (4/5)

- $P(B|\neg D) = \frac{P(\neg D|B)P(B)}{P(\neg D)} = \frac{(1-0.07)0.3}{P(\neg D)} = \frac{0.279}{P(\neg D)}$
- Sử dụng

$P(\neg B|\neg D) + P(B|\neg D) = 1$

$$\frac{0.602}{P(\neg D)} + \frac{0.279}{P(\neg D)} = 1$$

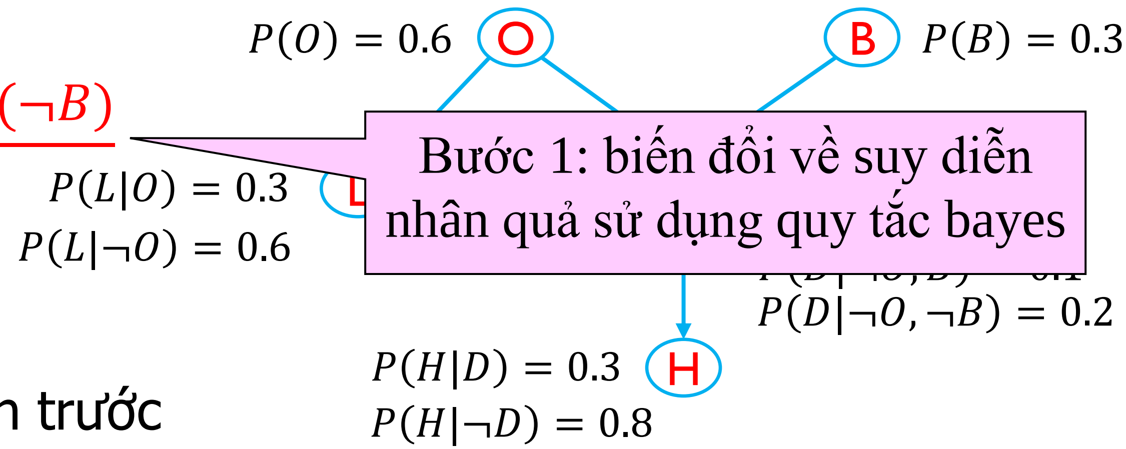
Do đó $P(\neg D) = 0.881$

Thay lại:

$$P(\neg B|\neg D) = \frac{0.602}{P(\neg D)} = \frac{0.602}{0.881} = 0.683$$

Suy diễn chuẩn đoán (5/5)

- Theo Bayes



$$P(\neg B|\neg D) = \frac{P(\neg D|\neg B)P(\neg B)}{P(\neg D)}$$

Bước 1: biến đổi về suy diễn nhân quả sử dụng quy tắc bayes

tính $P(\neg D|\neg B)$ như phần trước

$$\begin{aligned} P(\neg D|\neg B) &= P(\neg D|O, \neg B)P(O) + P(\neg D|\neg O, \neg B)P(\neg O) \\ &= (0.9)(0.6) + (0.8)(0.4) \\ &= 0.86 \end{aligned}$$

Bước 2: thực hiện giống suy diễn nhân quả

$$P(\neg B|\neg D) = \frac{(0.86)(0.7)}{P(\neg D)} = \frac{0.602}{P(\neg D)}$$

Để tính $P(\neg D)$, ta sẽ tính $P(B|\neg D)$

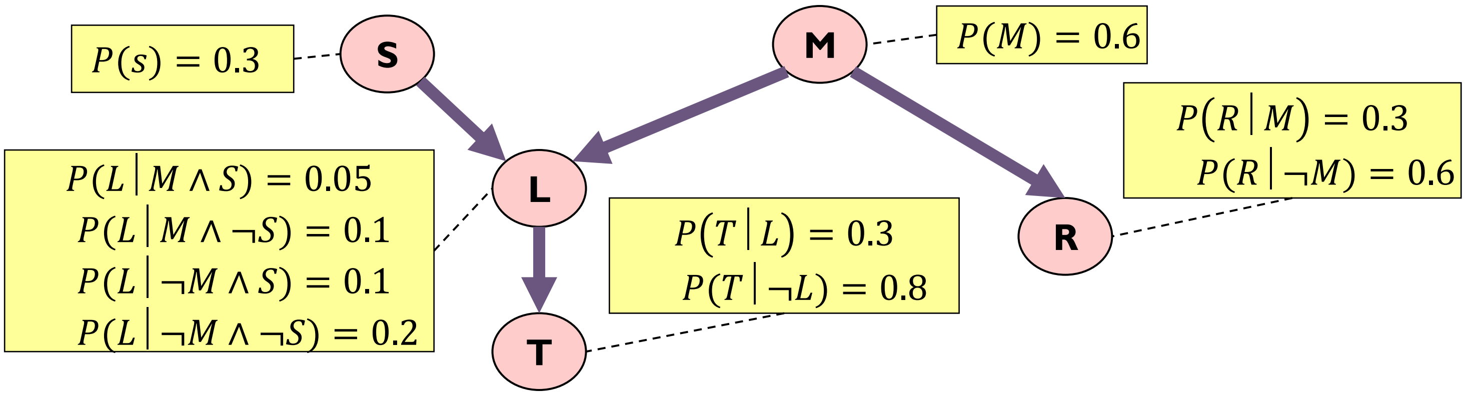
Phương pháp chung

- Áp dụng cho cả suy diễn nhân quả và suy diễn chuẩn đoán
  - **Bước 1**: Đưa xác suất điều kiện về các xác suất đồng thời
  - **Bước 2**: Sử dụng tính độc lập xác suất trên mạng Bayes, viết lại các xác suất đồng thời dưới dạng các xác suất điều kiện của nút con khi biết các giá trị bố mẹ
  - **Bước 3**: Sử dụng các giá trị xác suất từ bảng xác suất điều kiện để tính

Suy diễn bằng cách lấy mẫu

- Trong trường hợp tổng quát: suy diễn trên mạng Bayes là NP-đầy đủ (rất phức tạp)
- Có thể suy diễn xấp xỉ bằng cách lấy mẫu
- Sinh ra các bộ giá trị của biến có cùng xác suất đồng thời của mạng

Lấy mẫu (1/2)



- Chọn ngẫu nhiên S: $S = true$ với xác suất 0.3
- Chọn ngẫu nhiên M: $M = true$ với xác suất 0.6
- Chọn ngẫu nhiên L: xác suất $L = true$ phụ thuộc vào giá trị của S, M ở trên
  - Giả sử bước trên sinh $M = true, S = false$, $L = true$ với xác suất 0.1
- Chọn ngẫu nhiên R với xác suất phụ thuộc giá trị của M
- Chọn ngẫu nhiên T với xác suất phụ thuộc giá trị của L

Lấy mẫu (2/2)

- Giả sử cần tính: $P(R = True \,|\, T = True, S = False)$
- Lấy mẫu nhiều lần theo cách ở trên, mỗi bộ giá trị sinh ra được gọi là một mẫu
- Tính số lần xẩy ra những sự kiện sau:
  - Nc: số mẫu có $T = True$ và $S = False$
  - Ns: số mẫu có $R = True$, $T = True$ và $S = False$
  - N: tổng số mẫu
- Nếu N đủ lớn:
  - Nc/N: (xấp xỉ) xác suất $P(T = True\ and\ S = False)$
  - Ns/N: (xấp xỉ) xác suất $P(R = True, T = True, S = False)$
  - $P(R|T, \neg S) = P(R, T, \neg S)/P(T, \neg S) \approx Ns/Nc$

Lấy mẫu tổng quát

- Cần tính $P(E_1|E_2)$
- Lấy mẫu số lượng đủ lớn
- Tính số lượng:
  - Nc: số mẫu có E_2
  - Ns: số mẫu có E_1 và E_2
  - N: Tổng số mẫu
- Nếu N đủ lớn, ta có: $P(E_1|E_2) = \frac{Ns}{Nc}$

Bài tập

- Làm một số bài tập trong giáo trình