NHẬP MÔN TRÍ TUỆ NHÂN TẠO

SUY DIỄN XÁC SUẤT

ThS. Vũ Hoài Thư

Nội dung

1. Vấn đề suy diễn xác suất trong điều kiện không rõ ràng
2. Nguyên tắc suy diễn xác suất
3. Một số khái niệm về xác suất
4. Mạng Bayes
5. Suy diễn với mạng Bayes

Vấn đề suy diễn xác suất trong điều kiện không rõ ràng

Vấn đề suy diễn trong điều kiện không rõ ràng

- Logic
  - Cho phép biểu diễn tri thức và suy diễn
  - Đòi hỏi tri thức rõ ràng, đầy đủ, chắc chắn, không mâu thuẫn
- Thế giới thực
  - Luôn có yếu tố không rõ ràng, thiếu thông tin, có mâu thuẫn

Vấn đề suy diễn trong điều kiện không rõ ràng

Các yếu tố ảnh hưởng tới tính rõ ràng, chắc chắn của tri thức, thông tin

- Thông tin có chứa đựng yếu tố ngẫu nhiên: Khi chơi bài, tung đồng xu
- Lý thuyết không rõ ràng: Ví dụ không biết hết cơ chế chế gây bệnh
- Thiếu thông tin thực tế: Không đủ thông tin xét nghiệm của bệnh nhân
- Các yếu tố liên quan tới bài toán quá lớn, quá phức tạp: Không thể biểu diễn được mọi yếu tố
- Sai số khi lấy thông tin từ môi trường: Các thiết bị đo có sai số

Các cách tiếp cận

- Logic đa trị
  - Cho phép sử dụng nhiều giá trị hơn, ngoài "đúng" và "sai"
- Logic mờ
  - Biểu thức có thể nhận giá trị "đúng" với một giá trị trong khoảng [0,1]
- Lý thuyết khả năng
  - Các sự kiện hay công thức được gán một số thể hiện khả năng xảy ra sự kiện
- Suy diễn xác suất
  - Kết quả suy diễn trả về xác suất một sự kiện hay công thức nào đó đúng.

Nguyên tắc suy diễn xác suất

Nguyên tắc suy diễn xác suất (1/2)

- Thay vì suy diễn về tính "đúng" hoặc "sai" của mệnh đề (2 giá trị), suy diễn về "niềm tin" mệnh đề đó đúng hay sai (vô số giá trị)
    - Gắn cho mỗi mệnh đề một số đo giá trị niềm tin
    - Biểu diễn mức đo niềm tin như giá trị xác suất, sử dụng lý thuyết xác suất để làm việc với giá trị này
    - Với mệnh đề A:
        - Gán xác suất $P(A) : 0 \leq P(A) \leq 1$
        - $P(A) = 1$ nếu A đúng, $P(A) = 0$ nếu A sai.
    - Ví dụ:
        - $P(\text{Cảm=true}) = 0.6$: người bệnh bị cảm với xác suất 60%, "Cảm" là một biến ngẫu nhiên có thể nhận 1 trong 2 giá trị {True, False}
        - $P(\text{trời=nắng} \wedge \text{gió=mạnh}) = 0.8$: ta tin rằng trời nắng và gió mạnh với xác suất 80%, trời là biến ngẫu nhiên nhận các giá trị {nắng, mưa, u ám}, gió là biến ngẫu nhiên nhận giá trị {mạnh, yếu, trung bình}

Nguyên tắc suy diễn xác suất (2/2)

- Bản chất của xác suất sử dụng trong suy diễn
  - Bản chất thống kê: dựa trên thực nghiệm và quan sát
    - Không phải khi nào cũng xác định được
  - Xác suất dựa trên chủ quan: mức độ tin tưởng, niềm tin là sự kiện đó đúng hoặc sai của chúng chuyên gia, người dùng
    - Được sử dụng khi suy diễn xác suất
- Thu thập thông tin
  - Xác định các tham số liên quan tới bài toán: ví dụ “màu”, “đẹp”
  - Mỗi tham số là một biến ngẫu nhiên
  - Mỗi biến ngẫu nhiên có thể nhận một số giá trị rời rạc trong miền giá trị của biến đó Có thể là{True, False} hoặc nhiều giá trị hơn: {đỏ, xanh, vàng}
  - Ví dụ: $P(\text{màu}{=}\text{đỏ}) = 0.09; P(\neg\text{đẹp}) = 0.2$

Một số khái niệm về xác suất

Các tiên đề xác suất và một số tính chất cơ bản

Các tiên đề xác suất

1. $0 \leq P(A = a) \leq 1$, với mọi a thuộc miền giá trị của A
2. $P(True) = 1, P(False) = 0$
3. $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$

Một số tính chất

1. $P(\neg A) = 1 - P(A)$
2. $P(A) = P(A \wedge B) + P(A \wedge \neg B)$
3. $\sum_a P(A = a) = 1$: tổng lấy theo các giá trị a thuộc miền giá trị của A

- Ví dụ: A là biến màu nhận một trong 2 giá trị {đỏ, xanh}. Vậy $P(\text{A} = \text{đỏ}) + P(\text{A} = \text{xanh}) = 1$

Xác suất đồng thời (1/2)

- Có dạng: $P(V_1 = v_1, V_2 = v_2, \ldots V_n = v_n)$
- Phân bố xác suất đồng thời đầy đủ: bao gồm xác suất cho tất cả các tổ hợp giá trị của tất cả biến ngẫu nhiên
- Ví dụ: cho 3 biến Bool: Chim, Non, Bay

Chim (C)	Non (N)	Bay (B)	P
T	T	T	0.0
T	T	F	0.2
T	F	T	0.04
T	F	F	0.01
F	T	T	0.01
F	T	F	0.01
F	F	T	0.23
F	F	F	0.5

Xác suất đồng thời (2/2)

- Nếu có tất cả xác suất đồng thời, ta có thể tính xác suất cho mọi mệnh đề liên quan tới bài toán đang xét
- Ví dụ:
    - Xác suất một vật nào đó là "chim":
      $P(\text{Chim} = \text{T}) = P(C) = 0.0 + 0.2 + 0.04 + 0.01 = 0.25$
    - Xác suất "chim không biết bay":
      $P(\text{Chim} = \text{T}, \text{Bay} = \text{F}) = P(C, \neg B) =$
      $P(C, N, \neg B) + P(C, \neg N, \neg B) = 0.2 + 0.01 = 0.21$
- Tính xác suất một con chim là chim già?

Xác suất điều kiện (1/3)

- Đóng vai trò quan trọng trong suy diễn
  - Từ bằng chứng suy ra xác suất của kết quả
  - Ví dụ:
    - $P(A|B) = 1$ tương đương $B \Rightarrow A$ trong logic
    - $P(A|B) = 0.9$: tương đương $B \Rightarrow A$ với xác suất hay độ chắc chắn là 90%
    - Với nhiều bằng chứng (quan sát) $E_1, \ldots, E_n$ có thể tính $P(Q|E_1, \ldots, E_n)$ tương đương: niềm tin Q đúng là bao nhiêu nếu biết $E_1, \ldots, E_n$ và không biết gì thêm
- Định nghĩa xác suất điều kiện

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

Xác suất điều kiện (2/3)

- Ví dụ: cho 3 biến Bool: Chim, Non, Bay

Chim (C)	Non (N)	Bay (B)	P
T	T	T	0.0
T	T	F	0.2
T	F	T	0.04
T	F	F	0.01
F	T	T	0.01
F	T	F	0.01
F	F	T	0.23
F	F	F	0.5

- Tính $P(\neg\text{Chim} \mid \text{Bay})$?

Xác suất điều kiện (3/3)

Các tính chất của xác suất điều kiện

- $P(A, B) = P(A|B)P(B)$
- Quy tắc chuỗi:
  $P(A, B, C, D) = P(A|B, C, D)P(B|C, D)P(C|D)P(D)$
- Quy tắc chuỗi có điều kiện: $P(A, B|C) = P(A|B, C)P(B|C)$
- Quy tắc Bayes: $P(A|B) = \dfrac{P(A)P(B|A)}{P(B)}$
- Bayes có điều kiện: $P(A|B) = \dfrac{P(B|A, C)P(A|C)}{P(B|C)}$
- $P(A) = \sum_b \{P(A|B = b)P(B = b)\}$, tổng lấy theo tất cả giá trị b của biến ngẫu nhiên B.
- $P(\neg B|A) = 1 - P(B|A)$

Kết hợp nhiều bằng chứng

- Trường hợp tổng quát: cho bảng xác suất đồng thời, có thể tính:
  - $P(V_1 = v_1, \ldots, V_k = v_k | V_{k+1} = v_{k+1}, \ldots, V_n = v_n)$
  - Tổng các dòng có $V_1 = v_1, \ldots, V_n = v_n$ chia cho tổng các dòng có $V_{k+1} = v_{k+1}, \ldots, V_n = v_n$
- Ví dụ: Tính $P(\neg \mathrm{Chim} | \mathrm{Bay}, \neg \mathrm{Non})$

Tính độc lập xác suất

- A độc lập với B nếu $P(A|B) = P(A)$
  - Ý nghĩa: biết giá trị của B không thêm thông tin về A
  - Từ đây suy ra $P(A, B) = P(A)P(B)$
- A độc lập có điều kiện với B khi biết C nếu:
  - $P(A|B, C) = P(A|C)$ hoặc $P(B|A, C) = P(B|C)$
  - Ý nghĩa: nếu đã biết giá trị của C thì việc biết giá trị của B không cho ta thêm thông tin về A
  - Suy ra: $P(A, B|C) = P(A|C)P(B|C)$

Sử dụng quy tắc Bayes

- Quy tắc Bayes đóng vai trò quan trọng trong suy diễn
- Quy tắc Bayes:

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}$$

- Để suy diễn cần biết $P(A|B)$ nhưng thường $P(B|A)$ dễ tính hơn.
- Ví dụ: xác suất bị cúm khi đau đầu và xác suất đau đầu khi bị cúm

Ví dụ (1/2)

- Một người có kết quả xét nghiệm dương tính với bệnh B
- Thiết bị xét nghiệm không chính xác hoàn toàn
  - Thiết bị cho kết quả dương tính đối với 98% người có bệnh
  - Thiết bị cho kết quả dương tính đối với 3% người không có bệnh
- 0.8% dân số mắc bệnh này
- Hỏi: Người này có bị bệnh không?

Ví dụ (2/2)

- Ký hiệu hai sự kiện: B - có bệnh, A - xét nghiệm dương tính
- Theo dữ kiện bài toán ta có
  - $P(B) = 0.008, P(\neg B) = -0.008 = 0.992$
  - $P(A|B) = 0.98, P(\neg A|B) = 1 - P(A|B) = 0.02$
  - $P(A|\neg B) = 0.03, P(\neg A|\neg B) = 0.97$
- Cần so sánh các xác suất: $P(B|A)$ và $P(A|B)$
- Sử dụng quy tắc Bayes
  - $P(B|A) = \dfrac{P(A|B)P(B)}{P(A)} = \dfrac{0.98 * 0.008}{P(A)} = \dfrac{0.00784}{P(A)}$
  - $P(\neg B|A) = \dfrac{P(A|\neg B)P(\neg B)}{P(A)} = \dfrac{0.03 * 0.992}{P(A)} = \dfrac{0.02976}{P(A)}$
  - Có $P(\neg B|A) > P(B|A)$ => Không bị bệnh

Chuẩn tắc hoá

- Để so sánh P(B|A) và $P(\neg B|A)$ ta không cần tính cụ thể hai giá trị xác suất này, thay vào đó tính $\dfrac{P(B|A)}{P(\neg B|A)}$
  - Hai biểu thức có chung mẫu số $P(A)$
  - Kết luận có bệnh hay không phụ thuộc vào giá trị $\dfrac{P(B|A)}{P(\neg B|A)}$ lớn hơn hay nhỏ hơn 1
- Khi cần tính cụ thể xác suất này: dựa vào biểu thức $P(B|A) + P(\neg B|A) = 1$

Kết hợp quy tắc Bayes và tính độc lập xác suất

- Cần tính $P(A|B,C)$, biết B và C độc lập xác suất khi biết A
  - Theo quy tắc Bayes $P(A|B,C) = \dfrac{P(B,C|A)P(A)}{P(B,C)}$
  - Theo tính độc lập xác suất: $P(B,C|A) = P(B|A)P(C|A)$
  - Do đó: $P(A|B,C) = \dfrac{P(B|A)P(C|A)P(A)}{P(B,C)}$

Ví dụ

- Cho 3 biến nhị phân: gan BG, vàng da VD thiếu máu TM
- Giả sử VD độc lập với TM
- Biết $P(BG) = 10^{-7}$
- Có người đi khám bị VD
- Biết $P(VD) = 2^{-10}$ và $P(VD|BG) = 2^{-3}$
- Yêu cầu:
  a) Xác suất người khám bị bệnh gan là bao nhiêu?
  b) Cho biết thêm người đó bị thiếu máu và $P(TM) = 2^{-6}, P(TM|BG) = 2^{-1}$. Hãy tính xác suất người đó bị bệnh gan?

Mạng Bayes

Vấn đề biểu diễn xác suất

- Bài toán suy diễn
  - Cho bằng chứng $E_1, E_2, \ldots, E_n$
  - Cần xác định yêu cầu Q bằng cách tính $P(Q|E_1, E_2, \ldots, E_n)$
- 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 trong thực tế

=> Cần có cách biểu diễn và suy diễn hợp lý hơn

Ví dụ (1/2)

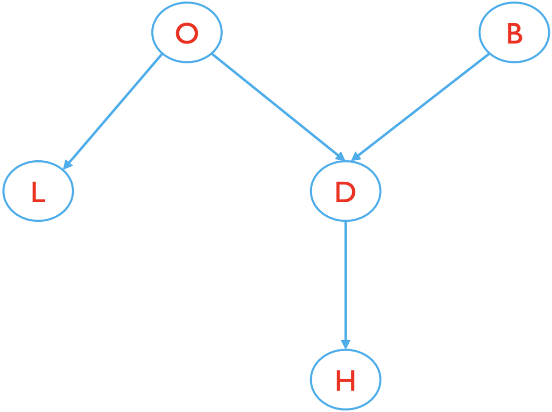
- Bài toán: Một người đi làm về, cần đoán trong nhà có người không?
- Biết rằng:
  - Nếu người nhà đi vắng thì thường (nhưng không luôn luôn) bật đèn ngoài sân
  - Khi không có người ở nhà thì thường buộc chó ở bên ngoài
  - Nếu chó bị ốm cũng bị buộc ở bên ngoài
  - Nếu chó ở ngoài thì có thể nghe tiếng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

Ví dụ (2/2)

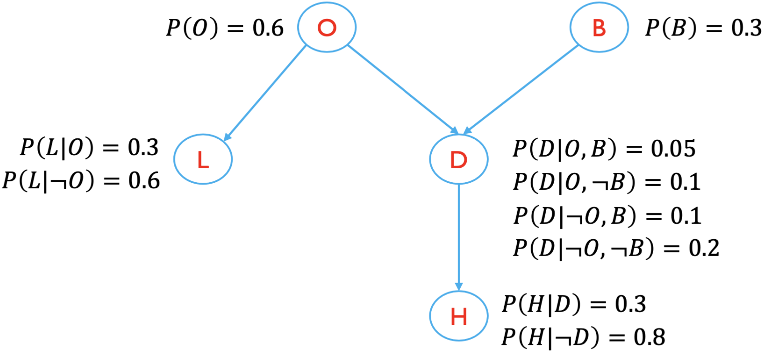
Xác định 5 biến ngẫu nhiên sau:

- 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

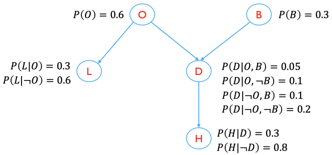

$P(O) = 0.6$
O
B
$P(B) = 0.3$
$P(L|O) = 0.3$
$P(L|\neg O) = 0.6$
L
D
$P(D|O,B) = 0.05$
$P(D|O,\neg B) = 0.1$
$P(D|\neg O,B) = 0.1$
$P(D|\neg O,\neg B) = 0.2$
H
$P(H|D) = 0.3$
$P(H|\neg D) = 0.8$

Định nghĩa mạng Bayes

- Mạng Bayes 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

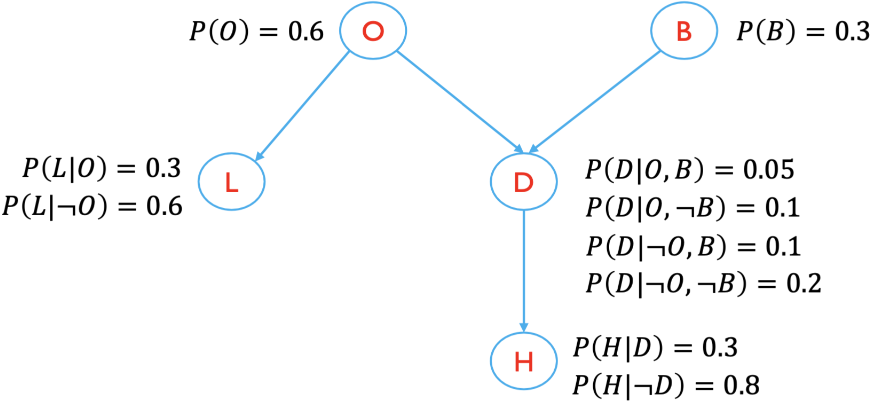
Tính độc lập xác suất của mạng Bayes

- 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 cha của V
  - Mỗi nút độc lập có điều kiện với tất cả các nút khác trên mạng nếu biết giá trị tất cả nút cha, nút con và nút cha của các nút con.
  - Ví dụ: H độc lập có điều kiện với L, O, B nếu biết D



Tính xác suất đồng thời cho mạng Bayes (1/2)

Ví dụ: Tính $P(H, \neg L, D, \neg O, B)$



Tính xác suất đồng thời cho mạng Bayes (2/2)

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

Tính xác suất đồng thời (tổng quát)

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

Hay:

$$P(x_1, \ldots, x_n) = \Pi_{i=1}^{n} P(X_i = x_i |\ parents(X_i))$$

Trong đó: $parents(X_i)$ là giá trị cụ thể các nút cha mẹ của nút 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: $X_1, X_2, \ldots, X_n$
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 về xây dựng mạng Bayes (1/4)

Giả sử có 3 biến ngẫu nhiên sau:

- B: Một người nào đó là Bill Gate
- T: Một người nào đó là tỷ phú
- H: Một người là nhà hảo tâm

Biết các giả thiết sau:

- Thế giới có 8 tỷ người
- Trên thế giới có 1000 người là tỷ phú
- Có 10% dân số trên thế giới là người hảo tâm
- Cứ 1000 tỷ phú thì có 700 người làm từ thiện
- Bill Gate là tỷ phú và cũng là nhà hảo tâm

Yêu cầu: Hãy xây dựng mạng Bayes và bảng xác suất có điều kiện theo thứ tự các biến là B, T, H

Ví dụ về xây dựng mạng Bayes (2/4)

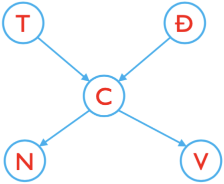
- Một người vừa lắp hệ thống báo động chống trộm ở nhà
- Hệ thống sẽ phát hiện tiếng động khi có trộm
- Tuy nhiên hệ thống có thể báo động sai nếu có chấn động do động đất
- Trong trường hợp nghe thấy hệ thống báo động, hai người hàng xóm tên là Nam và Việt sẽ gọi điện cho chủ nhà
- Do nhiều nguyên nhân 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 về xây dựng mạng Bayes (3/4)

- **Bước 1: Xác định các biến**
  Sử dụng 5 biến ngẫu nhiên: 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**

Ví dụ về xây dựng mạng Bayes (4/4)

- **Bước 3: Xây dựng cấu trúc mạng Bayes:**
  - Thêm nút **T**: Không có nút cha
  - Thêm nút **Đ**: Không có nút cha
  - Thêm nút **C**: Nếu có trộm và động đất => khả năng chuông kêu cao hơn => Cần thêm cả **T, Đ** vào tập nút cha của C
  - Thêm nút **N**: Nếu chuông kêu, khả năng Nam gọi điện tăng => Thêm **C** là nút cha của **N**
  - Thêm nút **V**: Nếu chuông kêu, khả năng Việt gọi điện tăng => Thêm **C** là nút cha của **V**



Ảnh hưởng của việc sắp xếp nút (1/2)

- 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 thức nhỏ là khó khăn
- Giả sử các biến được sắp xếp theo thứ tự khác: **N, V, C, T, Đ**

Ảnh hưởng của việc sắp xếp nút (2/2)

Xây dựng mạng theo thứ tự **N, V, C, T, Đ**

- Thêm nút N: không có nút cha
- Thêm nút V: Nếu Nam gọi điện, xs Việt gọi điện sẽ tăng lên => N có ảnh hưởng tới V. Thêm N vào tập cha của V
- Thêm C: Nếu N và V cùng gọi => khả năng chuông kêu cao hơn => cần thêm cả N và V vào tập nút cha của C.
- Thêm T: Nếu như biết trạng thái của chuông thì không cần quan tâm N và V => T chỉ có cha là C.
- Thêm Đ: Nếu có chuông, khả năng động đất sẽ tăng. Có trộm => nguyên nhân chuông kêu, khả năng động đất giảm => C và T đều là cha của Đ.

Bài tập 1

Cho ba triệu chứng sau: sốt (ký hiệu bằng biến ngẫu nhiên S), ho (ký hiệu H), và đau đầu (ký hiệu Đ) có thể có cùng một nguyên nhân là do một loại bệnh (ký hiệu B) gây ra. Xác suất ho khi bị bệnh và không bị bệnh lần lượt là 0.8 và 0.1; xác suất sốt khi bị bệnh và không bị bệnh là 0.9 và 0.05; xác suất đau đầu khi bị bệnh và không bị bệnh là 0.8 và 0.25. Số lượng bệnh nhân bị bệnh này chiếm 10% dân số.

Vẽ mạng Bayes và bảng xác suất điều kiện cho ví dụ này

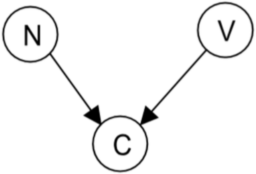
Bài tập 2

Cho biết 60% dân số thường xuyên ăn rau sống. Có 40% người ăn rau sống nhiễm bệnh đường ruột và 10% người không ăn rau sống bị bệnh. Nếu bị bệnh thì bị đau bụng với xác suất 80%, và bị chán ăn với xác suất 70%. Nếu không bị bệnh cũng có thể đau bụng với xác suất 10% và chán ăn với xác suất 20%. Ký hiệu ăn rau sống, nhiễm bệnh, đau bụng, chán ăn tương ứng bằng các biến ngẫu nhiên R, B, Đ, C.

Vẽ mạng Bayes và bảng xác suất điều kiện cho ví dụ này

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

Gỉa sử cho mạng sau:



- Nếu không biết giá trị của nút C: Theo tính chất của 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

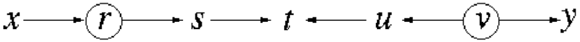
$$x \longrightarrow r \longrightarrow s \longrightarrow t \longleftarrow u \longleftarrow v \longrightarrow y$$

- 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.

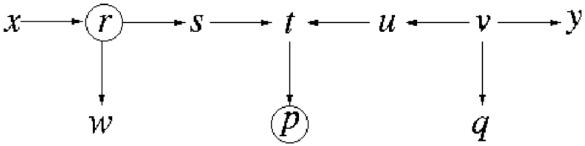
- Khi biết giá trị một số nút (tập nút E), tính chất độc lập hay phụ thuộc giữa các nút còn lại có thể thay đổi
- Tính độc lập hay phụ thuộc trong trường hợp này được gọi là d-phân cách có điều kiện theo tập biến E



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



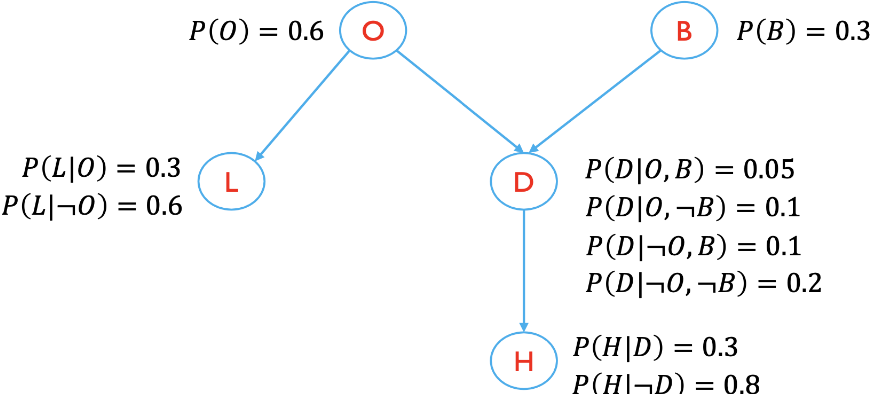
Suy diễn với mạng Bayes

Các nội dung đã học

- 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 suy diễn

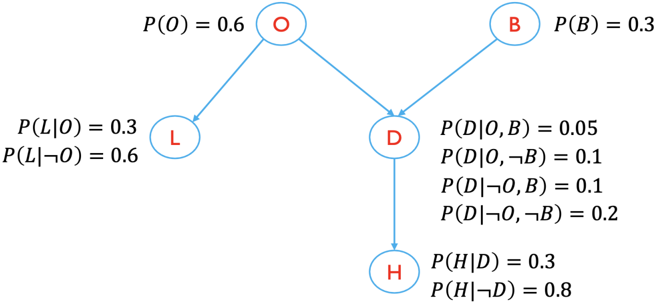
Ví dụ tính xác suất hậu nghiệm

Cần tính $P(L|B, \neg H)$



Ví dụ tính xác suất hậu nghiệm

Cần tính $P(L|B, \neg H)$



- Bước 1: Tính $P(L, B, \neg H)$
- Bước 2: Tính Tính $P(\neg L, B, \neg H)$
- Bước 3: Tính $\dfrac{P(L, B, \neg H)}{P(L, B, \neg H) + P(\neg L, B, \neg H)}$

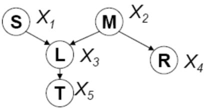
Trường hợp chung

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

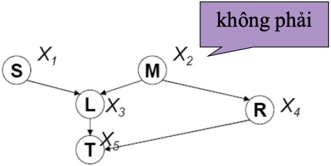
- Vấn đề:
  - Đòi hỏi liệt kê các xác suất đồng thời có chứa E_1 và E_2
  - Số lượng xác suất đồng thời như vậy tăng theo hàm mũ của số biến => Không thực thế
- Suy diễn nói chung trên mạng Bayes là bài toán NP đầy đủ!

Suy diễn trên thực tế

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



A poly tree



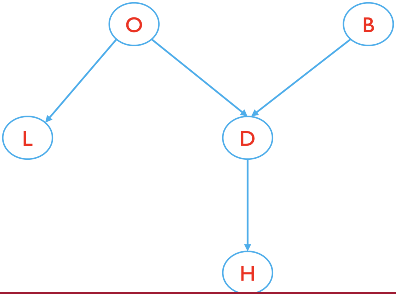
    - Tồn tại thuật toán với độ phức tạp tuyến tính cho poly tree
- Suy diễn xấp xỉ bằng cách lấy mẫu

Suy diễn với trường hợp riêng đơn giản nhất

- Suy diễn nhân quả
- Suy diễn chẩn đoán
- Suy diễn bằng cách lấy mẫu

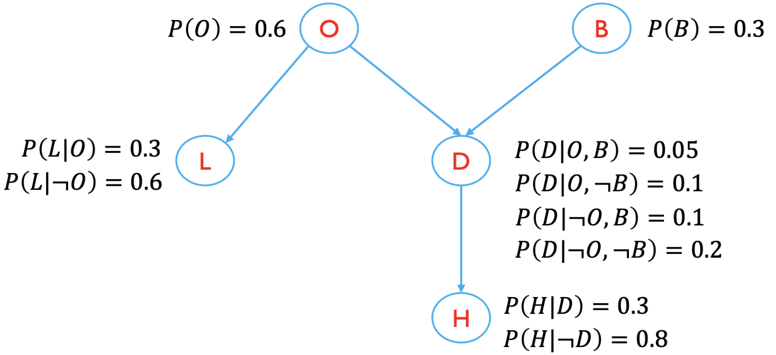
Suy diễn cho trường hợp riêng đơn giản

- 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 hai 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

- Cần tính: $P(D|B)$



Suy diễn nhân quả

$$
\begin{aligned}
& P(D|B) \\
& = P(D, O|B) + P(D, \neg O|B) \\
& = P(D|O, B) * P(O|B) + P(D|\neg O, B) * P(\neg O|B) \\
& = P(D|O, B) * P(O) + P(D|\neg O, B) * P(\neg O) \\
& = (0.05)(0.6) + (0.1)(1 - 0.6) = 0.07
\end{aligned}
$$

Chú ý: Sử dụng quy tắc chuỗi có điều kiện:
$P(A, B|C) = P(A|B, C) * P(B|C)$

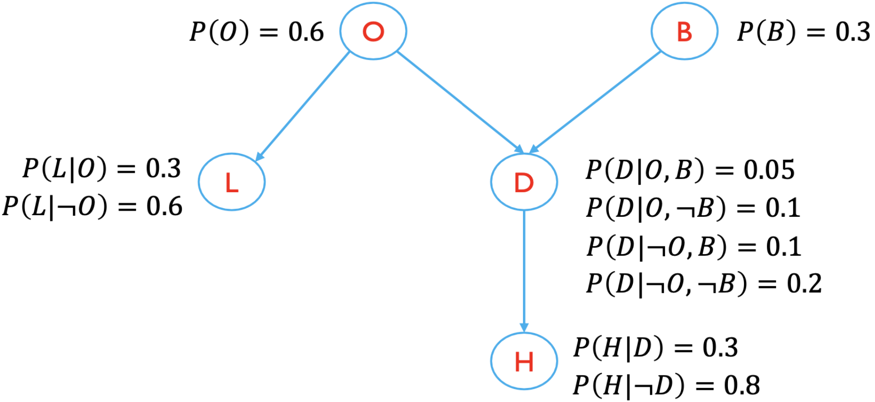
Suy diễn nhân quả

- Bước 1: *(xanh lá)* Xác suất Q cần tính được viết lại dưới dạng xác suất điều kiện của Q và cha của Q (không thuộc E), điều kiện theo E
- Bước 2: *(đỏ)* Viết lại các giá trị xác suất đồng thời dưới dạng xác suất của Q khi biết các giá trị cha mẹ của Q
- Bước 3: *(xanh dương)* Sử dụng các giá trị xác suất từ bảng xác suất có điều kiện để tính

Với ví dụ trước: $Q = \{D\}, E = \{B\}$. Vậy Bước 1: Đưa O vào (O không phụ thuộc E)

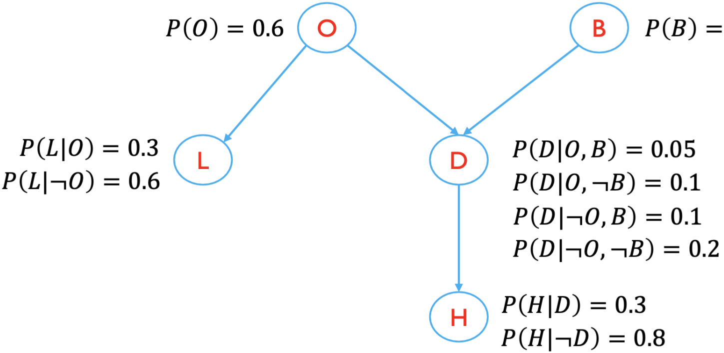
Suy diễn chẩn đoán (1/4)

Tính: $P(\neg B | \neg D)$



Suy diễn chẩn đoán (2/4)

- Theo Bayes: $P(\neg B|\neg D) = \dfrac{P(\neg D|\neg B).P(\neg B)}{P(\neg D)}$
- Tính $P(\neg D|\neg B)$ như phần trước sử dụng suy diễn nhân quả



Suy diễn chẩn đoán (3/4)

- $P(\neg D|\neg B) = P(\neg D|O, \neg B).P(O) + P(\neg D|\neg O, \neg B).P(\neg O)$
  $= (0.9)(0.6) + (0.8)(0.4) = 0.86$
- $P(\neg B|\neg D) = \dfrac{0.86 * 0.7}{P(\neg D)}$
- Cần tính $P(\neg D)$ => Tính $P(B|\neg D)$
- Tính:
  $P(B|\neg D) = \dfrac{P(\neg D|B).P(B)}{P(\neg D)} = \dfrac{(1 - 0.07) * 0.3}{P(\neg D)} = \dfrac{0.279}{P(\neg D)}$
- Sử dụng: $P(\neg B|\neg D) + P(B|\neg D) = 1 \rightarrow P(\neg D) = 0.881 \rightarrow$
  $P(\neg B|\neg D) = 0.683$

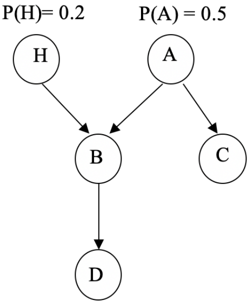
Suy diễn chẩn đoán (4/4)

Các bước thực hiện:

- Bước 1: Biến đổi về suy diễn nhân quả sử dụng quy tắc Bayes
- Bước 2: Thực hiện giống suy diễn nhân quả

Bài tập 1

Cho mạng Bayes như sau, các biến có thể nhận giá trị {T, F}



H	A	P(B=T | A, H)
F	F	0.7
F	T	0.2
T	F	0.1
T	T	0.5

A	P(C=T|A)
T	0.5
F	0.3

B	P(D=T|B)
T	0.3
F	0.5

a) Tính xác suất cả năm biến cùng nhận giá trị F

b) Tính $P(H = T|B)$

c) Mạng đã cho có dạng Polytree hay không?

Bài tập 2 (1/2)

Nhằm tìm hiểu nguyên nhân hỏng hóc của robot thám hiểm tự động một đội kỹ thuật đã thu thập được dữ liệu như sau:

- Chập điện thì do mưa và bụi, bụi gây nên mờ camera và chập điện dẫn tới hỏng robot.
- Nếu robot chập điện thì 20% là không hỏng, ngược lại nếu robot không chập điện thì 60% là không hỏng
- Nếu có bụi mà mưa thì 90% chập điện.
- Nếu không có bụi mà mà không mưa thì 20% chập điện.
- Nếu có bụi mà trời không mưa thì 70% chập điện.
- Nếu không bụi mà trời mưa thì 80% chập điện.
- Nếu có bụi thì 30% là mờ camera và nếu không bụi thì 80% là không mờ camera.
- Xác suất có bụi là 30% và có mưa là 20%.

Bài tập 2 (2/2)

a) Từ những câu trên hãy xây dựng mạng Bayes với các biến ngẫu nhiên dưới đây, mỗi biến có thể nhận giá trị {T, F}: **D**: Chập điện, **C**: Mờ camera, **B**: Bụi, **M**: Mưa, **H**: Hỏng

b) Tính xác suất đồng thời xảy ra 3 sự kiện sau: robot hỏng, mờ camera, và trời mưa.

c) Tính xác suất có bụi nếu biết camera bị mờ.

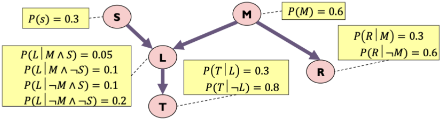
Suy diễn bằng cách lấy mẫu

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

Lấy mẫu

- Trong mạng bayes, lấy mẫu các biến theo thứ tự trên mạng tức là các nút cha mẹ sẽ được lấy mẫu trước, nút con sẽ được lấy mẫu sau. Mỗi biến sẽ được lấy mẫu theo xs điều kiện tuỳ thuộc vào giá trị nút cha mẹ đã được gán.

Ví dụ lấy mẫu (1/2)



- Chọn ngẫu nhiên $S : P(S = true) = 0.3$
- 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
  - Gỉa 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 xảy ra các sự kiện sau:
  - N_c: số mẫu có $T = true, S = false$
  - N_s: số mẫu có R=true, $T = true, S = false$
  - N: tổng số mẫu
- Nếu N đủ lớn:
  - N_c/N: xấp xỉ xác suất $P(T = true, S = false)$
  - N_c/N: xấp xỉ xác suất P(R=true, $T = true, S = false$)
  - $P(R|T, \neg S) = P(R, T, \neg S)/P(T, \neg S) \approx N_s/N_c$

Lấy mẫu tổng quát

- Cần tính $P(E_1|E_2)$
- Lấy mẫu số lượng đủ lớn
- Tính số lượng:
  - N_c: số mẫu có E_2
  - N_s: số mẫu có E_1 và E_2
  - N: tổng số mẫu
- Nếu N đủ lớn, ta có: $P(E_1|E_2) = \dfrac{N_s}{N_c}$