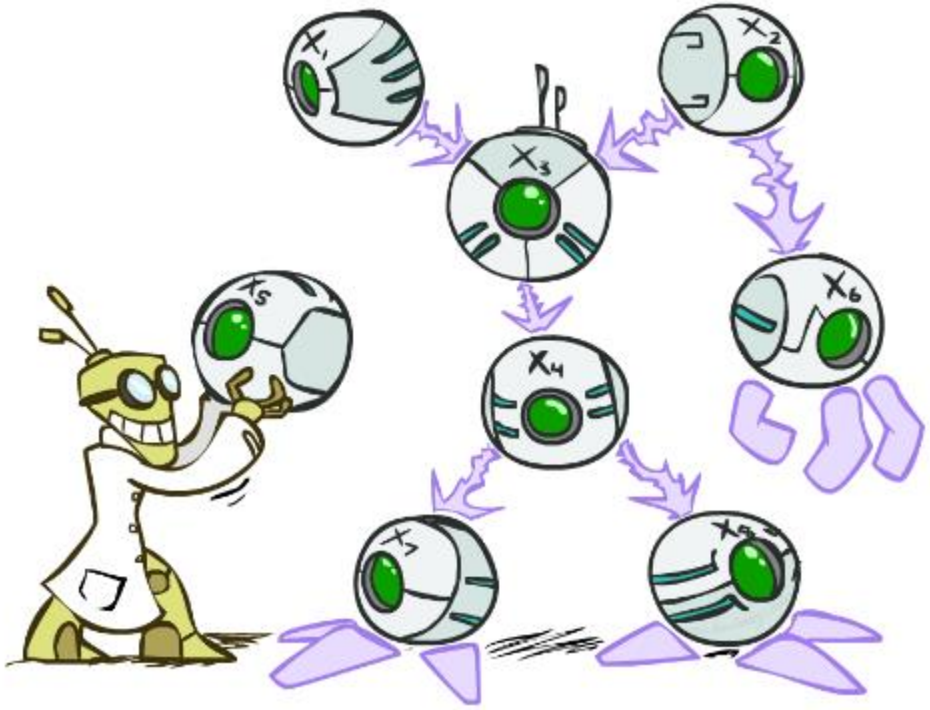
TRUONG DAI HOC KHOA HOC TU NHIEN
KHTN
TP. HO CHI MINH

MẠNG BAYES



TRÌNH BÀY: LÊ THÀNH CÔNG – 1612842

Outline

- Mô hình xác suất
- Luật Bayes
- Độc lập, độc lập có điều kiện và quy tắc mắc xích
- Mạng Bayes
- D-separation

Mô hình xác suất

- Các mô hình chỉ có thể mô tả một phần thế giới hoạt động
- Các mô hình luôn đơn giản hóa
  - Có thể không giải thích được cho mọi biến
  - Có thể không giải thích được mọi sự tương tác giữa các biến
  - "All models are wrong; but some are useful." – George E. P. Box
- Chúng ta làm gì với các mô hình xác suất?
  - Chúng ta cần suy luận biến chưa biết từ biến đã biết
  - Giải thích (Lý luận chuẩn đoán)
  - Dự đoán (Lý luận nhân quả)
  - Giá trị thông tin

Luật Bayes

- Từ luật nhân P(ab) = P(a|b).P(b) = P(b|a).P(a)

→ Luật Bayes: **P(a|b) = P(b|a).P(a) / P(b)**

- Đây là luật được sử dụng trong hầu hết các hệ thống AI hiện đại sử dụng lập luận xác suất, đặc biệt để đánh giá xác suất chuẩn đoán từ xác suất nguyên nhân
- VD: m là bệnh viêm màng não và s là bệnh cứng cổ

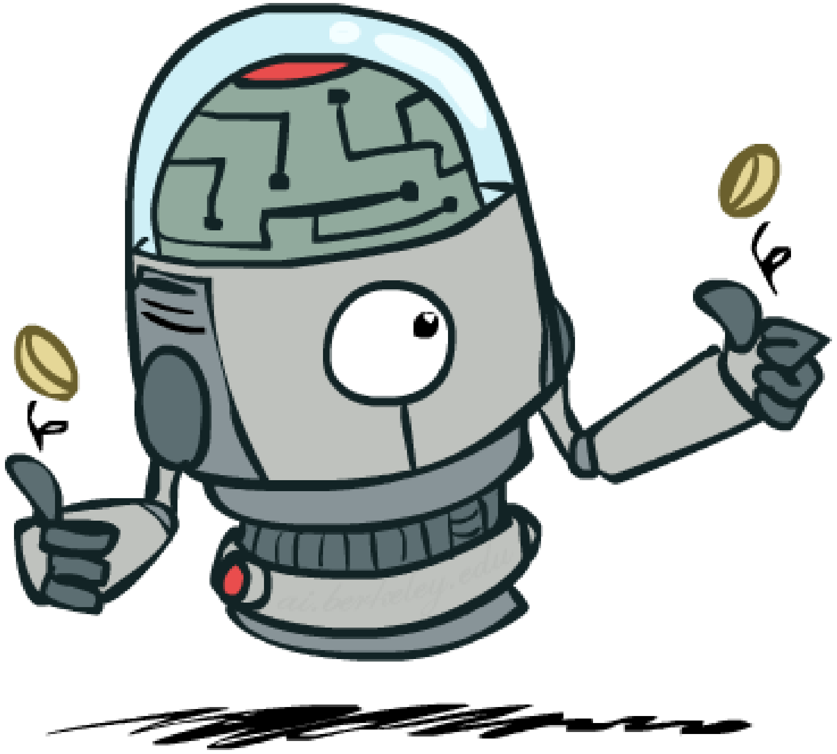
P(s|m) = 0.5

P(m) = 1/50000

$$P(m|s) = \frac{P(S|m)P(m)}{P(s)} = 0.5 * \frac{20}{50000} = 0.0002$$

P(s) = 1/20

Độc lập



Độc lập

- Hai biến được gọi là độc lập nếu:

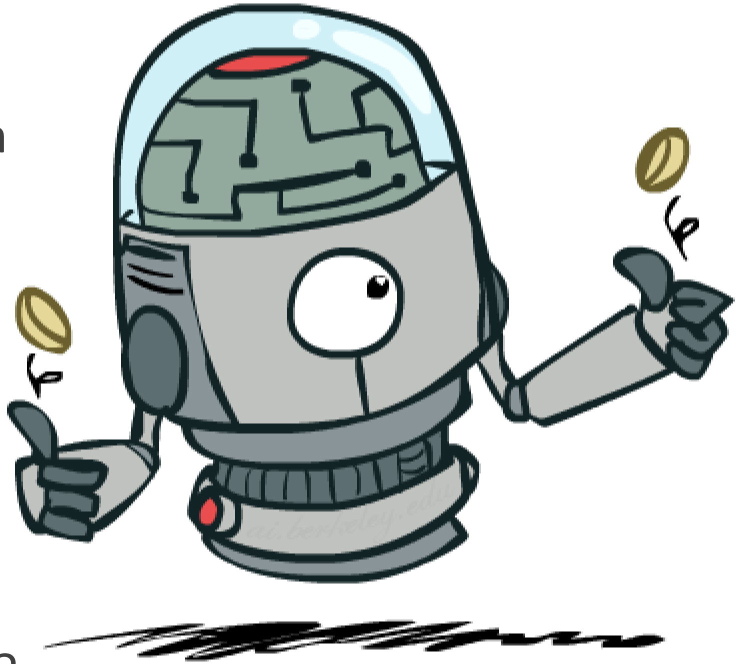
$$\forall x, y : P(x, y) = P(x)P(y)$$

  - Phân phối chung được phân tích thành tích 2 phân phối đơn giản hơn
  - Dạng khác:

$$\forall x, y : P(x|y) = P(x)$$

  - Ký hiệu độc lập: $X \perp\!\!\!\perp Y$

- Độc lập là một giả định mô hình đơn giản hóa
  - Phân phối chung trên thực tế hiếm khi độc lập hoàn toàn
  - Chúng ta có thể kết luận được gì {Weather, Traffic, Cavity, Toothache}?



Ví dụ: Độc lập?

$P(T)$

T	P
hot	0.5
cold	0.5

$P_1(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

$P_2(T, W)$

T	W	P
hot	sun	0.3
hot	rain	0.2
cold	sun	0.3
cold	rain	0.2

$P(W)$

W	P
sun	0.6
rain	0.4

Ví dụ: Độc lập?

- Tung N đồng xu độc lập cùng lúc

$P(X_1)$

H	0.5
T	0.5

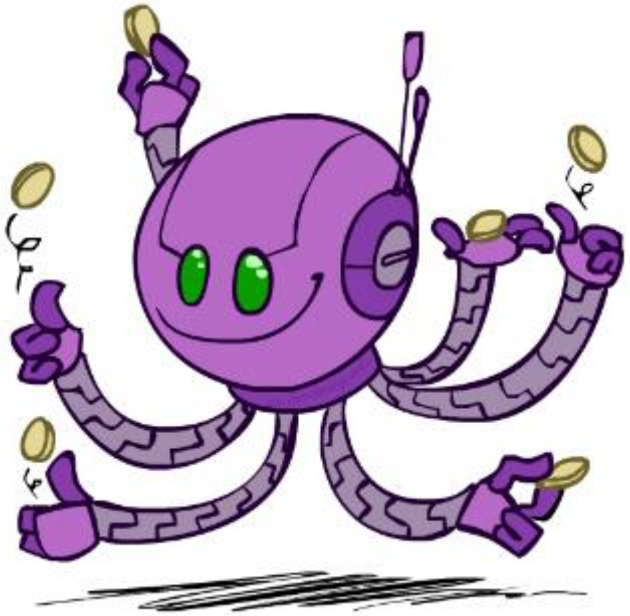
$P(X_2)$

H	0.5
T	0.5

$\cdots$

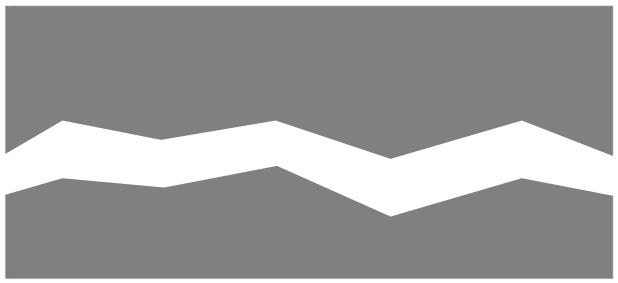
$P(X_n)$

H	0.5
T	0.5

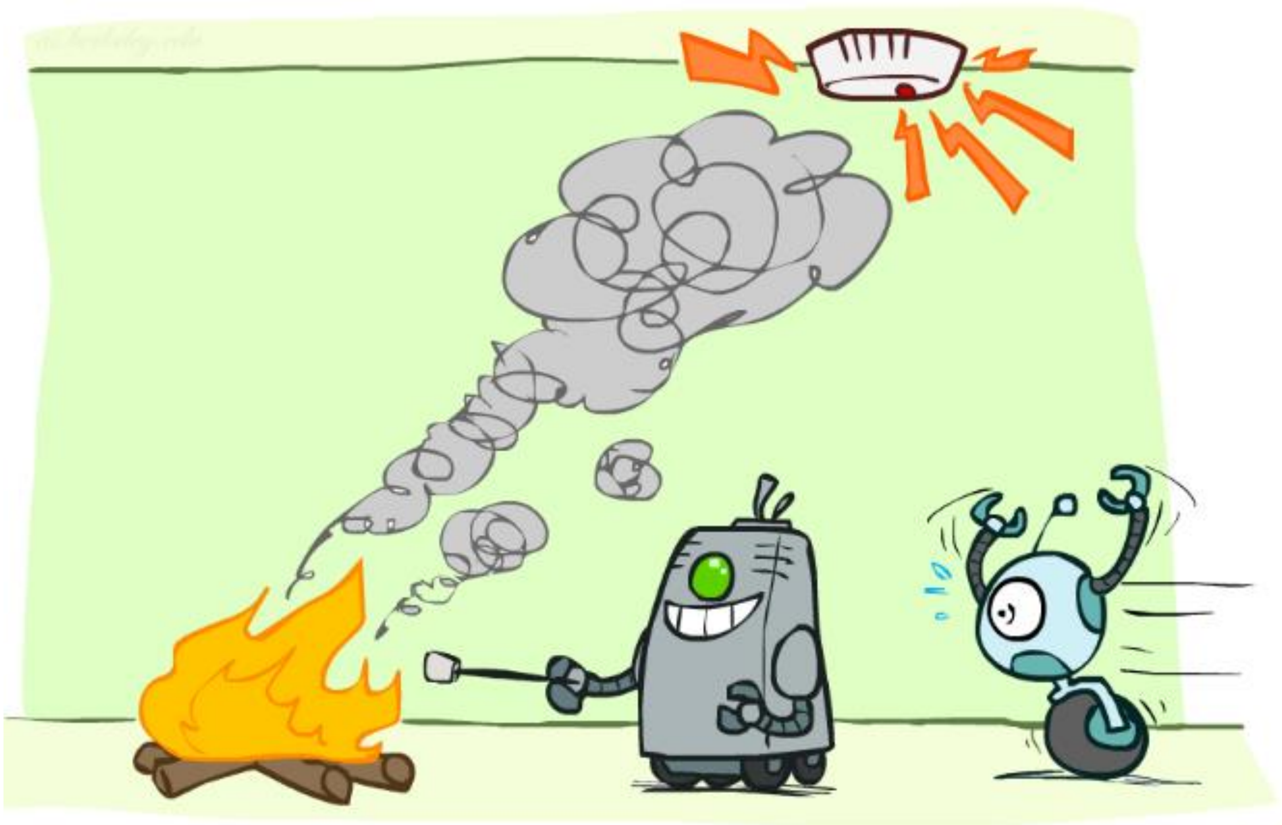


$P(X_1, X_2, \ldots X_n)$

2^n

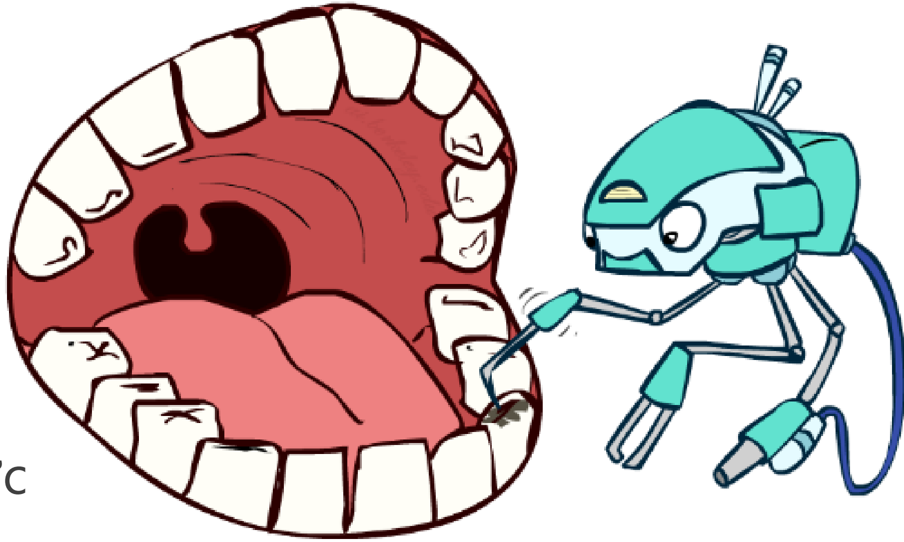


Độc lập có điều kiện



Độc lập có điều kiện

- P(Toothache, Cavity, Catch)
- Nếu bị lỗ sâu trong răng, xác suất robot tìm được nó không phụ thuộc vào việc có đau răng hay không
  - P(+catch|+toothache,+cavity) = P(+catch|+cavity)
- Tương tự nếu như không có lỗ sâu răng
  - P(+catch|+toothache,-cavity) = P(+catch|-cavity)
- Catch độc lập có điều kiện với Toothache và Cavity cho trước
  - P(Catch | Toothache, Cavity) = P(Catch, Cavity)
- Các phát biểu tương đương:
  - P(Toothache | Catch , Cavity) = P(Toothache | Cavity)
  - P(Toothache, Catch | Cavity) = P(Toothache | Cavity) P(Catch | Cavity)



Độc lập có điều kiện

- Độc lập không điều kiện khá hiếm? Tại sao?
- Độc lập có điều kiện là nền tảng kiến thức cơ bản và vững chắc nhất trong môi trường nghiên cứu không chắc chắn
- X độc lập có điều kiện với Y cho trước Z $X \perp\!\!\!\perp Y | Z$

  khi và chỉ khi:

$$\forall x, y, z : P(x, y|z) = P(x|z)P(y|z)$$

  hoặc khi và chỉ khi:

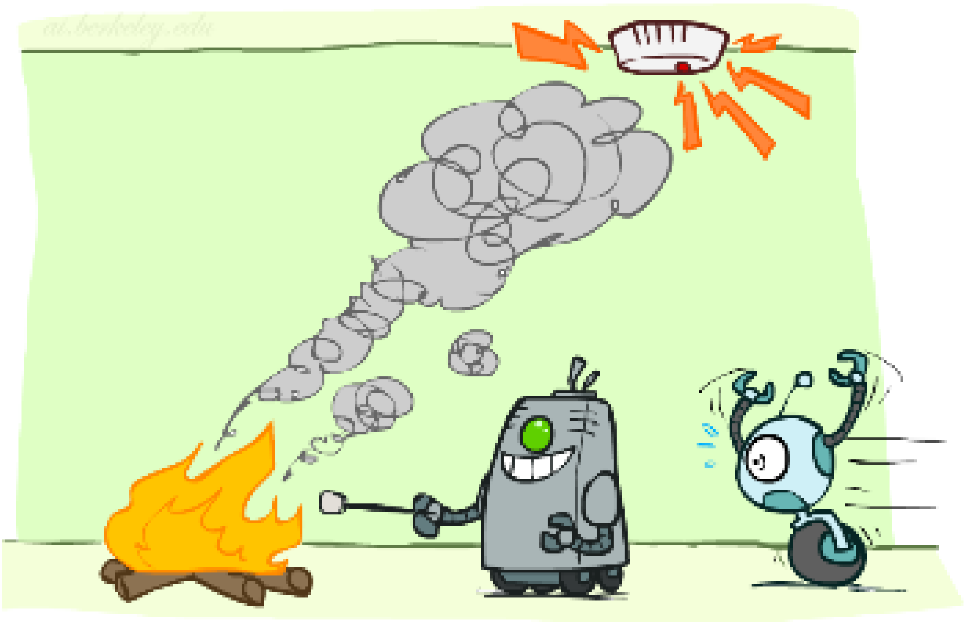
$$\forall x, y, z : P(x|z, y) = P(x|z)$$

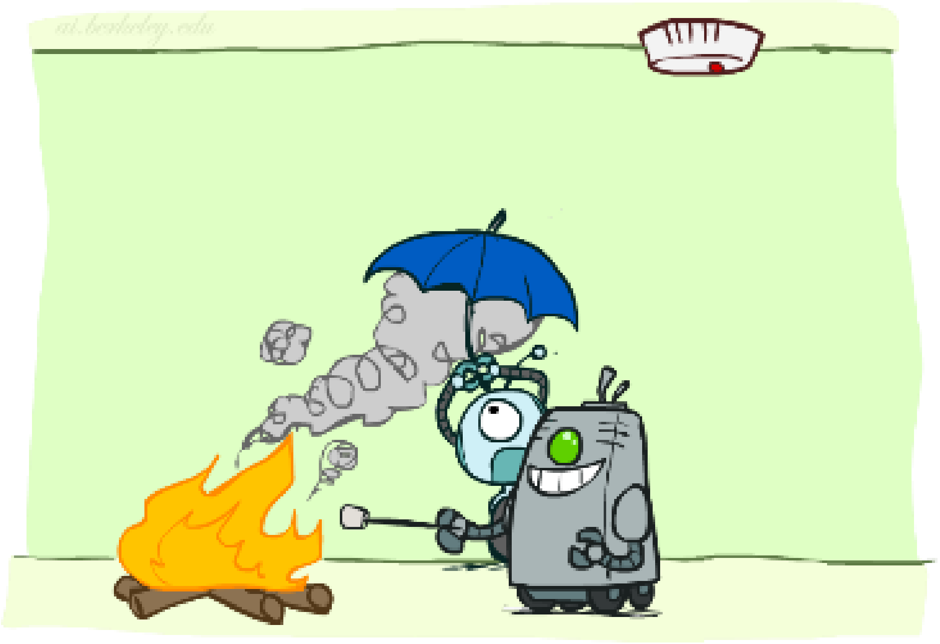
Độc lập có điều kiện

- Xét ngữ cảnh sau
  - Traffic
  - Umbrella
  - Raining

Độc lập có điều kiện

- Xét ngữ cảnh sau
  - Fire
  - Smoke
  - Alarm





Độc lập có điều kiện - quy tắc mắc xích

- Quy tắc mắc xích: $P(X_1, X_2, \ldots X_n) = P(X_1)P(X_2|X_1)P(X_3|X_1, X_2)\ldots$
- Phân giải tầm thường:

$$P(\text{Traffic}, \text{Rain}, \text{Umbrella}) =$$
$$P(\text{Rain})P(\text{Traffic}|\text{Rain})P(\text{Umbrella}|\text{Rain}, \text{Traffic})$$

- Kết hợp độc lập có điều kiện:

$$P(\text{Traffic}, \text{Rain}, \text{Umbrella}) =$$
$$P(\text{Rain})P(\text{Traffic}|\text{Rain})P(\text{Umbrella}|\text{Rain})$$

- Mạng Bayes giúp thể hiện mối quan hệ độc lập có điều kiện

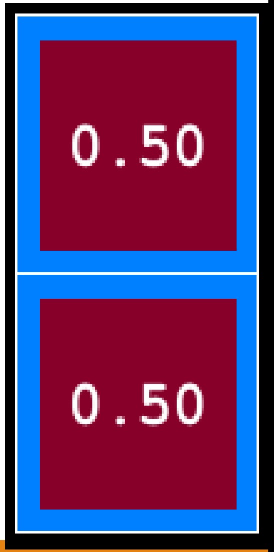
Quy tắc mắc xích - Ghostbusters

- Mỗi cảm biến phụ thuộc vào vị trí con ma
- Hai cảm biến độc lập có điều kiện với vị trí ma cho trước
- T : Vị trí trên màu đỏ
- B : Vị trí dưới màu đỏ
- G : Ma ở trên tầng trên
- Cho:

$P(+g) = 0.5$

$P(-g) = 0.5$

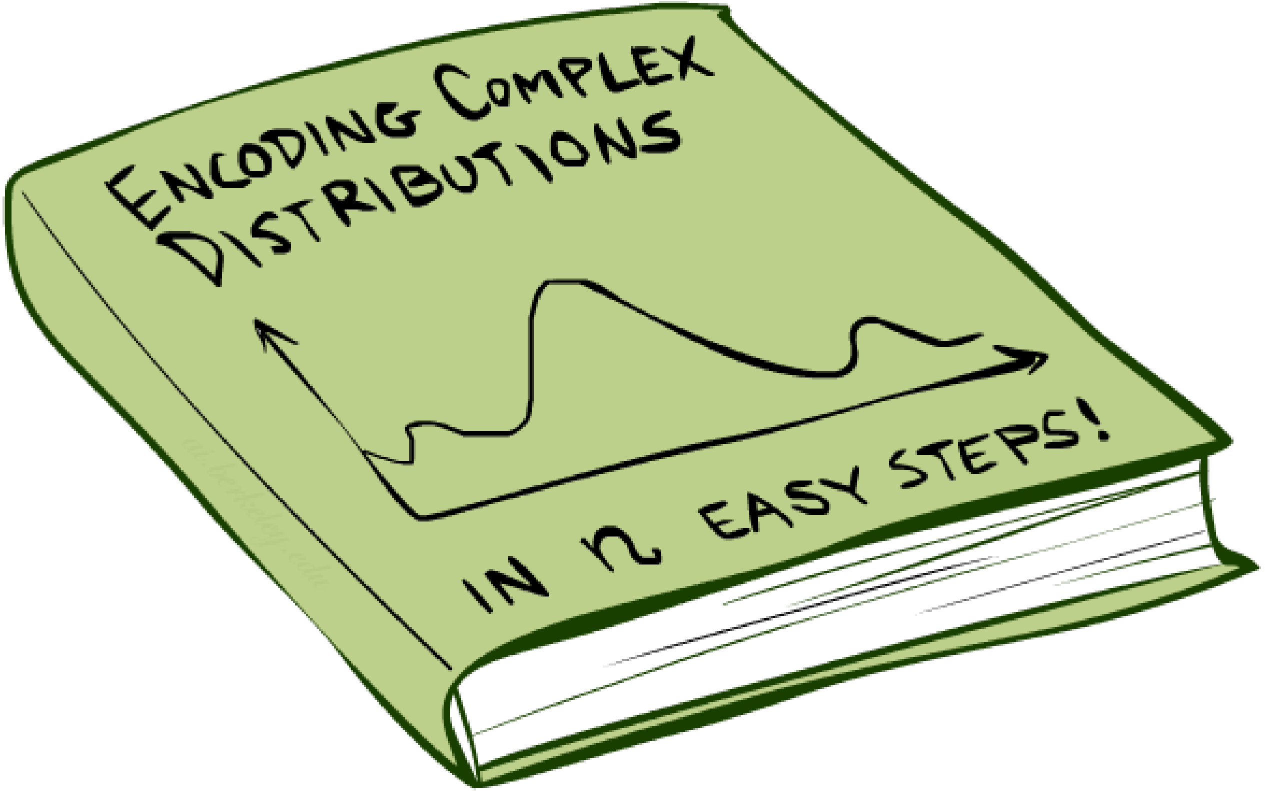
$P(+t \mid +g) = 0.8$
$P(+t \mid -g) = 0.4$
$P(+b \mid +g) = 0.4$
$P(+b \mid -g) = 0.8$



$P(T,B,G) = P(G)\,P(T|G)\,P(B|G)$

T	B	G	P(T,B,G)
+t	+b	+g	0.16
+t	+b	-g	0.16
+t	-b	+g	0.24
+t	-b	-g	0.04
-t	+b	+g	0.04
-t	+b	-g	0.24
-t	-b	+g	0.06
-t	-b	-g	0.06

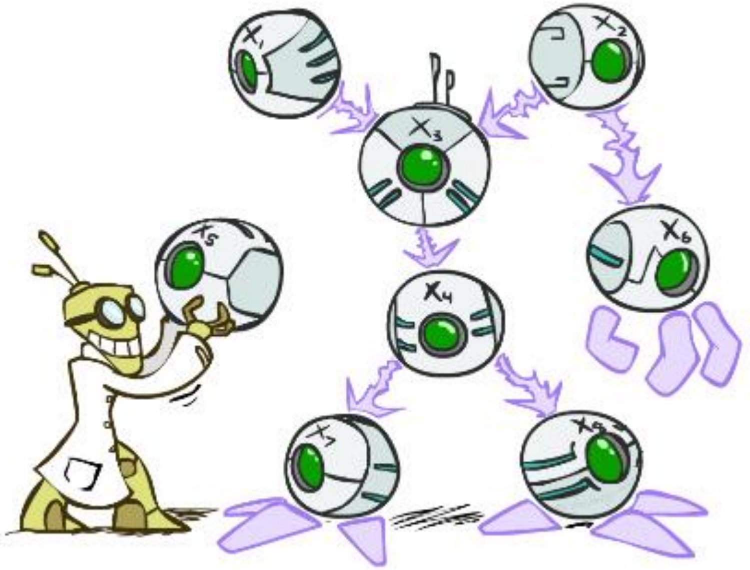
Mạng Bayes – Bức tranh toàn cảnh



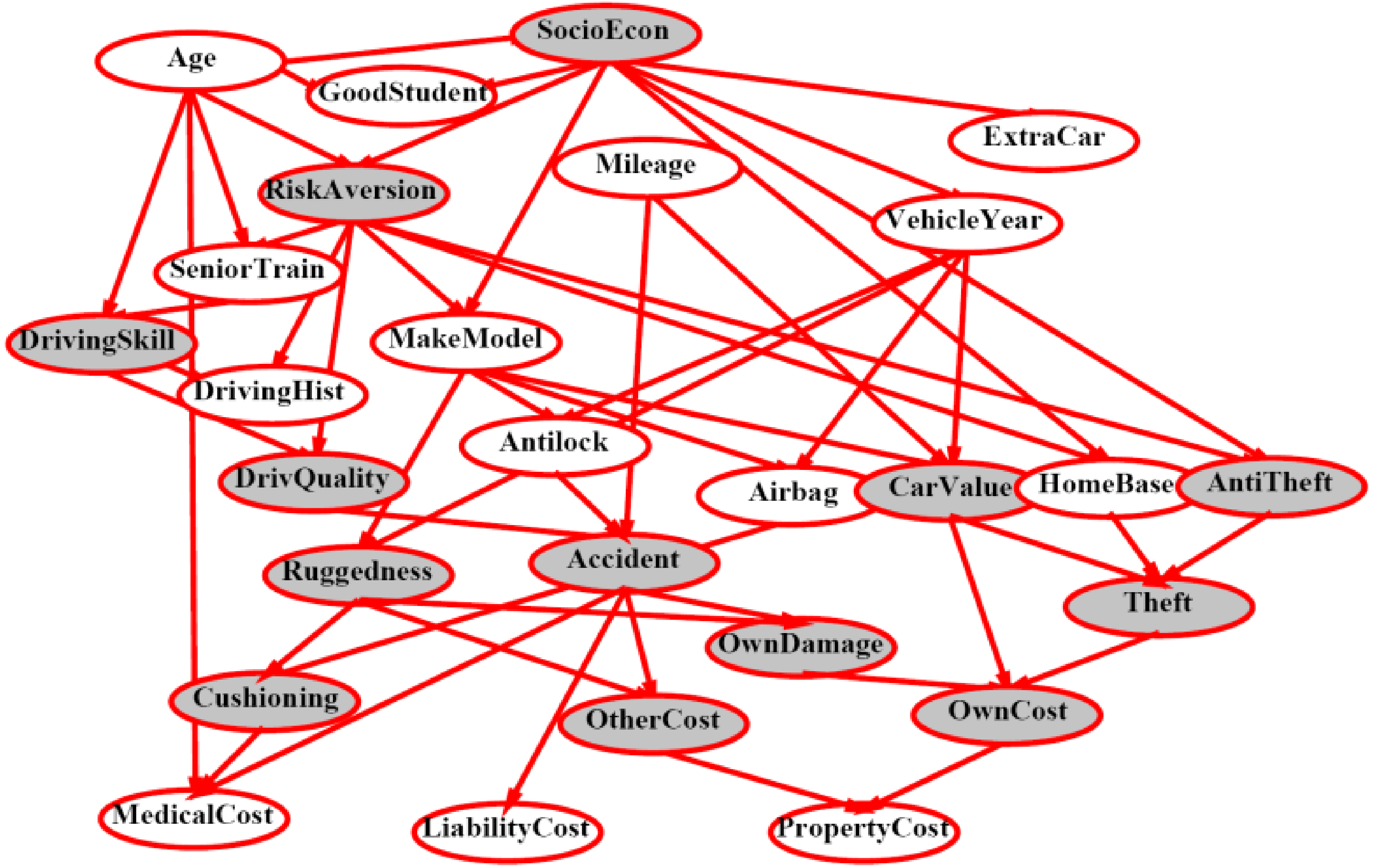
Mạng Bayes – Bức tranh toàn cảnh

- Sử dụng bảng phân phối chung đầy đủ làm mô hình xác suất gặp 2 vấn đề sau
  - Quá lớn để thể hiện rõ ràng nếu nhiều biến
  - Khó để học, ước tính một cách thực tiễn nhiều biến ở một thời điểm

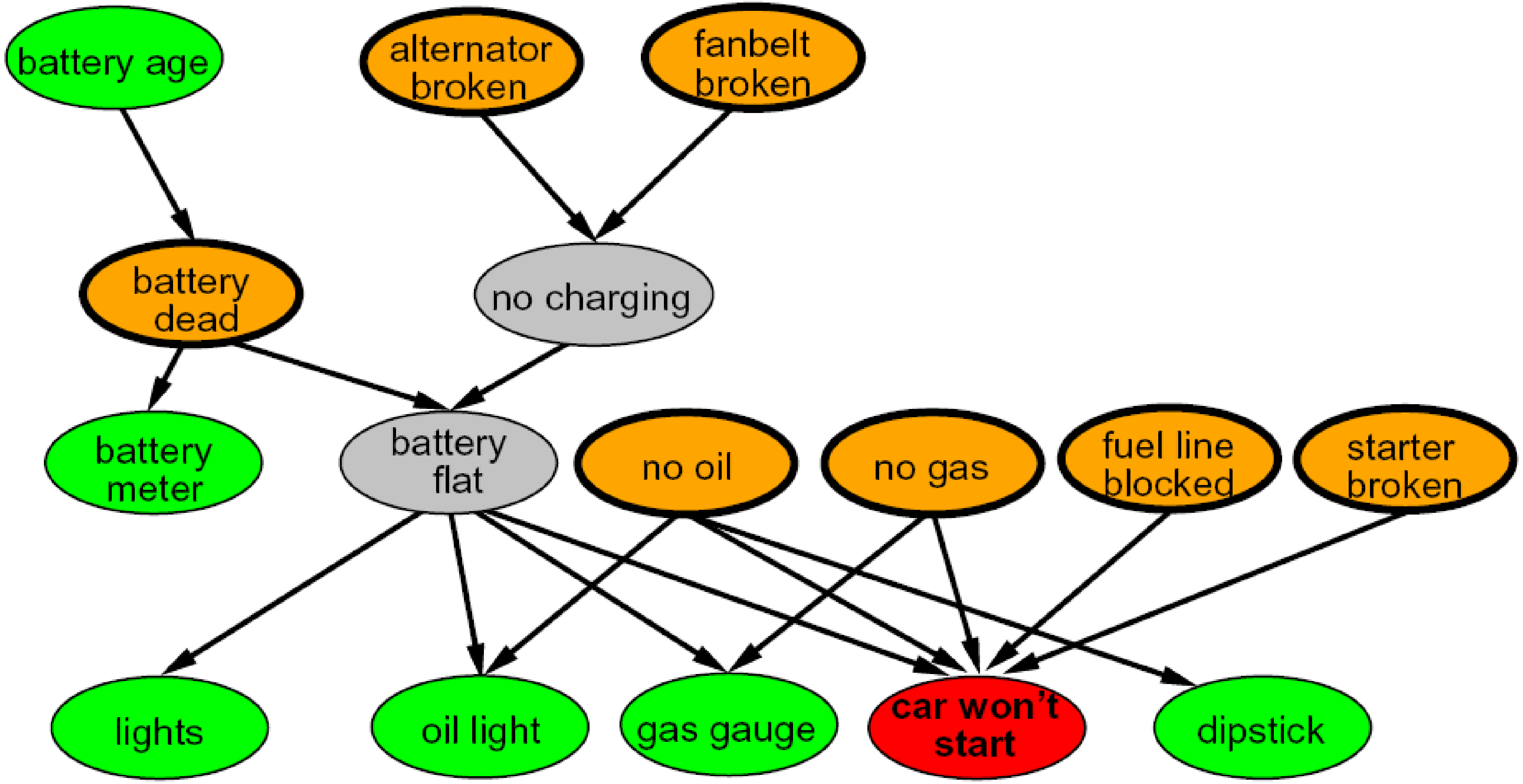
- Mạng Bayes là một kỹ thuật mô tả các phân phối phức tạp (model) nhờ sử dụng các phân phối đơn giản (độc lập có điều kiện)
  - Thường được gọi là mô hình đồ thị
  - Mô tả các biến tương tác cục bộ với nhau như thế nào
  - Tương tác cục bộ mắc xích tạo thành toàn cục, tương tác gián tiếp



Ví dụ mạng Bayes – Bảo hiểm

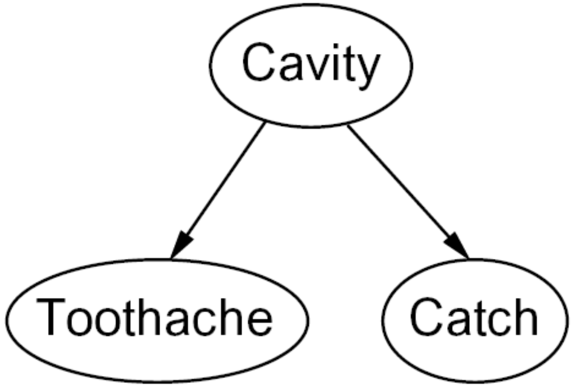


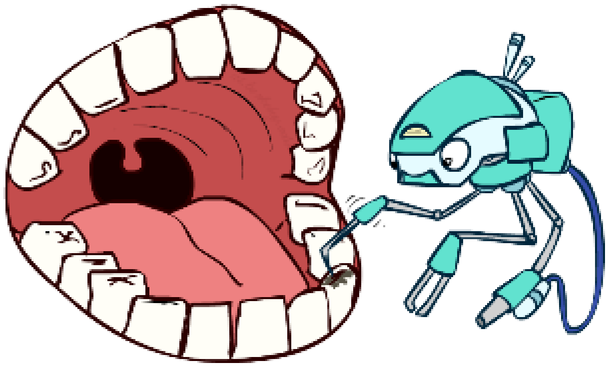
Ví dụ mạng Bayes – Xe hơi



Ký hiệu mô hình đồ thị

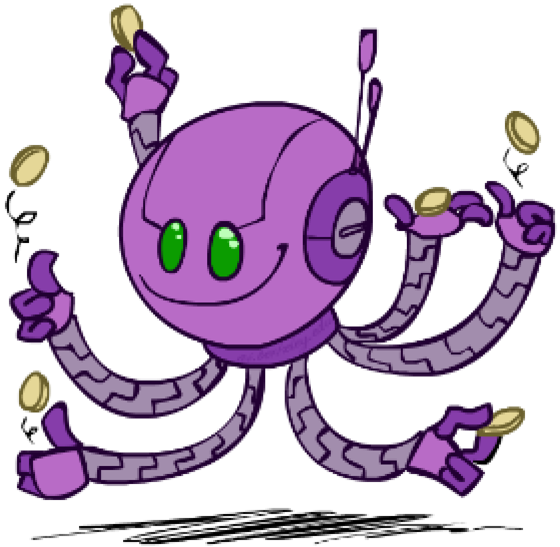
- Nodes: các biến
  - Có thể được quan sát hoặc chưa được quan sát
- Arcs: thể hiện sự tương tác
  - Giống như ràng buộc CSP
  - Xác định ảnh hưởng trực tiếp giữa 2 biến
  - Dùng để biểu diễn độc lập có điều kiện
- Cứ tưởng tượng mũi tên như quan hệ nhân quả trực tiếp





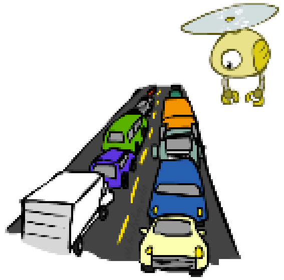
Ví dụ: Tung đồng xu

- Tung N đồng xu độc lập

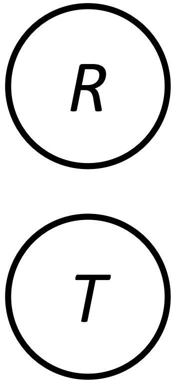


- Không có sự tương tác giữa các biến: độc lập hoàn toàn

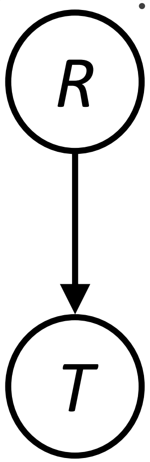
Ví dụ: Giao thông



- Các biến:
  - R: Rain
  - T: Traffic
- Mô hình 1: độc lập



- Model 2: mưa làm kẹt xe



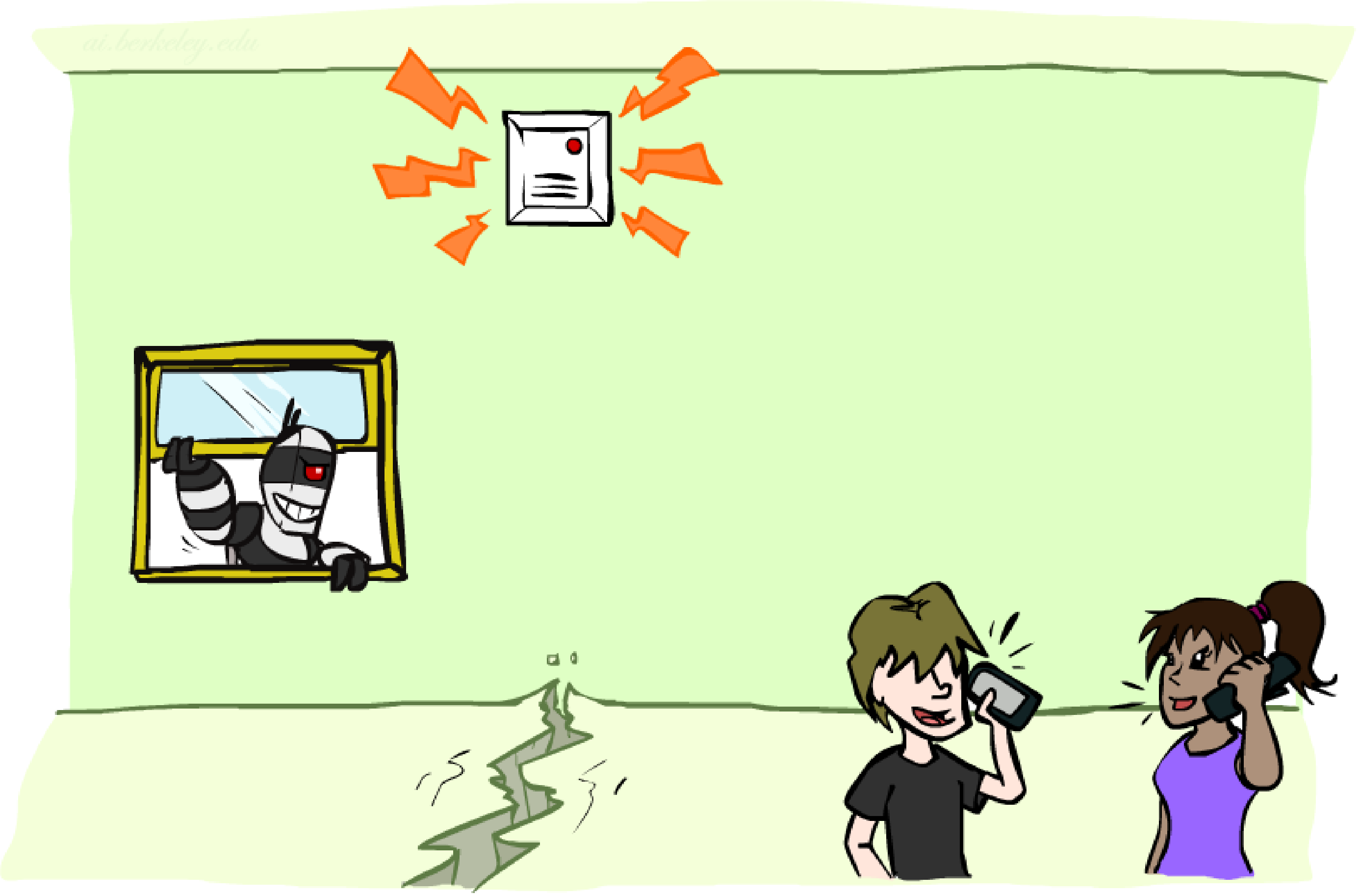
- Tại sao sử dụng mô hình 2 sẽ tốt hơn?

Ví dụ: Giao thông II

- Xây dựng một mô hình đồ thị
- Biến:
  - T: Traffic
  - R: It rains
  - L: Low pressure
  - D: Roof drips
  - B: Ball game
  - C: Cavity

Ví dụ: Mạng lưới báo động

- Xây dựng một mô hình đồ thị
- Biến:
  - B: Burglary
  - A: Alarm goes off
  - M: Mary calls
  - J: John calls
  - E: Earthquake!



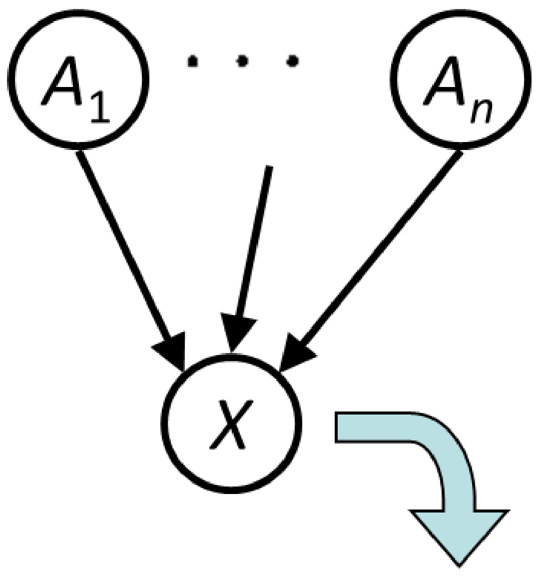
Ngữ nghĩa mạng Bayes

Ngữ nghĩa mạng Bayes

- Một tập hợp các node, mỗi node một biến
- Đồ thị có hướng, không chu trình
- Phân phối có điều kiện tại mỗi node:
  - Công thức tính tại X với a_1 … a_n là các node cha:

$$P(X|a_1 \ldots a_n)$$

  - CPT: conditional probability table



$$P(X|A_1 \ldots A_n)$$

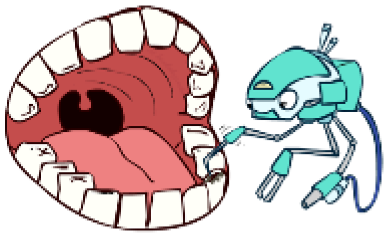
A Bayes net = Topology (graph) + Local Conditional Probabilities

Xác suất trong mạng Bayes

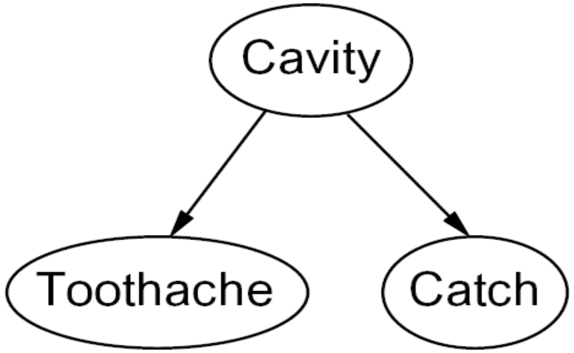
- Mạng Bayes ngầm ám chỉ các phân phối chung
  - Là tích của các phân phối có điều kiện cục bộ
  - Để tính xác suất dạng đủ, nhân tất cả các điều kiện liên quan

$$P(x_1, x_2, \ldots x_n) = \prod_{i=1}^{n} P(x_i | parents(X_i))$$

  - Ví dụ:



$P(+cavity, +catch, -toothache)$



Xác suất trong mạng Bayes

- Liệu công thức này có trả về xác suất chung hợp lý?

$$P(x_1, x_2, \ldots x_n) = \prod_{i=1}^{n} P(x_i | parents(X_i))$$

- Quy tắc mắc xích (hợp lệ mọi phân phối): $P(x_1, x_2, \ldots x_n) = \prod_{i=1}^{n} P(x_i | x_1 \ldots x_{i-1})$
- Giả thiết độc lập có điều kiện: $P(x_i | x_1, \ldots x_{i-1}) = P(x_i | parents(X_i))$
- Kết quả: $P(x_1, x_2, \ldots x_n) = \prod_{i=1}^{n} P(x_i | parents(X_i))$
- Không phải mạng Bayes nào cũng thể hiện được mọi phân phối

Ví dụ: Tung đồng xu

X_1 X_2 $\cdots$ X_n

$P(X_1)$

h	0.5
t	0.5

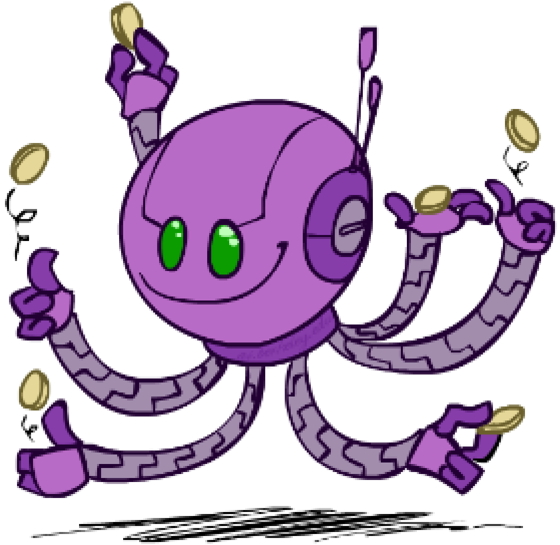
$P(X_2)$

h	0.5
t	0.5

$\cdots$

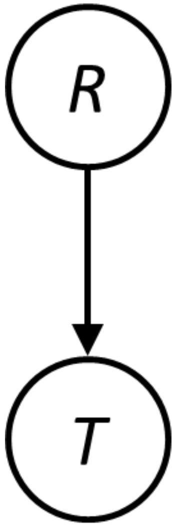
$P(X_n)$

h	0.5
t	0.5



$$P(h, h, t, h) =$$

Ví dụ: Giao thông



$$P(R)$$

+r	1/4
-r	3/4

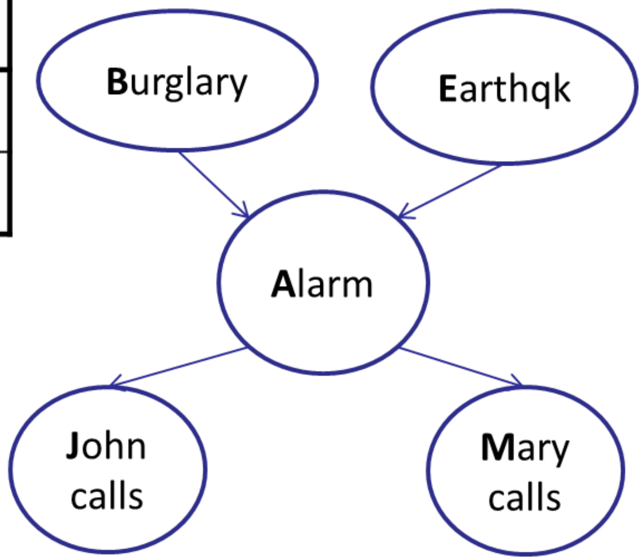
$$P(T|R)$$

+r	+t	3/4
	-t	1/4
-r	+t	1/2
	-t	1/2

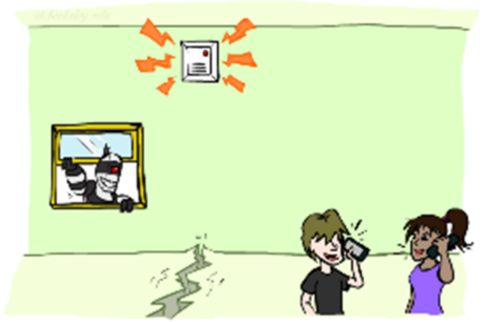
$$P(+r, -t) =$$

Ví dụ: Mạng lưới báo động

B	P(B)
+b	0.001
-b	0.999



E	P(E)
+e	0.002
-e	0.998



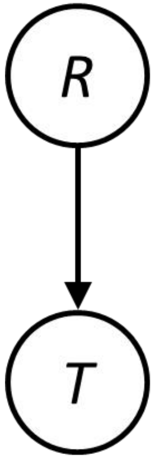
B	E	A	P(A|B,E)
+b	+e	+a	0.95
+b	+e	-a	0.05
+b	-e	+a	0.94
+b	-e	-a	0.06
-b	+e	+a	0.29
-b	+e	-a	0.71
-b	-e	+a	0.001
-b	-e	-a	0.999

A	J	P(J|A)
+a	+j	0.9
+a	-j	0.1
-a	+j	0.05
-a	-j	0.95

A	M	P(M|A)
+a	+m	0.7
+a	-m	0.3
-a	+m	0.01
-a	-m	0.99

Ví dụ: Giao thông

- Hướng nhân quả



$P(R)$

+r	1/4
-r	3/4

$P(T|R)$

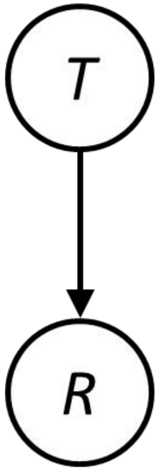
+r	+t	3/4
	-t	1/4
-r	+t	1/2
	-t	1/2

$P(T,R)$

+r	+t	3/16
+r	-t	1/16
-r	+t	6/16
-r	-t	6/16

Ví dụ: Giao thông

- Nhân quả đảo



$P(T)$

+t	9/16
-t	7/16

$P(R|T)$

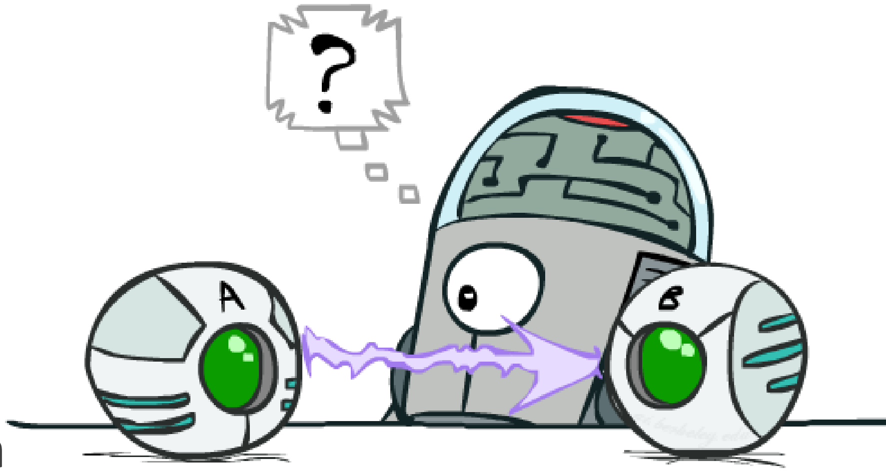
+t	+r	1/3
	-r	2/3

-t	+r	1/7
	-r	6/7

+r	+t	3/16
+r	-t	1/16
-r	+t	6/16
-r	-t	6/16

Quan hệ nhân quả?

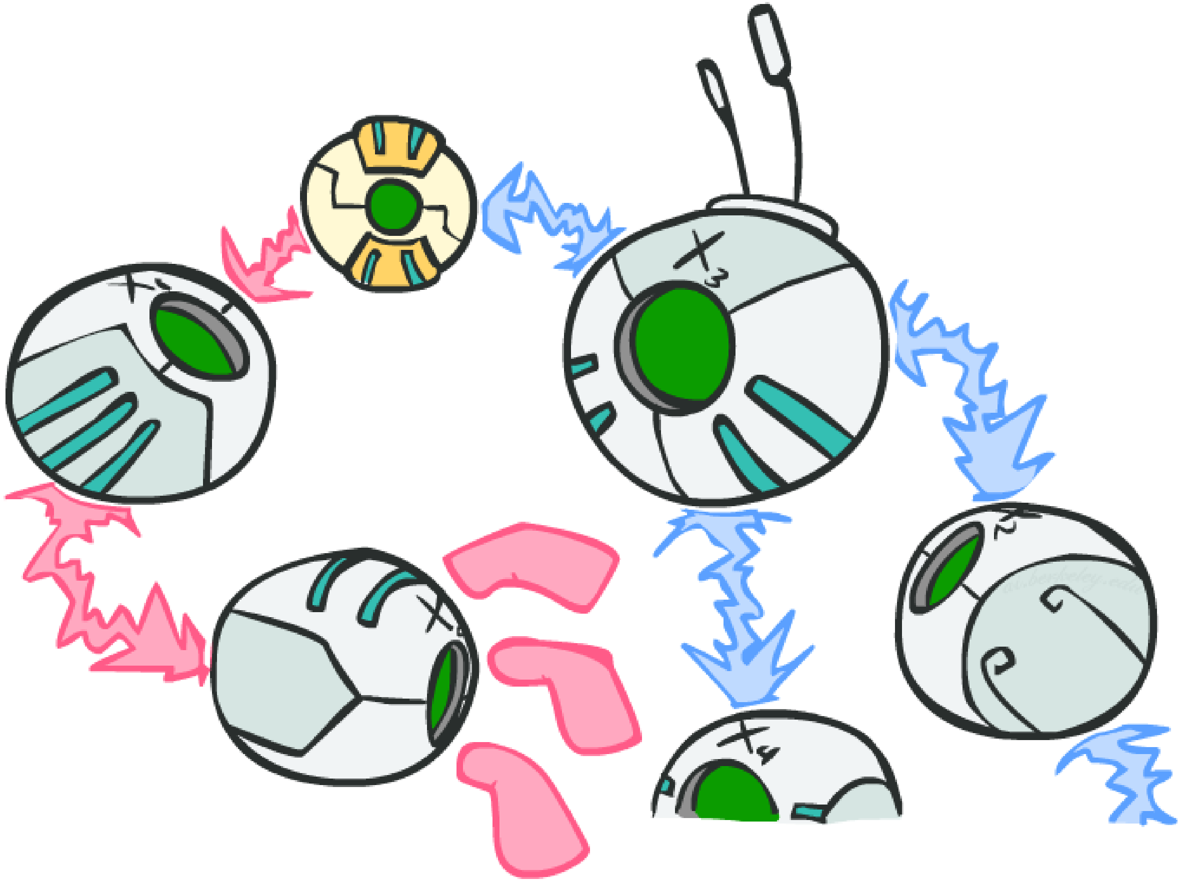
- Khi mạng Bayes phản ánh mẫu nhân quả thực sự:
  - Thường đơn giản (ít node cha)
  - Thường dễ suy nghĩ, dễ nhớ
  - Dễ suy luận từ các chuyên gia



- Mạng Bayes không cần thực sự quan hệ nhân quả
  - Thỉnh thoảng không có mạng nhân quả tồn tại trong miền
  - Kết thúc bằng mũi tên, chỉ sự tương quan, không quả quan hệ nhân quả
- Ý nghĩa thực sự của mũi tên?
  - Cấu trúc liên kết thực sự mã hóa độc lập có điều kiện

$$P(x_i|x_1, \ldots x_{i-1}) = P(x_i|\mathit{parents}(X_i))$$

D-separation: Outline



D-separation: Outline

- Nghiên cứu bộ ba các tính chất độc lập
- Đánh giá các trường hợp phức tạp liên quan bộ ba
- D-seperation: Một điều kiện/thuật toán trả lời truy vấn

Causal chain

- Mô hình causal chain

X: Low pressure Y: Rain Z: Traffic

$$P(x, y, z) = P(x)P(y|x)P(z|y)$$

- Liệu có đảm bảo X độc lập với Z ? Không
  - Một ví dụ về các CPT mà X không độc lập với Z là đủ để cho thấy tính độc lập này không được đảm bảo.
  - Ví dụ:
    - Áp suất thấp gây mưa gây kẹt xe
    - Áp suất cao thì không mưa không kẹt xe
    - Bằng số:

      P(+y | +x) = 1, P(-y | - x) = 1,

      P(+z | +y) = 1, P(-z | -y) = 1

Causal chain

- Mô hình causal chain

X: Low pressure Y: Rain Z: Traffic

$$P(x, y, z) = P(x)P(y|x)P(z|y)$$

- Liệu đảm bảo X độc lập có điều kiện với Z và Y cho trước?

$$P(z|x, y) = \frac{P(x, y, z)}{P(x, y)}$$

$$= \frac{P(x)P(y|x)P(z|y)}{P(x)P(y|x)}$$

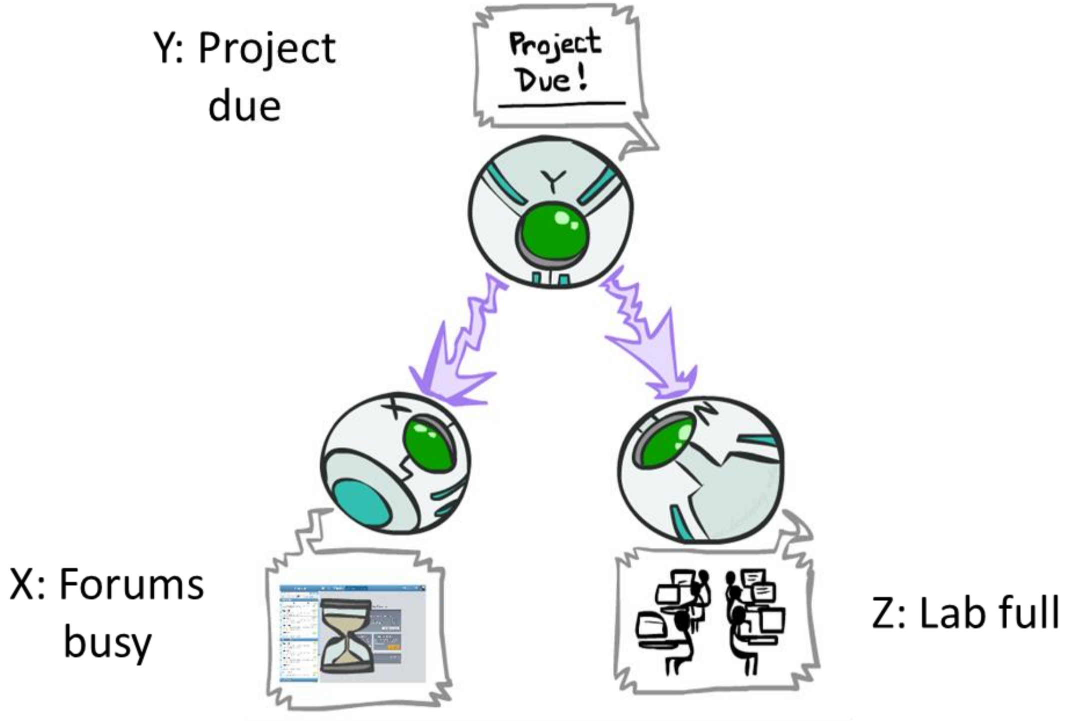
$$= P(z|y)$$

Có!

- Nếu biến đã quan sát ở giữa sẽ ngăn chặn ảnh hưởng hai biến ngoài cùng

Common cause

- Mô hình common cause



$$P(x,y,z) = P(y)P(x|y)P(z|y)$$

- Liệu có đảm bảo X độc lập với Z ? Không
  - Một ví dụ về các CPT mà X không độc lập với Z là đủ để cho thấy tính độc lập này không được đảm bảo.
  - Ví dụ:
    - Đồ án đến làm các forums bận và phòng lab đầy
    - Bằng số:

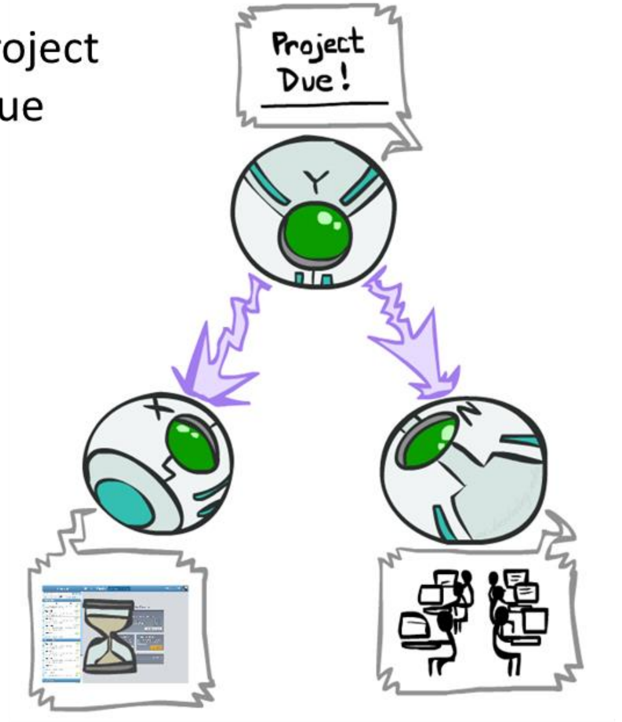
      P(+x | +y) = 1, P(-x | - y) = 1,

      P(+z | +y) = 1, P(-z | -y) = 1

Common cause

- Mô hình common cause

Y: Project due



X: Forums busy

Z: Lab full

$$P(x,y,z) = P(y)P(x|y)P(z|y)$$

- Liệu đảm bảo X độc lập có điều kiện với Z và Y cho trước?

$$P(z|x,y) = \frac{P(x,y,z)}{P(x,y)}$$

$$= \frac{P(x)P(y|x)P(z|y)}{P(x)P(y|x)}$$

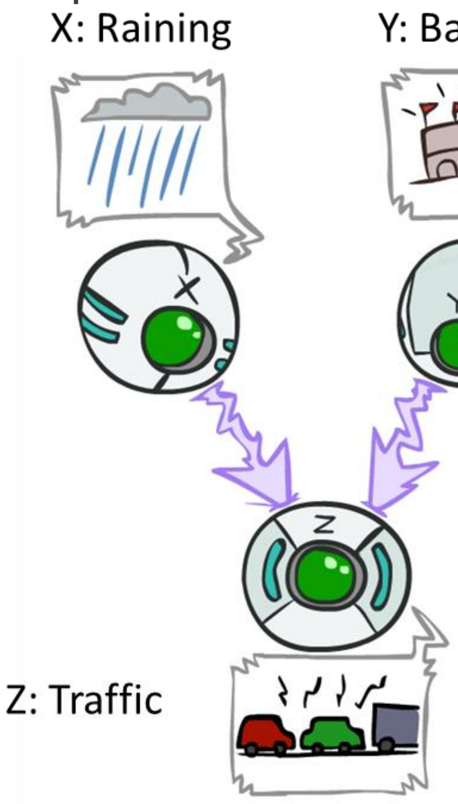
$$= P(z|y)$$

Có!

- Nếu biến nguyên nhân đã quan sát ở giữa sẽ ngăn chặn ảnh hưởng hai biến kết quả ngoài cùng

Common effect

- Mô hình common effect gồm 2 nguyên nhân 1 kết quả

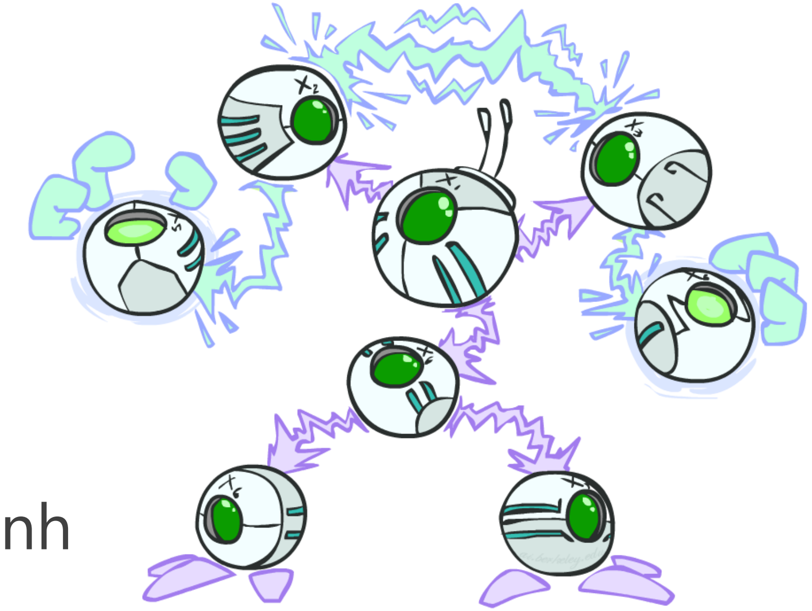


- Liệu X và Y có độc lập?
  - Có: trận bóng và mưa gây kẹt xe, nhưng chúng không tương quan
- X và Y có độc lập có điều kiện?
  - Không: nếu biết Z, thì X và Y không độc lập với nhau.
- Việc quan sát biến kết quả gây ảnh hưởng các biến nguyên nhân

Trường hợp khái quát

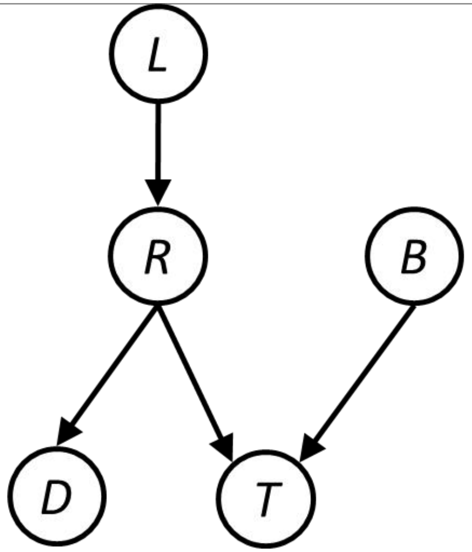
Trường hợp khái quát

- Câu hỏi chung: trong 1 mạng Bayes, liệu 2 biến có độc lập với nhau?
- Cách giải: phân tích đồ thị
- Bất cứ ví dụ phức tạp nào cũng có thể phân rã thành các bộ ba cơ bản



Khả năng tiếp cận

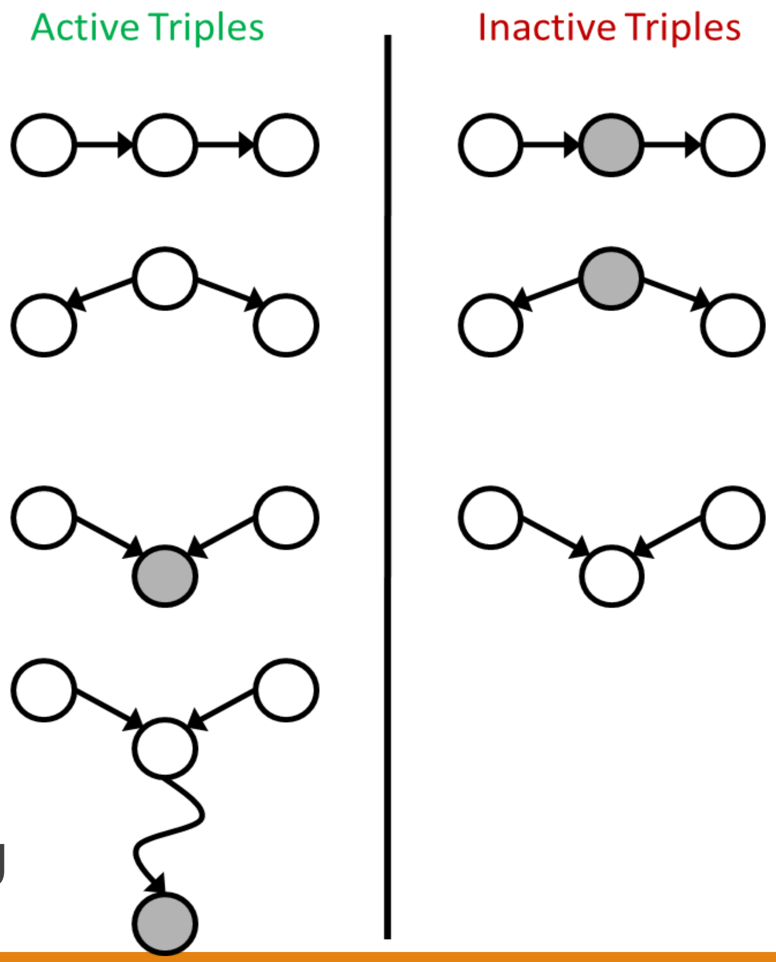
- Nếu 2 node được kết nối gián tiếp và không bị chắn bởi node đã quan sát thì chúng độc lập có điều kiện



- Hầu như hoạt động, nhưng không hoàn toàn
  - Đồ thị bị ngắt chỗ nào?
  - Trả lời: cấu trúc v ở T không phải là liên kết trong đường đi trừ phi "active"

Đường hoạt động/không hoạt động

- Liệu X và Y có độc lập có điều kiện với biến X cho trước?
  - Có, nếu X và Y bị ngắt bởi Z
  - Xét tất cả đường gián tiếp X đến Y
  - Không có đường hoạt động = độc lập
- Một đường gọi là hoạt động nếu mỗi bộ ba con của nó hoạt động:
  - Causal chain A -> B -> C khi B không quan sát
  - Common cause A <- B -> C khi B không quan sát
  - Common effect A -> B <- C khi B hoặc hậu duệ của B được quan sát
- Nếu một đoạn bị không hoạt động thì đường không hoạt động



D-separation

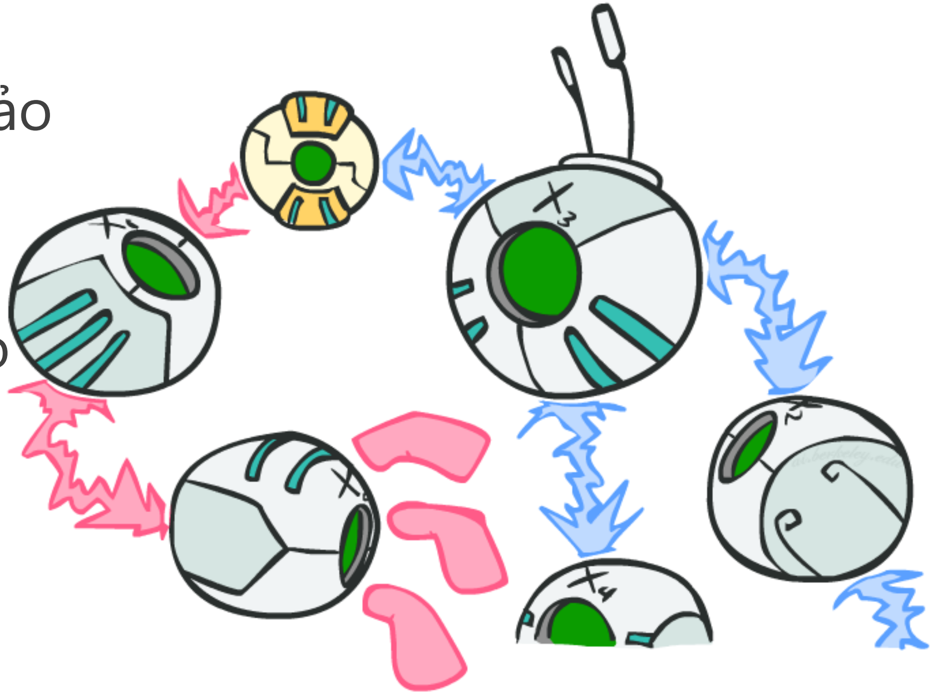
- Truy vấn: $X_i \perp\!\!\!\perp X_j | \{X_{k_1}, ..., X_{k_n}\}$?

- Kiểm tra tất cả đường gián tiếp từ X_i đến X_j
  - Nếu 1 hay nhiều hoạt động, độc lập không đảm bảo

$$X_i \not\perp\!\!\!\perp X_j | \{X_{k_1}, ..., X_{k_n}\}$$

  - Nếu tất cả đều không hoạt động, độc lập đảm bảo

$$X_i \perp\!\!\!\perp X_j | \{X_{k_1}, ..., X_{k_n}\}$$

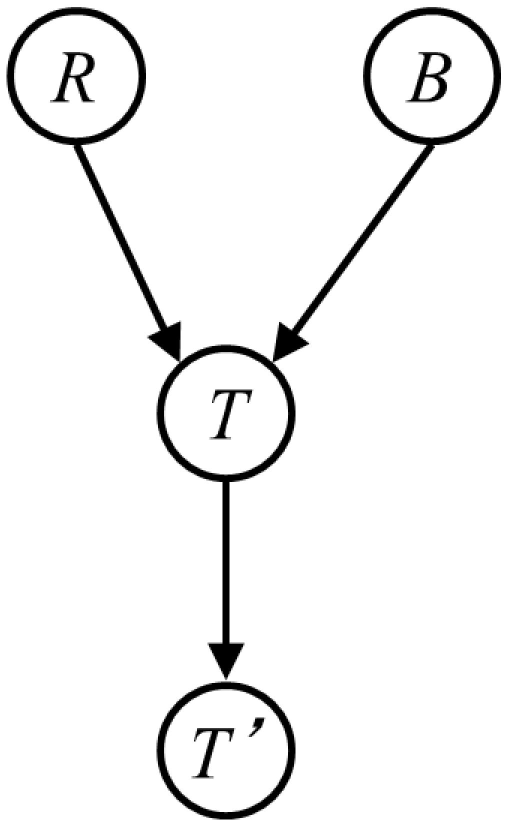


Ví dụ

$R \perp\!\!\!\perp B$ *Yes*

$R \perp\!\!\!\perp B | T$

$R \perp\!\!\!\perp B | T'$



Ví dụ

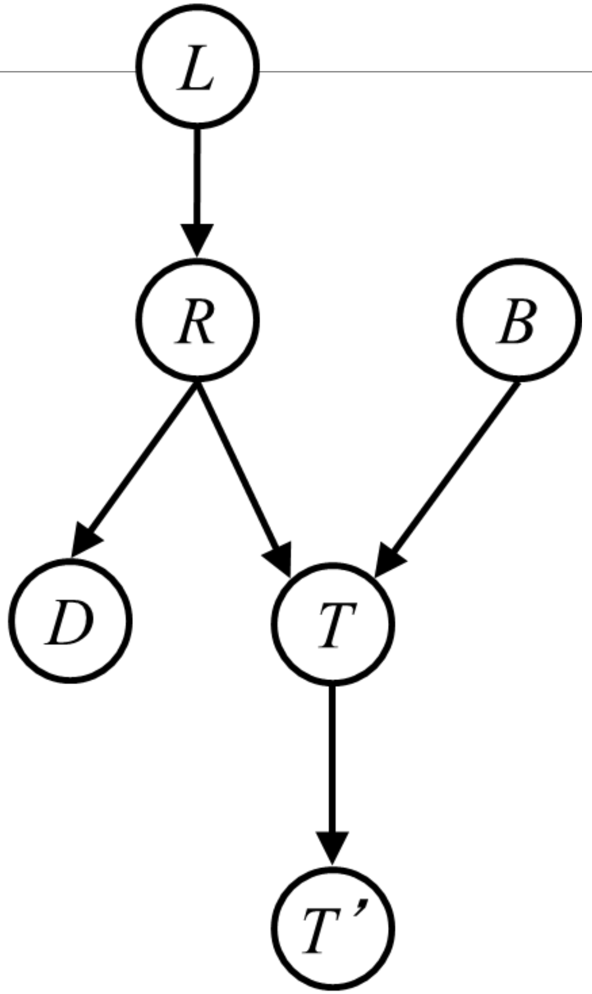
$L \perp\!\!\!\perp T' | T$ *Yes*

$L \perp\!\!\!\perp B$ *Yes*

$L \perp\!\!\!\perp B | T$

$L \perp\!\!\!\perp B | T'$

$L \perp\!\!\!\perp B | T, R$ *Yes*



Ví dụ

Biến:

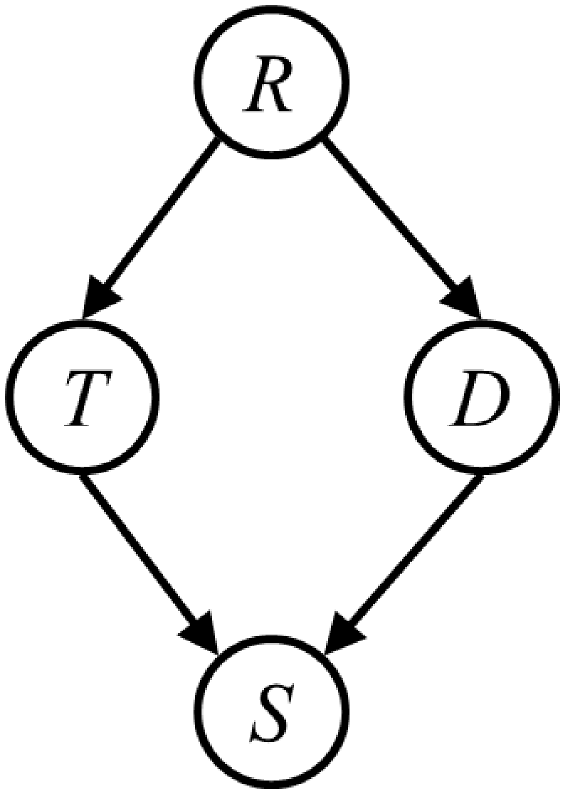
- R: Raining
- T: Traffic
- D: Roof drips
- S: I'm sad

Câu hỏi:

$$T \perp\!\!\!\perp D$$

$$T \perp\!\!\!\perp D | R$$ Yes

$$T \perp\!\!\!\perp D | R, S$$

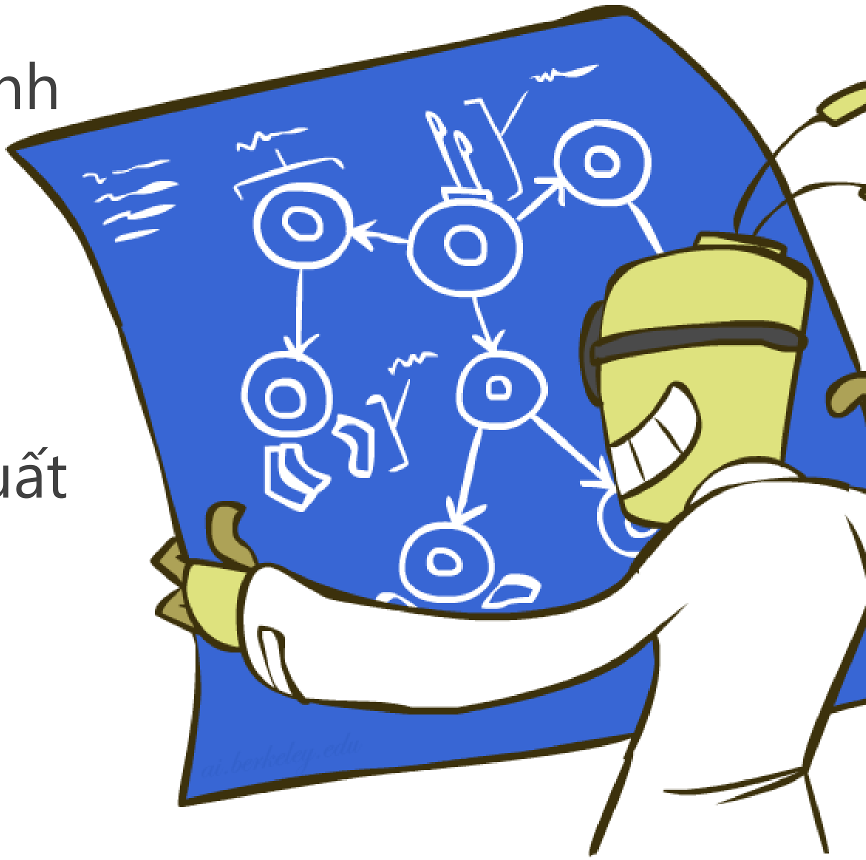


Hàm ý cấu trúc

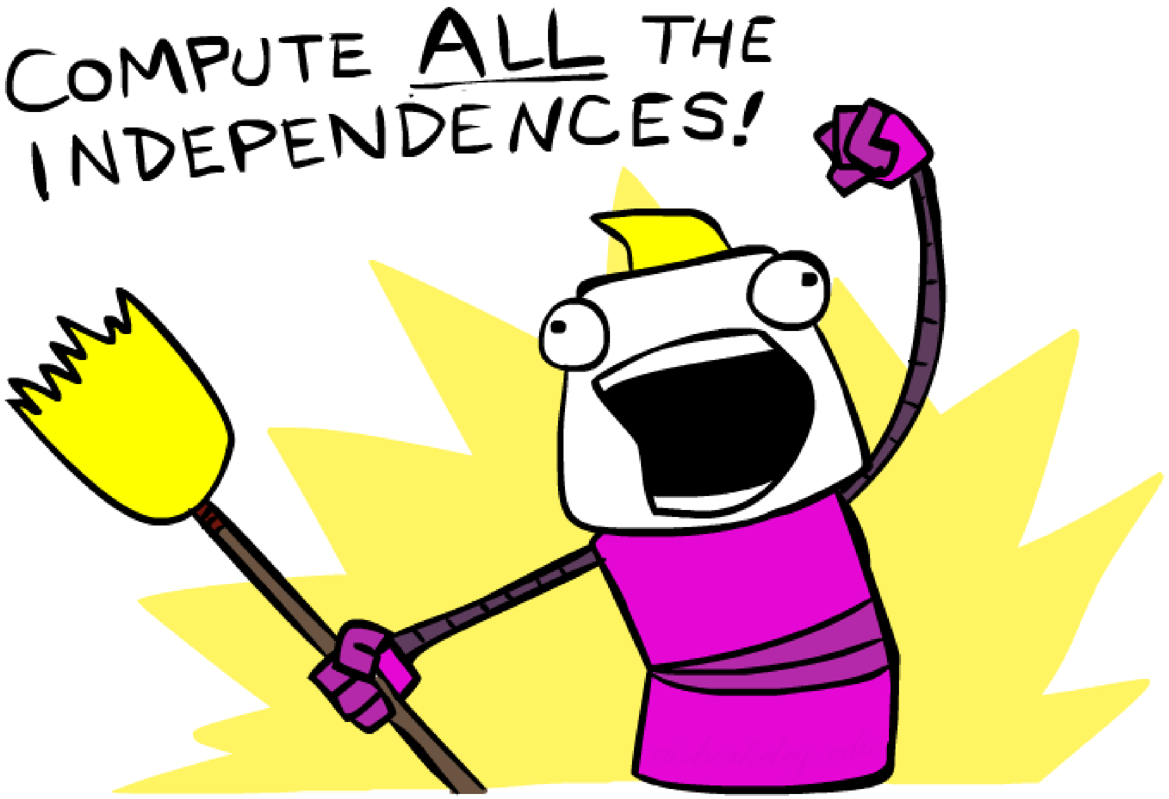
- Cho một cấu trúc mạng Bayes, có thể chạy thuật toán d-separation để xây dựng danh sách hoàn chỉnh độc lập có điều kiện đúng theo kiểu:

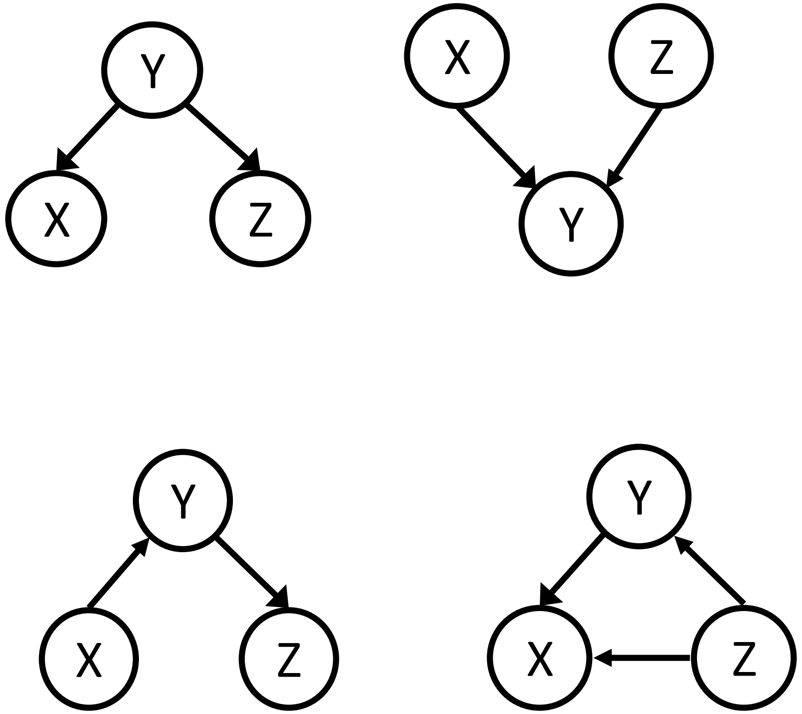
$$X_i \perp\!\!\!\perp X_j | \{X_{k_1}, ..., X_{k_n}\}$$

- Danh sách này quyết định tập các phân phối xác suất có thể thể hiện



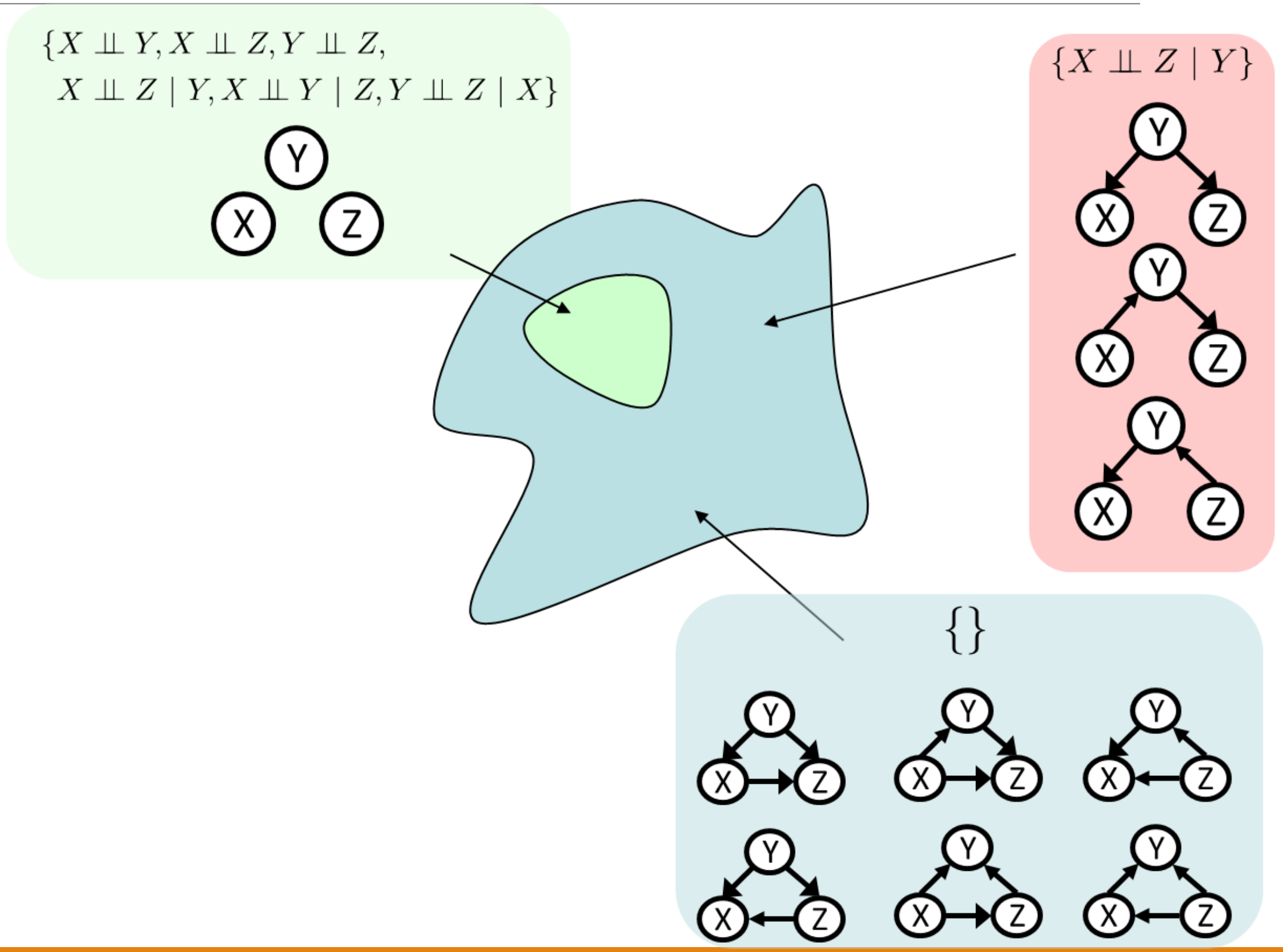
Tìm tất cả độc lập





Phân phối giới hạn liên kết

- Cho một số đồ thị G, chỉ những phân phối chắc chắn chung mới có thể biểu diễn
- Cấu trúc đồ thị đảm bảo các độc lập (có điều kiện) chắc chắn
- Thêm các cung làm tăng tập phân phối nhưng tốn chi phí
- Điều kiện đầy đủ có thể biểu diễn bất cứ phân phối nào



Tổng kết

- Mạng Bayes biểu diễn các phân phối chung nhanh chóng, gọn nhẹ
- Các phân phối đảm bảo độc lập có thể suy luận từ đồ thị mạng Bayes
- D-separation giúp xác định được độc lập có điều kiện từ đồ thị
- Mạng Bayes có thể có nhiều độc lập (có điều kiện) chưa được phát hiện cho đến khi tìm ra được phân phối đặc biệt của nó

Nguồn tham khảo

Slide bài giảng tiếng Anh Bayesian Networks I & II