Chuyển hóa quan hệ đồ thị thành xác suất có điều kiện trong mạng Bayesian

GS. Nguyễn Phước Lộc PhD, MD, MBA

Đại học An Giang, Việt Nam

Email: ng_phloc@yahoo.com

Homepage: www.locnguyen.net

Tóm tắt

Mạng Bayesian - Bayesian network (BN) là một công cụ toán học mạnh dùng cho những ứng dụng tiên đoán và chẩn đoán. Nhiều mạng nhỏ hợp thành mạng lớn. Mỗi mạng nhỏ được tạo từ một đồ thị đơn giản, gồm một nút con và nhiều nút cha. Quan hệ giữa nút con và nút cha được lượng hóa bằng một trọng số. Mỗi quan hệ đều có một ngữ nghĩa như quan hệ điều kiện, chẩn đoán, kết hợp. Nghiên cứu tập trung vào *chuyển hóa quan hệ* đồ thị thành xác suất có điều kiện để xây dựng mạng BN từ đồ thị. *Quan hệ chẩn đoán* là chủ đề chính, trong đó tôi đề xuất *mệnh đề chẩn đoán đủ*, nhằm xác nhận tính hợp lệ của quan hệ chẩn đoán. Sự biến đổi quan hệ mô phỏng theo những cổng lô-gic như AND, OR và XOR, là đặc trưng của nghiên cứu.

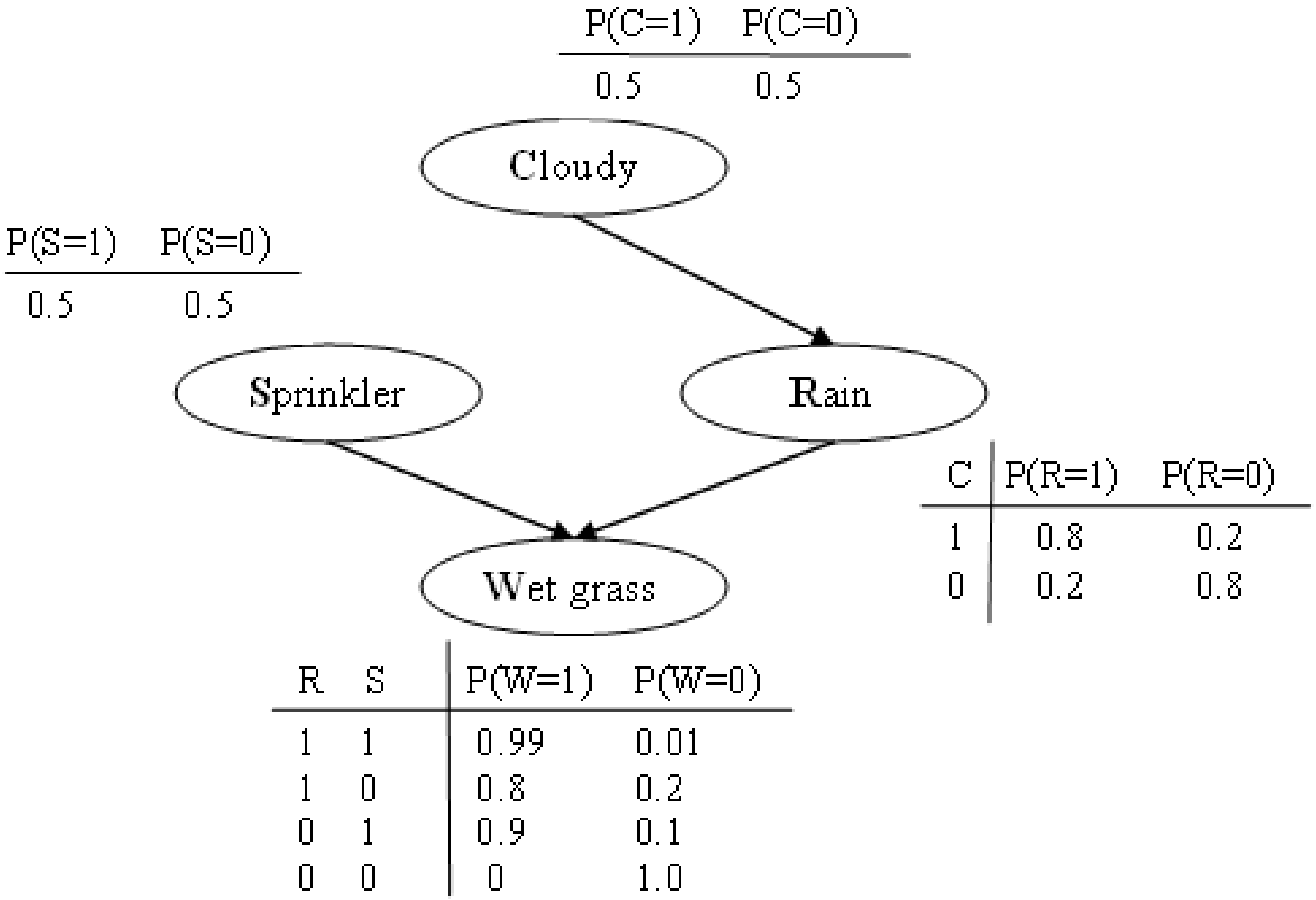
Table of contents

1. Giới thiệu

- Mạng Bayesian (BN) là đồ thị có hướng không vòng (DAG) gồm một tập nút và một tập cạnh. Mỗi nút là biến ngẫu nhiên. Mỗi cạnh biểu diễn quan hệ giữa hai nút.
- Mỗi quan hệ có một *trọng số* (weight). Những quan hệ quan trọng bao gồm điều kiện tiên quyết, chẩn đoán và kết hợp.
- Sự khác biệt giữa BN và đồ thị bình thường là trọng số của quan hệ trong BN được thay bằng bảng xác suất điều kiện (CPT). Mỗi mục của CPT là xác suất điều kiện của nút con với nút cha.
- Có hai cách tiếp cận để xây dựng một BN:
    - Cách thứ nhất: huấn luyện BN từ dữ liệu bằng thuật toán học máy.
    - Cách thứ hai: các chuyên gia định nghĩa một vài mẫu đồ thị với quan hệ đặc thù và sau đó, BN được xây dựng dựa vào những mẫu này cùng với các bảng CPT. **Đây là hướng tiếp cận của nghiên cứu này**.

1. Giới thiệu

- **Sau đây là một ví dụ BN**, sự kiện "cloudy" là nguyên nhân của sự kiện "rain" tiếp theo là nguyên nhân của sự kiện "grass is wet" (Murphy, 1998).
- Các biến ngẫu nhiên nhị phân (là những nút) *C*, *S*, *R*, và *W* biểu diễn các sự kiện "cloudy", "sprinkler", "rain" và "wet grass". Mỗi nút kèm theo một bảng CPT.
- Với bằng chứng "wet grass" W=1, ta có kết luận rằng "sprinkler" là nguyên nhân chính vì $P(S=1|W=1) = 0.7 > P(R=1|W=1) = 0.67$.



1. Giới thiệu

- *Chuyển hóa quan hệ* (relationship conversion) là xác định những xác suất điều kiện dựa trên trọng số và ngữ nghĩa của quan hệ, đặc biệt tập trung vào các quan hệ lô-gic X-gate (Wikipedia, 2016) như AND-gate, OR-gate và SIGMA-gate. Trong nghiên cứu này, suy diễn X-gate được kế thừa và cảm hứng từ nhiễu OR-gate trong cuốn sách “Learning Bayesian Networks” của tác giả (Neapolitan, 2003, pp. 157-159).
- Các tác giả (Díez & Druzdzel, 2007) cũng nghiên cứu các suy diễn OR/MAX, AND/MIN, nhiễu XOR nhưng họ tập trung vào mô hình chuẩn tắc, mô hình xác định, mô hình ICI trong khi tôi tập trung vào cổng lô-gic và quan hệ đồ thị.
- Đồ thị thừa số (Wikipedia, Factor graph, 2015) biểu diễn chuỗi thừa số của một hàm toàn cục thành nhiều hàm cục bộ. Thuật toán lan truyền của Pearl (Pearl, 1986) là một biến thể rất thành công trong suy diễn BN. Tuy nhiên tôi không sử dụng đồ thị thừa số cho việc xây dựng BN. Khái niệm “suy diễn X-gate” ngụ ý chuyển đổi đồ thị đơn giản thành BN.

1. Giới thiệu

- Nghiên cứu của tôi tập trung vào xác suất ứng dụng gắn liền với BN, cổng lô-gic và mô hình hóa mạng Bayesian người dùng (Millán & Pérez-de-la-Cruz, 2002). Những kết quả nghiên cứu chia sẻ với hai tác giả **Eva Millán** và **José Luis Pérez-de-la-Cruz**.
- Ứng dụng mặc định của nghiên cứu là đánh giá kết quả học tập của người học trong môi trường học điện tử.
- Bằng chứng (evidence) là những bài kiểm tra, bài tập còn giả thuyết (hypothesis) là những khái niệm, kiến thức.
- *Quan hệ chẩn đoán* rất quan trọng vì nó được dùng để đánh giá kiến thức người học. Trong phần tiếp theo, chúng ta bắt đầu nghiên cứu sự chuyển hóa quan hệ với khái niệm quan hệ chẩn đoán.

2. Quan hệ chẩn đoán

- Theo (Millán, Loboda, & Pérez-de-la-Cruz, 2010, p. 1666) và tôi, quan hệ chẩn đoán nên bắt đầu từ giả thuyết đến bằng chứng. Ví dụ, bệnh là giả thuyết *X* và triệu chứng là bằng chứng *D*.

$$X \xrightarrow{1} D$$

- Trọng số *w* của quan hệ giữa *X* và *D* là 1. Công thức 2.1 đặc tả CPT của biến nhị phân *D* nhằm chuyển hóa quan hệ chẩn đoán đơn giản nhất thành BN đơn giản nhất.

$$P(D|X) = \begin{cases} D \text{ if } X = 1 \\ 1 - D \text{ if } X = 0 \end{cases} \quad (2.1)$$

- "Bằng chứng *D* được dùng để chẩn đoán giả thuyết *X* nếu *mệnh đề chẩn đoán đủ* được thỏa mãn. Mệnh đề này, gọi ngắn gọn là **điều kiện chẩn đoán**, được phát biểu như sau "*D tương đương với X trong quan hệ chẩn đoán nếu P(X|D) = kP(D|X) với giả thiết rằng X phân bố đều và* ***hệ số chuyển hóa*** *k độc lập với D. Nói cách khác, k là hằng số với D và vì thế D được gọi là bằng chứng đủ*".

2. Quan hệ chẩn đoán

- Tôi khảo sát ba trường hợp phổ biến khác của bằng chứng D như $D \in \{0, 1, 2, \ldots, \eta\}$, $D \in \{a, a+1, a+2, \ldots, b\}$ và $D \in [a, b]$.
- Công thức 2.6 tổng kết CPT của bằng chứng trong quan hệ chẩn đoán đơn, thỏa mãn **điều kiện chẩn đoán**.

$$P(D|X) = \begin{cases} \dfrac{D}{S} \text{ if } X = 1 \\ \dfrac{M}{S} - \dfrac{D}{S} \text{ if } X = 0 \end{cases} \quad (2.6)$$

$$k = \frac{N}{2}$$

Với

$$N = \begin{cases} 2 \text{ if } D \in \{0,1\} \\ \eta + 1 \text{ if } D \in \{0,1,2,\ldots,\eta\} \\ b - a + 1 \text{ if } D \in \{a, a+1, a+2, \ldots, b\} \\ b - a \text{ if } D \text{ continuous and } D \in [a, b] \end{cases} \quad M = \begin{cases} 1 \text{ if } D \in \{0,1\} \\ \eta \text{ if } D \in \{0,1,2,\ldots,\eta\} \\ b + a \text{ if } D \in \{a, a+1, a+2, \ldots, b\} \\ b + a \text{ if } D \text{ continuous and } D \in [a, b] \end{cases}$$

$$S = \sum_D D = \frac{NM}{2} = \begin{cases} 1 \text{ if } D \in \{0,1\} \\ \dfrac{\eta(\eta+1)}{2} \text{ if } D \in \{0,1,2,\ldots,\eta\} \\ \dfrac{(b+a)(b-a+1)}{2} \text{ if } D \in \{a, a+1, a+2, \ldots, b\} \\ \dfrac{b^2 - a^2}{2} \text{ if } D \text{ continuous and } D \in [a, b] \end{cases}$$

2. Quan hệ chẩn đoán

Chúng ta thử ví dụ chứng minh một trường hợp đặc biệt của công thức 2.6, CPT của bằng chứng $D \in \{0, 1, 2, \ldots, \eta\}$, thỏa mãn điều kiện chẩn đoán.

$$P(D|X) = \begin{cases} \frac{D}{S} \text{ if } X = 1 \\ \frac{\eta}{S} - \frac{D}{S} \text{ if } X = 0 \end{cases} \text{ where } D \in \{0,1,2,\ldots,\eta\} \text{ and } S = \sum_{D=0}^{\eta} D = \frac{\eta(\eta+1)}{2}$$

Thật vậy, ta có:

$$P(D|X=0) + P(D|X=1) = \frac{D}{S} + \frac{\eta - D}{S} = \frac{2}{(\eta+1)}$$

$$\sum_{D=0}^{\eta} P(D|X=1) = \sum_{D=0}^{\eta} \frac{D}{S} = \frac{\sum_{D=0}^{\eta} D}{S} = \frac{S}{S} = 1$$

$$\sum_{D=0}^{\eta} P(D|X=0) = \sum_{D=0}^{\eta} \frac{\eta - D}{S} = \frac{\eta(\eta+1) - S}{S} = 1$$

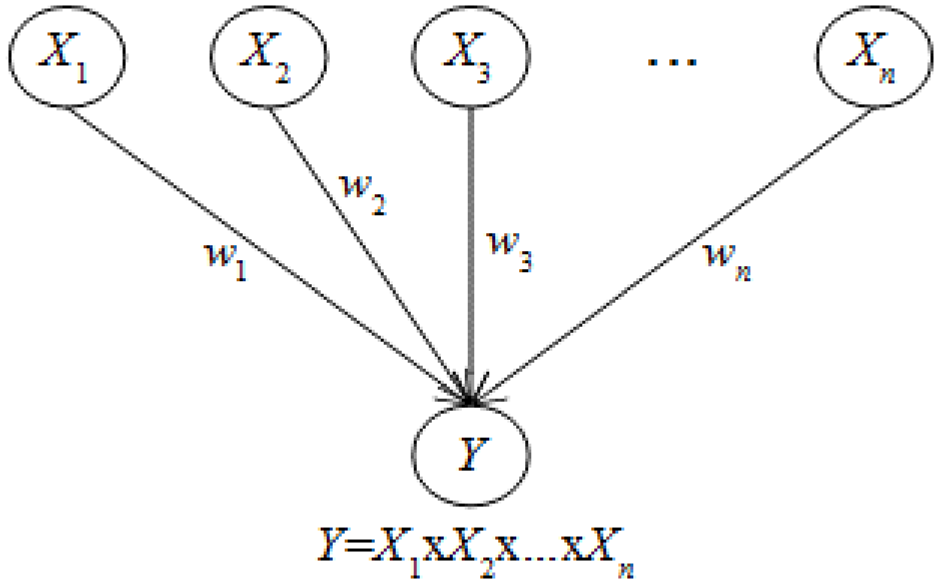
Giả sử *X* phân bố đều: $P(X=0) = P(X=1)$. We have:

$$P(X|D) = \frac{P(D|X)P(X)}{P(D)} = \frac{P(D|X)}{P(D|X=0) + P(D|X=1)} = \frac{\eta+1}{2} P(D|X)$$

Vậy hệ số chuyển hóa *k* là $\frac{\eta+1}{2}$.

3. Suy diễn X-gate

- Mở rộng quan hệ chẩn đoán với nhiều giả thuyết, trong đó đồ thị đơn giản gồm một biến con *Y* và *n* biến cha X_i. Mỗi quan hệ từ X_i đến *Y* được lượng hóa bởi trọng số w_i với $0 \le w_i \le 1$.
- Những quan hệ đồ thị gắn liền với các cổng lô-gic X-gate như AND-gate, OR-gate và SIGMA-gate, được chuyển thành CPT của BN đơn giản. Sự chuyển hóa quan hệ xác định *suy diễn X-gate*. Đồ thị đơn giản được gọi *đồ thị X-gate* hay *mạng X-gate*.

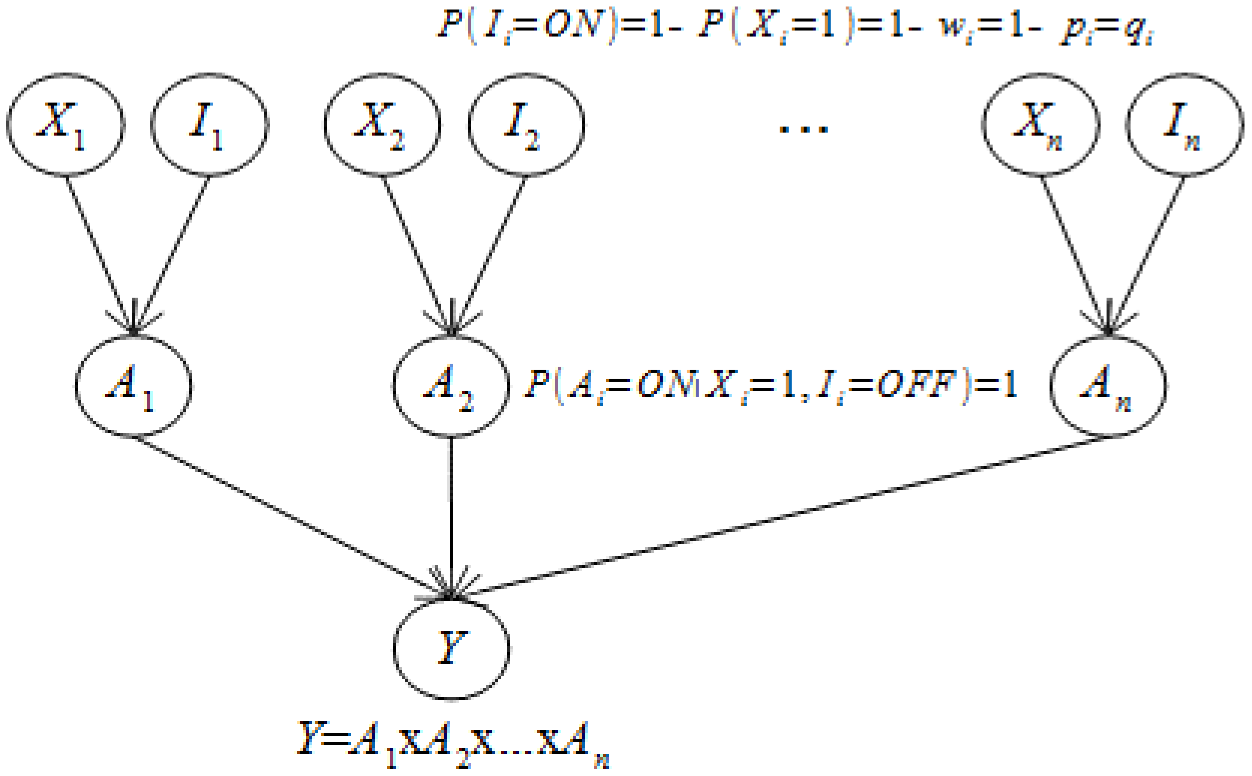


Mạng X-gate

3. Suy diễn X-gate

Suy diễn X-gate dựa trên ba giả định sau được đề cập trong (Neapolitan, 2003, p. 157)

- *Ức chế X-gate*: Với mỗi quan hệ từ nguồn X_i đến đích Y, có một biến ức chế I_i ngăn X_i tích hợp vào Y.
- *Độc lập ức chế*: tất cả biến I_i đều độc lập lẫn nhau.
- *Sự đáp ứng*: Mạng X-gate được thiết lập bởi các biến đáp ứng A_i cho X_i và I_i. Mỗi suy diễn X-gate là sự kết hợp đặc thù của những A_i.



Mạng X-gate mở rộng với biến đáp ứng

3. Suy diễn X-gate

Xác suất của biến ức chế I_i và biến đáp ứng A_i với các công thức 3.2, 3.3

$$P(I_i = OFF) = p_i = w_i$$
$$P(I_i = ON) = 1 - p_i = 1 - w_i$$
$$P(A_i = ON|X_i = 1, I_i = OFF) = 1$$
$$P(A_i = ON|X_i = 1, I_i = ON) = 0$$
$$P(A_i = ON|X_i = 0, I_i = OFF) = 0$$
$$P(A_i = ON|X_i = 0, I_i = ON) = 0$$
$$P(A_i = OFF|X_i = 1, I_i = OFF) = 0$$
$$P(A_i = OFF|X_i = 1, I_i = ON) = 1$$
$$P(A_i = OFF|X_i = 0, I_i = OFF) = 1$$
$$P(A_i = OFF|X_i = 0, I_i = ON) = 1$$

$$P(A_i = ON|X_i = 1) = p_i = w_i$$
$$P(A_i = ON|X_i = 0) = 0$$
$$P(A_i = OFF|X_i = 1) = 1 - p_i = 1 - w_i$$
$$P(A_i = OFF|X_i = 0) = 1$$

3. Suy diễn X-gate

$$P(Y|X_1, X_2, \dots, X_n) = \frac{P(Y, X_1, X_2, \dots, X_n)}{P(X_1, X_2, \dots, X_n)}$$

(do luật Bayes)

$$= \frac{\sum_{A_1,A_2,\dots,A_n} P(Y, X_1, X_2, \dots, X_n | A_1, A_2, \dots, A_n) * P(A_1, A_2, \dots, A_n)}{P(X_1, X_2, \dots, X_n)}$$

(do luật xác suất tổng)

$$= \sum_{A_1,A_2,\dots,A_n} P(Y, X_1, X_2, \dots, X_n | A_1, A_2, \dots, A_n) * \frac{P(A_1, A_2, \dots, A_n)}{P(X_1, X_2, \dots, X_n)}$$

$$= \sum_{A_1,A_2,\dots,A_n} P(Y | A_1, A_2, \dots, A_n) * P(X_1, X_2, \dots, X_n | A_1, A_2, \dots, A_n) * \frac{P(A_1, A_2, \dots, A_n)}{P(X_1, X_2, \dots, X_n)}$$

(do Y độc lập điều kiện với X_i cho A_i)

$$= \sum_{A_1,A_2,\dots,A_n} P(Y | A_1, A_2, \dots, A_n) * \frac{P(X_1, X_2, \dots, X_n, A_1, A_2, \dots, A_n)}{P(X_1, X_2, \dots, X_n)}$$

$$= \sum_{A_1,A_2,\dots,A_n} P(Y | A_1, A_2, \dots, A_n) * P(A_1, A_2, \dots, A_n | X_1, X_2, \dots, X_n)$$

(do luật Bayes)

$$= \sum_{A_1,A_2,\dots,A_n} P(Y | A_1, A_2, \dots, A_n) \prod_{i=1}^{n} P(A_i | X_1, X_2, \dots, X_n)$$

(do các A_i độc lập lẫn nhau)

$$= \sum_{A_1,A_2,\dots,A_n} P(Y | A_1, A_2, \dots, A_n) \prod_{i=1}^{n} P(A_i | X_i)$$

(do mỗi A_i chỉ phụ thuộc vào X_i)

$$\boxed{P(Y|X_1, X_2, \dots, X_n) = \sum_{A_1,A_2,\dots,A_n} P(Y | A_1, A_2, \dots, A_n) \prod_{i=1}^{n} P(A_i | X_i)}$$

Xác suất X-gate (3.4)

Suy diễn X-gate được đặc tả bởi xác suất X-gate $P(Y=1 \mid X_1, X_2, \dots, X_n)$ từ (Neapolitan, 2003, p. 159)

3. Suy diễn X-gate

- Cho Ω = {X_1, X_2,…, X_n} với |Ω|=*n* là lực lượng của Ω.
- Đặt *a*(Ω) là một chỉnh hợp của, tạo từ *n* thể hiện {X_1=x_1, X_2=x_2,…, X_n=x_n} với x_i là 1 hoặc 0. Số lượng tất cả *a*(Ω) là $2^{|\Omega|}$. Ví dụ, với Ω = {X_1, X_2}, có 2^2=4 chỉnh hợp như sau: *a*(Ω) = {X_1=1, X_2=1}, *a*(Ω) = {X_1=1, X_2=0}, *a*(Ω) = {X_1=0, X_2=1}, *a*(Ω) = {X_1=0, X_2=0}. Đặt *a*(Ω:{X_i}) là một chỉnh hợp của Ω với X_i cố định.
- Đặt *c*(Ω) và *c*(Ω:{X_i}) tương ứng là số lượng các chỉnh hợp *a*(Ω) and *a*(Ω:{X_i}).
- Toán tử X-gate được ký hiệu là x. Ví dụ, $\mathrm{x} = \odot$ cho AND-gate, $\mathrm{x} = \oplus$ cho OR-gate, $\mathrm{x} = \mathrm{not}\odot$ cho NAND-gate, $\mathrm{x} = \mathrm{not}\oplus$ cho NOR-gate, $\mathrm{x} = \otimes$ cho XOR-gate, $\mathrm{x} = \mathrm{not}\otimes$ cho XNOR-gate, $\mathrm{x} = \uplus$ cho U-gate, $\mathrm{x} = +$ cho SIGMA-gate. Cho toán tử x, đặt *s*(Ω:{X_i}) tương ứng *s*(Ω) là tổng tất cả $P(X_1 \mathrm{x} X_2 \mathrm{x} \ldots \mathrm{x} X_n)$ duyệt qua mọi chỉnh hợp của Ω với X_i tự do và X_i cố định.

$$s(\Omega) = \sum_a P(X_1 \oplus X_2 \oplus \ldots \oplus X_n | a(\Omega))$$

$$s(\Omega:\{X_i\}) = \sum_a P(X_1 \oplus X_2 \oplus \ldots \oplus X_n | a(\Omega:\{X_i\}))$$

3. Suy diễn X-gate

```
public class ArrangementGenerator {
    private ArrayList<int[]> arrangements; private int n, r;
    private ArrangementGenerator(int n, int r) {
        this.n = n; this.r = r; this.arrangements = new ArrayList();
    }

    private void create(int[] a, int i) {
        for(int j = 0; j < n; j++) {
            a[i] = j;
            if(i < r - 1) create(a, i + 1);
            else if(i == r -1) {
                int[] b = new int[a.length];
                for(int k = 0; k < a.length; k++) b[k] = a[k];
                arrangements.add(b);
            }
        }
    }

}
```

mã lập trình phát sinh tất cả chỉnh hợp

3. Suy diễn X-gate

- Mối liên hệ giữa $s(\Omega:\{X_i=1\})$ và $s(\Omega:\{X_i=0\})$, giữa $c(\Omega:\{X_i=1\})$ và $c(\Omega:\{X_i=0\})$.

$$s(\Omega:\{X_i = 1\}) + s(\Omega:\{X_i = 0\}) = s(\Omega)$$
$$c(\Omega:\{X_i = 1\}) + c(\Omega:\{X_i = 0\}) = c(\Omega)$$

- Đặt K là tập hợp các X_i có giá trị 1 và let L là tập hợp các X_i có giá trị 0. K và L bù lẫn nhau.

$$\begin{cases} K = \{i: X_i = 1\} \\ L = \{i: X_i = 0\} \\ K \cap L = \emptyset \\ K \cup L = \{1,2,\dots,n\} \end{cases}$$

3. Suy diễn X-gate

- Điều kiện AND-gate (3.7)

$$P(Y = 1 | A_i = OFF \text{ for some } i) = 0$$

- **Suy diễn AND-gate** (3.8)

$$P(X_1 \odot X_2 \odot \dots \odot X_n) = P(Y = 1 | X_1, X_2, \dots, X_n)$$

$$= \begin{cases} \prod_{i=1}^{n} p_i \text{ if all } X_i \ (s) \text{ are } 1 \\ 0 \text{ if there exists at least one } X_i = 0 \end{cases}$$

$$P(Y = 0 | X_1, X_2, \dots, X_n) = \begin{cases} 1 - \prod_{i=1}^{n} p_i \text{ if all } X_i \ (s) \text{ are } 1 \\ 1 \text{ if there exists at least one } X_i = 0 \end{cases}$$

3. Suy diễn X-gate

- Điều kiện OR-gate (3.9)

$$P(Y=1|A_i=ON \text{ for some } i)=1$$

- **Suy diễn OR-gate** (3.10)

$$P(X_1 \oplus X_2 \oplus \ldots \oplus X_n) = 1 - P(Y=0|X_1, X_2, \ldots, X_n)$$

$$= \begin{cases} 1 - \prod_{i \in K} (1-p_i) \text{ if } K \neq \emptyset \\ 0 \text{ if } K = \emptyset \end{cases}$$

$$P(Y=0|X_1, X_2, \ldots, X_n) = \begin{cases} \prod_{i \in K} (1-p_i) \text{ if } K \neq \emptyset \\ 1 \text{ if } K = \emptyset \end{cases}$$

3. Suy diễn X-gate

- **Suy diễn NAND-gate** and **suy diễn NOR-gate** (3.11)

$$P\big(\text{not}(X_1 \odot X_2 \odot \ldots \odot X_n)\big) = \begin{cases} 1 - \prod_{i \in L} p_i \ \text{ if } L \neq \emptyset \\ 0 \ \text{ if } L = \emptyset \end{cases}$$

$$P\big(\text{not}(X_1 \oplus X_2 \oplus \ldots \oplus X_n)\big) = \begin{cases} \prod_{i=1}^{n} q_i \ \text{ if } K = \emptyset \\ 0 \text{ if } K \neq \emptyset \end{cases}$$

3. Suy diễn X-gate

- Hai điều kiện XOR-gate (3.12)

$$P\left(Y=1 \middle| \begin{cases} A_i = ON \text{ for } i \in O \\ A_i = OFF \text{ for } i \notin O \end{cases}\right) = P(Y=1|A_1=ON, A_2=OFF, \dots, A_{n-1}=ON, A_n=OFF) = 1$$

$$P\left(Y=1 \middle| \begin{cases} A_i = ON \text{ for } i \in E \\ A_i = OFF \text{ for } i \notin E \end{cases}\right) = P(Y=1|A_1=OFF, A_2=ON, \dots, A_{n-1}=OFF, A_n=ON) = 1$$

- Đặt O là tập những X_i có chỉ mục lẻ. Đặt O_1 và O_2 là hai tập con của O, trong đó X_i lần lượt có giá trị 1 và 0. Đặt E là tập những X_i có chỉ mục chẵn. Đặt E_1 và E_2 là hai tập con của E, trong đó X_i lần lượt có giá trị 1 và 0. **Suy diễn XOR-gate** (3.13) là:

$$P(X_1 \otimes X_2 \otimes \dots \otimes X_n) = P(Y=1|X_1, X_2, \dots, X_n) = \begin{cases} \left(\prod_{i\in O_1} p_i\right)\left(\prod_{i\in E_1}(1-p_i)\right) + \left(\prod_{i\in E_1} p_i\right)\left(\prod_{i\in O_1}(1-p_i)\right) \text{ if } O_2 = \emptyset \text{ and } E_2 = \emptyset \\ \left(\prod_{i\in O_1} p_i\right)\left(\prod_{i\in E_1}(1-p_i)\right) \text{if } O_2 = \emptyset \text{ and } E_1 \neq \emptyset \text{ and } E_2 \neq \emptyset \\ \prod_{i\in O_1} p_i \text{ if } O_2 = \emptyset \text{ and } E_1 = \emptyset \\ \left(\prod_{i\in E_1} p_i\right)\left(\prod_{i\in O_1}(1-p_i)\right) \text{ if } E_2 = \emptyset \text{ and } O_1 \neq \emptyset \text{ and } O_2 \neq \emptyset \\ \prod_{i\in E_1} p_i \text{ if } E_2 = \emptyset \text{ and } O_1 = \emptyset \\ 0 \text{ if } O_2 \neq \emptyset \text{ and } E_2 \neq \emptyset \\ 0 \text{ if } n < 2 \text{ or } n \text{ is odd} \end{cases}$$

3. Suy diễn X-gate

- Hai điều kiện XNOR-gate (3.14)

$$P(Y=1|A_i=ON,\forall i)=1$$
$$P(Y=1|A_i=OFF,\forall i)=1$$

- **Suy diễn XNOR-gate** (3.15)

$$P\big(\text{not}(X_1 \otimes X_2 \otimes \ldots \otimes X_n)\big)=P(Y=1|X_1,X_2,\ldots,X_n)$$
$$=\begin{cases}\prod_{i=1}^{n} p_i+\prod_{i=1}^{n}(1-p_i) \text{ if } L=\emptyset \\ \prod_{i\in K}(1-p_i) \text{ if } L\neq\emptyset \text{ and } K\neq\emptyset \\ 1 \text{ if } L\neq\emptyset \text{ and } K=\emptyset\end{cases}$$

3. Suy diễn X-gate

Đặt U là tập các chỉ mục sao cho A_i=ON và đặt $\alpha \geq 0$ và $\beta \geq 0$ là những số được định nghĩa trước. **Suy diễn U-gate** được định nghĩa trên α, β và lực lượng của U. Công thức 3.16 đặc tả những **điều kiện U-gate** phổ biến.

$\boldsymbol{\lvert U\rvert=\alpha}$	$P(Y=1\mid A_1,A_2,\ldots,A_n)=1$ nếu có chính xác α biến A_i = ON (s). Ngược lại, $P(Y=1\mid A_1,A_2,\ldots,A_n)=0$.
$\boldsymbol{\lvert U\rvert\geq\alpha}$	$P(Y=1\mid A_1,A_2,\ldots,A_n)=1$ nếu có ít nhất α variables A_i = ON (s). Ngược lại, $P(Y=1\mid A_1,A_2,\ldots,A_n)=0$.
$\boldsymbol{\lvert U\rvert\leq\beta}$	$P(Y=1\mid A_1,A_2,\ldots,A_n)=1$ nếu có nhiều nhất β variables A_i = ON (s). Ngược lại, $P(Y=1\mid A_1,A_2,\ldots,A_n)=0$.
$\boldsymbol{\alpha\leq\lvert U\rvert\leq\beta}$	$P(Y=1\mid A_1,A_2,\ldots,A_n)=1$ nếu số lượng A_i = ON (s) từ α đến β. Ngược lại, $P(Y=1\mid A_1,A_2,\ldots,A_n)=0$.

3. Suy diễn X-gate

Đặt P_U là xác suất U-gate, công thức 3.17 sau đặc tả **suy diễn U-gate** và lực lượng của $\mathcal{U}$ với $\mathcal{U}$ là tập hợp tất cả tập con (U) của K

- Đặt $S_U = \sum_{U \in \mathcal{U}} \prod_{i \in U} p_i \prod_{j \in K \backslash U} (1 - p_j)$ và $P_U = P(X_1 \uplus X_2 \uplus \ldots \uplus X_n) = P(Y = 1 | X_1, X_2, \ldots, X_n)$
- Theo quy ước, $\prod_{i \in U} p_i = 1$ if $|U| = 0$ and $\prod_{j \in K \backslash U} (1 - p_j) = 1$ if $|U| = |K|$
- $|U|$=0: ta có $P_U = \begin{cases} \prod_{j=1}^{n} (1 - p_j) \text{ if } |K| > 0 \\ 1 \text{ if } |K| = 0 \end{cases}$ and $|\mathcal{U}| = 1$
- $|U|$≥0: ta có $P_U = \begin{cases} S_U \text{ if } |K| > 0 \\ 1 \text{ if } |K| = 0 \end{cases}$ and $|\mathcal{U}| = 2^{|K|}$. Trường hợp $|U|$≥0 giống như trường hợp $|U|$≤n.
- $|U|$=n: ta có $P_U = \begin{cases} \prod_{i=1}^{n} p_i \text{ if } |K| = n \\ 0 \text{ if } |K| < n \end{cases}$ and $|\mathcal{U}| = \begin{cases} 1 \text{ if } |K| = n \\ 0 \text{ if } |K| < n \end{cases}$

3. Suy diễn X-gate

Suy diễn U-gate (tiếp tục)

- $|U|=\alpha$ and $0<\alpha<n$: ta có $P_U = \begin{cases} S_U \text{ if } |K| \geq \alpha \\ 0 \text{ if } |K| < \alpha \end{cases}$ and $|\mathcal{U}| = \begin{cases} \binom{|K|}{\alpha} \text{ if } |K| \geq \alpha \\ 0 \text{ if } |K| < \alpha \end{cases}$
- $|U|\geq\alpha$ and $0<\alpha<n$: ta có $P_U = \begin{cases} S_U \text{ if } |K| \geq \alpha \\ 0 \text{ if } |K| < \alpha \end{cases}$ and $|\mathcal{U}| = \begin{cases} \sum_{j=\alpha}^{|K|} \binom{|K|}{j} \text{ if } |K| \geq \alpha \\ 0 \text{ if } |K| < \alpha \end{cases}$
- $|U|\leq\beta$ and $0<\beta<n$: ta có $P_U = \begin{cases} S_U \text{ if } |K| > 0 \\ 1 \text{ if } |K| = 0 \end{cases}$ and $|\mathcal{U}| = \begin{cases} \sum_{j=0}^{\min(\beta,|K|)} \binom{|K|}{j} \text{ if } |K| > 0 \\ 1 \text{ if } |K| = 0 \end{cases}$
- $\alpha\leq|U|\leq\beta$ and $0<\alpha<n$ and $0<\beta<n$: ta có $P_U = \begin{cases} S_U \text{ if } |K| \geq \alpha \\ 0 \text{ if } |K| < \alpha \end{cases}$ and $|\mathcal{U}| = \begin{cases} \sum_{j=\alpha}^{\min(\beta,|K|)} \binom{|K|}{j} \text{ if } |K| \geq \alpha \\ 0 \text{ if } |K| < \alpha \end{cases}$

3. Suy diễn X-gate

- Điều kiện U-gate trên $|U|$ là bất kỳ, chỉ phụ thuộc vào các A_i (*ON* hoặc *OFF*) và các kết hợp những A_i. Ví dụ, AND-gate và OR-gate là những trường hợp đặc biệt của U-gate với $|U|=n$ và $|U|\geq 1$. XOR-gate và XNOR-gate là những trường hợp đặc biệt của U-gate với các điều kiện đặc thù trên A_i.
- Trong nghiên cứu này, U-gate là cổng phi tuyến tổng quát nhất, trong đó xác suất U-gate chứa tích các trọng số.
- Tiếp theo, chúng ta nghiên cứu SIGMA-gate chỉ chứa kết hợp tuyến tính của những trọng số.

3. Suy diễn X-gate

- Suy diễn SIGMA-gate (Nguyen, 2016) biểu diễn quan hệ kết hợp, thỏa mãn **điều kiện SIGMA-gate** đặc tả bởi công thức 3.18 sau.

$$P(Y) = P\left(\sum_{i=1}^{n} A_i\right) \quad (3.18)$$

  Với tập những A_i đủ và loại trừ lẫn nhau.

$$\sum_{i=1}^{n} w_i = 1 \text{ and } A_i \cap A_j = \emptyset, \forall i \neq j$$

- Tổng sigma $\sum_{i=1}^{n} A_i$ chỉ rằng Y là hợp của những A_i rời và do đó, không biểu diễn phép cộng số học.

$$Y = \sum_{i=1}^{n} A_i = \bigcup_{i=1}^{n} A_i$$

$$P(Y) = P\left(\sum_{i=1}^{n} A_i\right) = P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

3. Suy diễn X-gate

- Công thức 3.19 sau đặc tả định lý **suy diễn SIGMA-gate** (Nguyen, 2016). Cơ sở của định lý này đã được các tác giả (Millán & Pérez-de-la-Cruz, 2002, pp. 292-295) đề cập.

$$P(X_1 + X_2 + \cdots + X_n) = P\left(\sum_{i=1}^{n} X_i\right) = P(Y = 1 | X_1, X_2, \ldots, X_n) = \sum_{i \in K} w_i$$

$$P(Y = 0 | X_1, X_2, \ldots, X_n) = 1 - \sum_{i \in K} w_i = \sum_{i \in L} w_i$$

Với tập những A_i đủ và loại trừ lẫn nhau.

$$\sum_{i=1}^{n} w_i = 1 \text{ and } A_i \cap A_j = \emptyset, \forall i \neq j$$

- Phần sau là chứng minh của định lý SIGMA-gate.

3. Suy diễn X-gate

Chứng minh định lý SIGMA-gate

$$(Y|X_1, X_2, \ldots, X_n) = P\left(\sum_{i=1}^{n} A_i \,|X_1, X_2, \ldots, X_n\right)$$

(do điều kiện SIGMA – gate)

$$= \sum_{i=1}^{n} P(A_i|X_1, X_2, \ldots, X_n)$$

(do các A_i loại trừ lẫn nhau)

$$= \sum_{i=1}^{n} P(A_i|X_i)$$

(do A_i chỉ phụ thuộc vào X_i)

Suy ra

$$P(Y = 1|X_1, X_2, \ldots, X_n) = \sum_{i=1}^{n} P(A_i = ON|X_i)$$

$$= \left(\sum_{i \in K} P(A_i = ON|X_i = 1)\right)$$

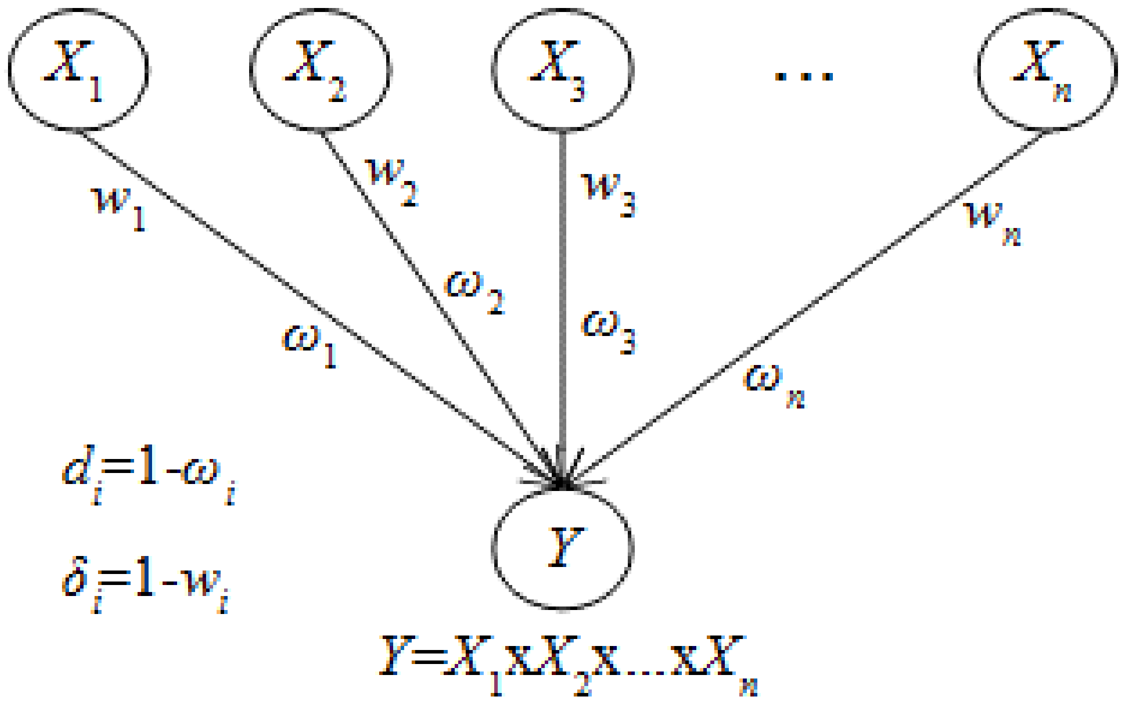
$$+ \left(\sum_{i \notin K} P(A_i = ON|X_i = 0)\right)$$

$$= \sum_{i \in K} w_i + \sum_{i \notin K} 0 = \sum_{i \in K} w_i$$

(do công thức 3.3)

3. Suy diễn X-gate

- Thông thường, mỗi cạnh trong đồ thị đơn giản gắn với một trọng số "*theo chiều kim đồng hồ*". Ví dụ, sự kiện X_i=1 gây ra sự kiện A_i=*ON* có trọng số "theo chiều kim đồng hồ" w_i.
- Tôi định nghĩa trọng số "*ngược chiều kim đồng hồ*". Ví dụ, sự kiện X_i=0 gây ra sự kiện A_i=*OFF* có trọng số "ngược chiều kim đồng hồ" ω_i.
- Do đó, mỗi cạnh sẽ gắn với hai trọng số w_i và ω_i, tạo thành *đồ thị đơn trọng số kép*.



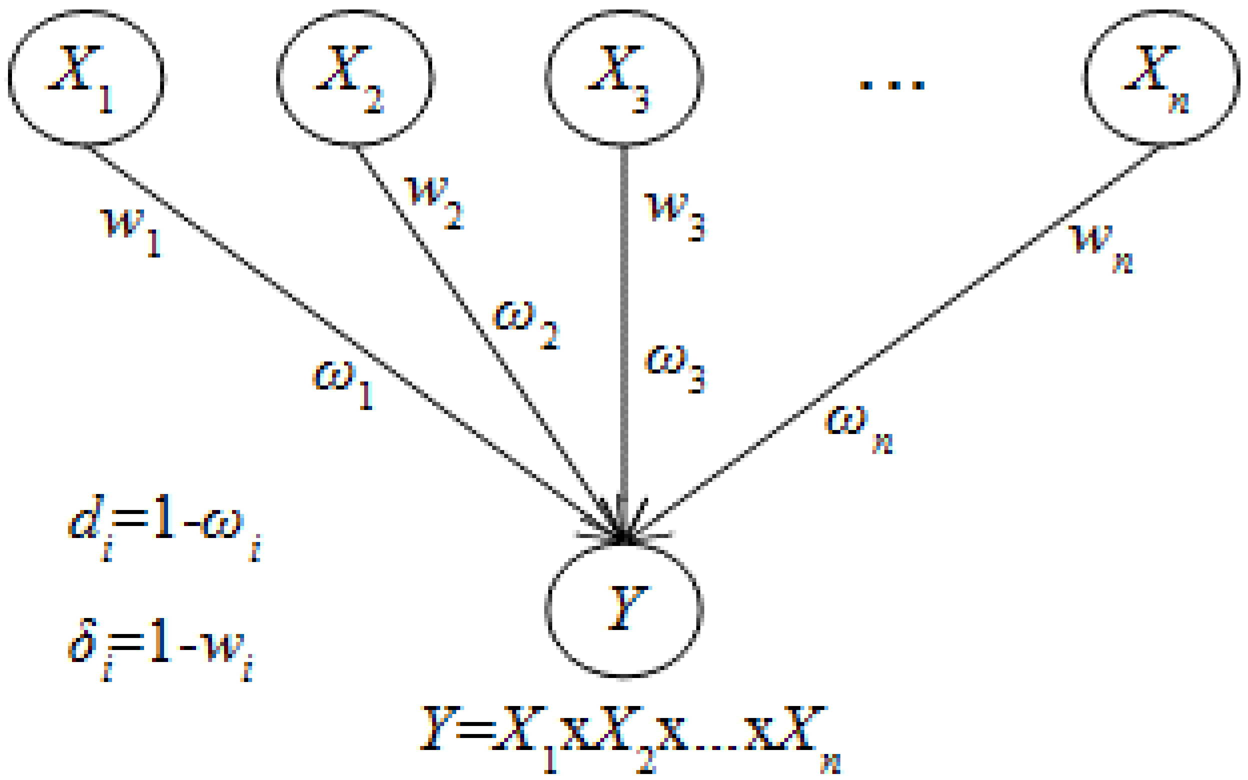
Đồ thị đơn trọng số kép

3. Suy diễn X-gate

$$P(A_i = ON|X_i = 1) = p_i = w_i$$
$$P(A_i = ON|X_i = 0) = 1 - \rho_i = 1 - \omega_i$$
$$P(A_i = OFF|X_i = 1) = 1 - p_i = 1 - w_i$$
$$P(A_i = OFF|X_i = 0) = \rho_i = \omega_i$$

$$d_i = 1 - \omega_i$$
$$\delta_i = 1 - w_i$$
$$W_i = w_i + d_i$$
$$\mathcal{W}_i = \omega_i + \delta_i$$

Đồ thị đơn trọng số kép



3. Suy diễn X-gate

Từ đồ thị đơn trọng số kép, chúng ta có những **suy diễn kép** cho AND-gate, OR-gate, NAND-gate, NOR-gate, XOR-gate, XNOR-gate, và U-gate

- $P(X_1 \odot X_2 \odot \ldots \odot X_n) = \prod_{i \in K} p_i \prod_{i \in L} d_i$
- $P(X_1 \oplus X_2 \oplus \ldots \oplus X_n) = 1 - \prod_{i \in K} \delta_i \prod_{i \in L} \rho_i$
- $P\left(\text{not}(X_1 \odot X_2 \odot \ldots \odot X_n)\right) = 1 - \prod_{i \in L} \rho_i \prod_{i \in K} \delta_i$
- $P\left(\text{not}(X_1 \oplus X_2 \oplus \ldots \oplus X_n)\right) = \prod_{i \in L} d_i \prod_{i \in K} p_i$
- $P(X_1 \otimes X_2 \otimes \ldots \otimes X_n) = \prod_{i \in O_1} p_i \prod_{i \in O_2} d_i \prod_{i \in E_1} \delta_i \prod_{i \in E_2} \rho_i + \prod_{i \in E_1} p_i \prod_{i \in E_2} d_i \prod_{i \in O_1} \delta_i \prod_{i \in O_2} \rho_i$
- $P\left(\text{not}(X_1 \otimes X_2 \otimes \ldots \otimes X_n)\right) = \prod_{i \in K} p_i \prod_{i \in L} d_i + \prod_{i \in K} \delta_i \prod_{i \in L} \rho_i$
- $P(X_1 \uplus X_2 \uplus \ldots \uplus X_n) = \sum_{U \in \mathcal{U}} \left(\prod_{i \in U \cap K} p_i \prod_{i \in U \cap L} d_i\right)\left(\prod_{i \in \bar{U} \cap K} \delta_i \prod_{i \in \bar{U} \cap L} \rho_i\right)$

3. Suy diễn X-gate

- Công thức 3.22 đặc tả **suy diễn kép SIGMA-gate**.

$$P(X_1 + X_2 + \cdots + X_n) = \sum_{i \in K} w_i + \sum_{i \in L} d_i$$

Với tập những X_i đủ và loại trừ lẫn nhau.

$$\sum_{i=1}^{n} W_i = 1 \text{ and } X_i \cap X_j = \emptyset, \forall i \neq j$$

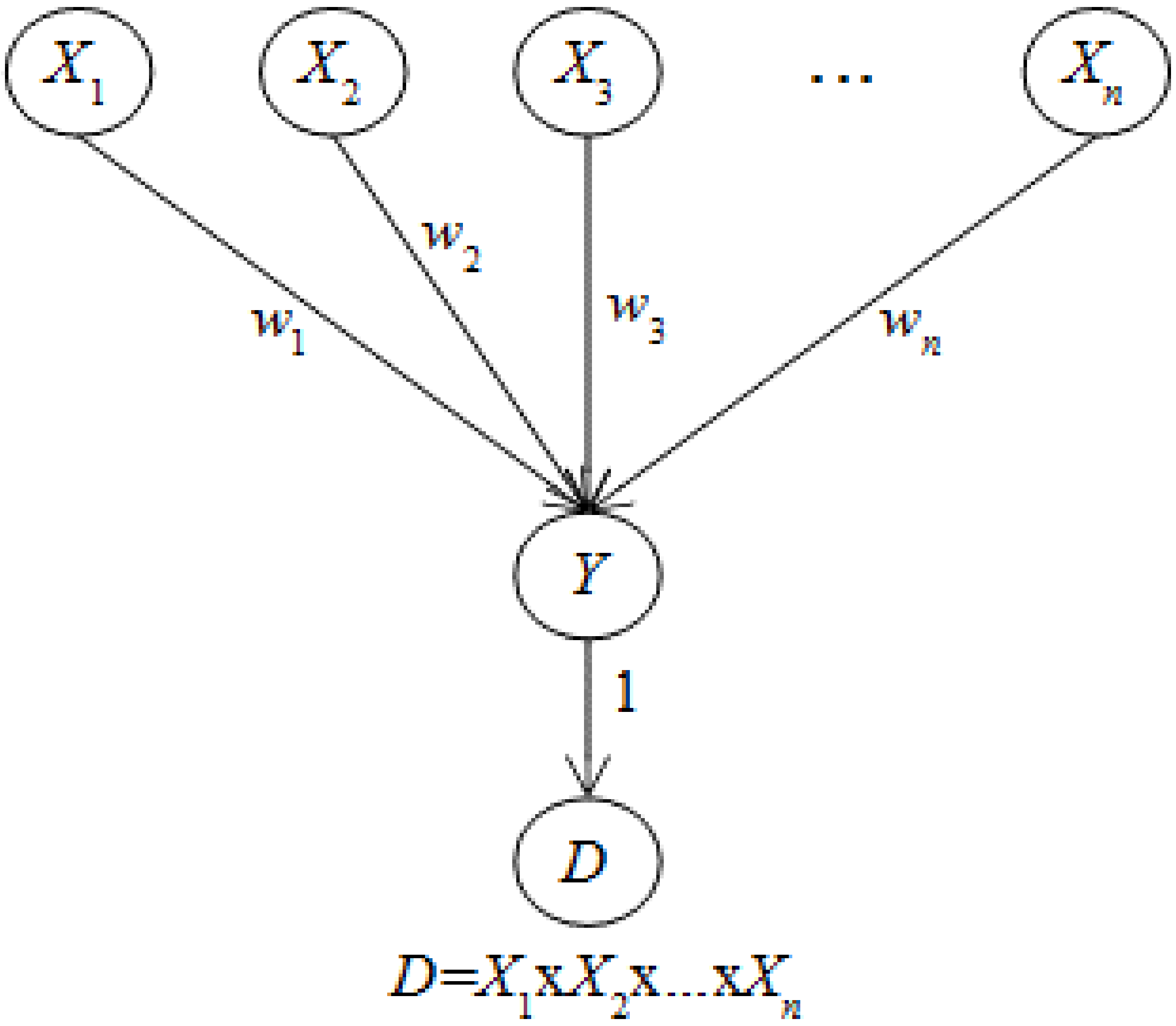
- Sau đây là chứng minh suy diễn kép SIGMA-gate.

$$P(X_1 + X_2 + \cdots + X_n) = \sum_{i=1}^{n} P(A_i = ON | X_i)$$

$$= \sum_{i \in K} P(A_i = ON | X_i = 1) + \sum_{i \in L} P(A_i = ON | X_i = 0) = \sum_{i \in K} w_i + \sum_{i \in L} d_i$$

4. Quan hệ chẩn đoán nhiều giả thuyết

- Trong đồ thị đơn giản, nếu thay biến đích *Y* bởi bằng chứng *D*, ta sẽ được *quan hệ chẩn đoán nhiều giả thuyết*, gọi ngắn gọn là *quan hệ chẩn đoán X-gate* gắn chặt với suy diễn X-gate.
- Mạng X-gate đòi hỏi biến đích phải nhị phân trong khi bằng chứng *D* có giá trị số. Do đó biến phụ nhị phân *Y* được thêm vào và *D* gắn trực tiếp vào *Y*. Cuối cùng ta được mạng chẩn đoán X-gate, gọi tắt là **mạng X-D**.



4. Quan hệ chẩn đoán nhiều giả thuyết

$$P(X_1, X_2, \dots, X_n, D) = \frac{P(D, X_1, X_2, \dots, X_n)}{P(X_1, X_2, \dots, X_n)} \prod_{i=1}^{n} P(X_i)$$

(do luật Bayes)

$$= \frac{\sum_Y P(D, X_1, X_2, \dots, X_n | Y) P(Y)}{P(X_1, X_2, \dots, X_n)} \prod_{i=1}^{n} P(X_i)$$

(do luật xác suất tổng)

$$= \frac{\sum_Y P(D, X_1, X_2, \dots, X_n | Y) P(Y)}{P(X_1, X_2, \dots, X_n)} \prod_{i=1}^{n} P(X_i)$$

$$= \left(\sum_Y P(D, X_1, X_2, \dots, X_n | Y) * \frac{P(Y)}{P(X_1, X_2, \dots, X_n)} \right) * \prod_{i=1}^{n} P(X_i)$$

$$= \left(\sum_Y P(D|Y) * \frac{P(X_1, X_2, \dots, X_n | Y) P(Y)}{P(X_1, X_2, \dots, X_n)} \right) * \prod_{i=1}^{n} P(X_i)$$

(do D độc lập điều kiện với tất cả X_i, cho Y)

$$= \left(\sum_Y P(D|Y) * \frac{P(Y, X_1, X_2, \dots, X_n)}{P(X_1, X_2, \dots, X_n)} \right) * \prod_{i=1}^{n} P(X_i)$$

$$= \sum_Y P(D|Y) P(Y | X_1, X_2, \dots, X_n) \prod_{i=1}^{n} P(X_i)$$

(do luật Bayes)

$$= \sum_Y P(X_1, X_2, \dots, X_n, Y, D)$$

$$\boxed{P(X_1, X_2, \dots, X_n, D) = P(D | X_1, X_2, \dots, X_n) \prod_{i=1}^{n} P(X_i)}$$

Xác suất hợp của mạng X-D (4.1)

4. Quan hệ chẩn đoán nhiều giả thuyết

Những xác suất cơ bản liên quan đến mạng X-D với phân bố đều

- $P(\Omega, Y, D) = P(X_1, X_2, \dots, X_n, Y, D) = P(D|Y)P(Y|X_1, X_2, \dots, X_n) \prod_{i=1}^{n} P(X_i)$
- $P(D|X_i) = \frac{P(X_i, D)}{P(X_i)} = \frac{\sum_{\{\Omega, Y, D\} \backslash \{X_i, D\}} P(\Omega, Y, D)}{\sum_{\{\Omega, Y, D\} \backslash \{X_i\}} P(\Omega, Y, D)}$
- $P(X_i|D) = \frac{P(X_i, D)}{P(D)} = \frac{\sum_{\{\Omega, Y, D\} \backslash \{X_i, D\}} P(\Omega, Y, D)}{\sum_{\{\Omega, Y, D\} \backslash \{D\}} P(\Omega, Y, D)}$
- $P(X_i, D) = \frac{1}{2^n S}\left((2D - M)s(\Omega: \{X_i\}) + 2^{n-1}(M - D)\right)$
- $P(D) = \frac{1}{2^n S}\left((2D - M)s(\Omega) + 2^n(M - D)\right)$

4. Quan hệ chẩn đoán nhiều giả thuyết

Xác suất điều kiện, xác suất hậu nghiệm và hệ số chuyển hóa của mạng X-D theo công thức 4.4

- $P(D|X_i = 1) = \frac{P(X_i=1,D)}{P(X_i=1)} = \frac{(2D-M)s(\Omega:\{X_i=1\})+2^{n-1}(M-D)}{2^{n-1}S}$
- $P(D|X_i = 0) = \frac{P(X_i=0,D)}{P(X_i=0)} = \frac{(2D-M)s(\Omega:\{X_i=0\})+2^{n-1}(M-D)}{2^{n-1}S}$
- $P(X_i = 1|D) = \frac{P(X_i=1,D)}{P(D)} = \frac{(2D-M)s(\Omega:\{X_i=1\})+2^{n-1}(M-D)}{(2D-M)s(\Omega)+2^{n}(M-D)}$
- $P(X_i = 0|D) = 1 - P(X_i = 1|D) = \frac{(2D-M)s(\Omega:\{X_i=0\})+2^{n-1}(M-D)}{(2D-M)s(\Omega)+2^{n}(M-D)}$
- $k = \frac{P(X_i|D)}{P(D|X_i)} = \frac{2^{n-1}S}{(2D-M)s(\Omega)+2^{n}(M-D)}$

4. Quan hệ chẩn đoán nhiều giả thuyết

Định lý chẩn đoán

Mạng X-D là sự kết hợp giữa quan hệ chẩn đoán và suy diễn X-gate:

$$P(Y=1|X_1,X_2,\ldots,X_n)=P(X_1 \text{x} X_2 \text{x} \ldots \text{x} X_n)$$

$$P(D|Y)=\begin{cases}\dfrac{D}{S} \text{ if } Y=1\\ \dfrac{M}{S}-\dfrac{D}{S} \text{ if } Y=0\end{cases}$$

Điều kiện chẩn đoán của mạng X-D được thỏa mãn khi và chỉ khi

$$s(\Omega)=\sum_a P(Y=1|a(\Omega))=2^{|\Omega|-1}, \forall\Omega\neq\emptyset$$

Hệ số chuyển hóa lúc này là:

$$k=\frac{N}{2}$$

Lưu ý, các trọng số p_i=w_i và ρ_i=ω_i, đầu vào của $s(\Omega)$, là những biến trừu tượng. Do đó, đẳng thức $s(\Omega) = 2^{|\Omega|-1}$ ngụ ý tất cả biến trừu tượng đều bị khử và $s(\Omega)$ không phụ thuộc vào trọng số.

4. Quan hệ chẩn đoán nhiều giả thuyết

Chứng minh định lý chẩn đoán

Hệ số chuyển hóa được viết lại: $k = \frac{2^{n-1}S}{2D(s(\Omega)-2^{n-1})+M(2^n-s(\Omega))}$

Trường hợp nhị phân với *D*=0 và *S*=1, ta có: $2^{n-1} = 2^{n-1} * 1 = 2^{n-1}S = aD^j = a * 0^j = 0$

Xảy ra mâu thuẫn, nên không thể biến đổi *k* về dạng: $k = \frac{aD^j}{bD^j}$

Do đó, nếu *k* là hằng số với *D* thì, $2D(s(\Omega) - 2^{n-1}) + M\left(2^n - s(\Omega)\right) = C \neq 0, \forall D$

Với *C* là hằng số. Ta có: $\sum_D \left(2D(s(\Omega) - 2^{n-1}) + M\left(2^n - s(\Omega)\right)\right) = \sum_D C \Rightarrow 2S(s(\Omega) - 2^{n-1}) + NM\left(2^n - s(\Omega)\right) = NC \Rightarrow 2^nS = NC$

Suy ra $k = \frac{2^{n-1}S}{2D(s(\Omega)-2^{n-1})+M(2^n-s(\Omega))} = \frac{NC}{2C} = \frac{N}{2}$

Nghĩa là $2^nS = N\left(2D(s(\Omega) - 2^{n-1}) + M\left(2^n - s(\Omega)\right)\right) = 2ND(s(\Omega) - 2^{n-1}) + 2S\left(2^n - s(\Omega)\right)$

$$\Rightarrow 2ND(s(\Omega) - 2^{n-1}) - 2S(s(\Omega) - 2^{n-1}) = 0$$

$$\Rightarrow (ND - S)(s(\Omega) - 2^{n-1}) = 0$$

Giả sử *ND*=*S* ta có: $ND = S = 2NM \Rightarrow D = 2M$

Xảy ra mâu thuẫn vì *M* là giá trị lớn nhất của *D*. Do đó, nếu *k* là hằng số với *D* thì $s(\Omega) = 2^{n-1}$. Ngược lại, nếu $s(\Omega) = 2^{n-1}$ thì *k* là: $k = \frac{2^{n-1}S}{2D(2^{n-1}-2^{n-1})+M(2^n-2^{n-1})} = \frac{N}{2}$

Tóm lại, sự kiện *k* là hằng số với *D* tương đương với sự kiện $s(\Omega) = 2^{n-1}$.

4. Quan hệ chẩn đoán nhiều giả thuyết

Những xác suất và hệ số chuyển hóa liên quan đến mạng X-D và suy diễn AND-gate (*mạng AND-D*), theo công thức 4.5.

$$P(D|X_i = 1) = \frac{(2D - M)\prod_{i=1}^{n} p_i + 2^{n-1}(M - D)}{2^{n-1}S}$$

$$P(D|X_i = 0) = \frac{M - D}{S}$$

$$P(X_i = 1|D) = \frac{(2D - M)\prod_{i=1}^{n} p_i + 2^{n-1}(M - D)}{(2D - M)\prod_{i=1}^{n} p_i + 2^{n}(M - D)}$$

$$P(X_i = 0|D) = \frac{2^{n-1}(M - D)}{(2D - M)\prod_{i=1}^{n} p_i + 2^{n}(M - D)}$$

$$k = \frac{2^{n-1}S}{(2D - M)\prod_{i=1}^{n} p_i + 2^{n}(M - D)}$$

Để thuận tiện, chúng ta kiểm định điều kiện chẩn đoán với trường hợp đơn giản nhất Ω = $\{X_1, X_2\}$, $p_1 = p_2 = w_1 = w_2 = 0.5$, $D \in \{0,1,2,3\}$. Bằng cách áp dụng định lý chẩn đoán, mạng AND-D không thỏa điều kiện chẩn đoán vì $s(\Omega) = 0.25$.

4. Quan hệ chẩn đoán nhiều giả thuyết

- AND-gate, OR-gate, XOR-gate và XNOR-gate không thỏa điều kiện chẩn đoán và không nên dùng chúng để kiểm định giả thuyết. Tuy nhiên chưa thể xác định U-gate và SIGMA-gate có thỏa điều kiện chẩn đoán hay không.
- Công thức 4.6 đặc tả xác suất của **mạng SIGMA-D** như sau:

$$P(D|X_i = 1) = \frac{(2D - M)w_i + M}{2S}$$

$$P(D|X_i = 0) = \frac{(M - 2D)w_i + M}{2S}$$

$$P(X_i = 1|D) = \frac{(2D - M)w_i + M}{2M}$$

$$P(X_i = 0|D) = \frac{(M - 2D)w_i + M}{2M}$$

$$k = \frac{N}{2}$$

- Bằng cách áp dụng định lý chẩn đoán cho mạng SIGMA-D, ta có $s(\Omega) = 2^{n-1}\sum_i (w_i + d_i) = 2^{n-1}$, vậy *mạng SIGMA-D thỏa điều kiện chẩn đoán*.

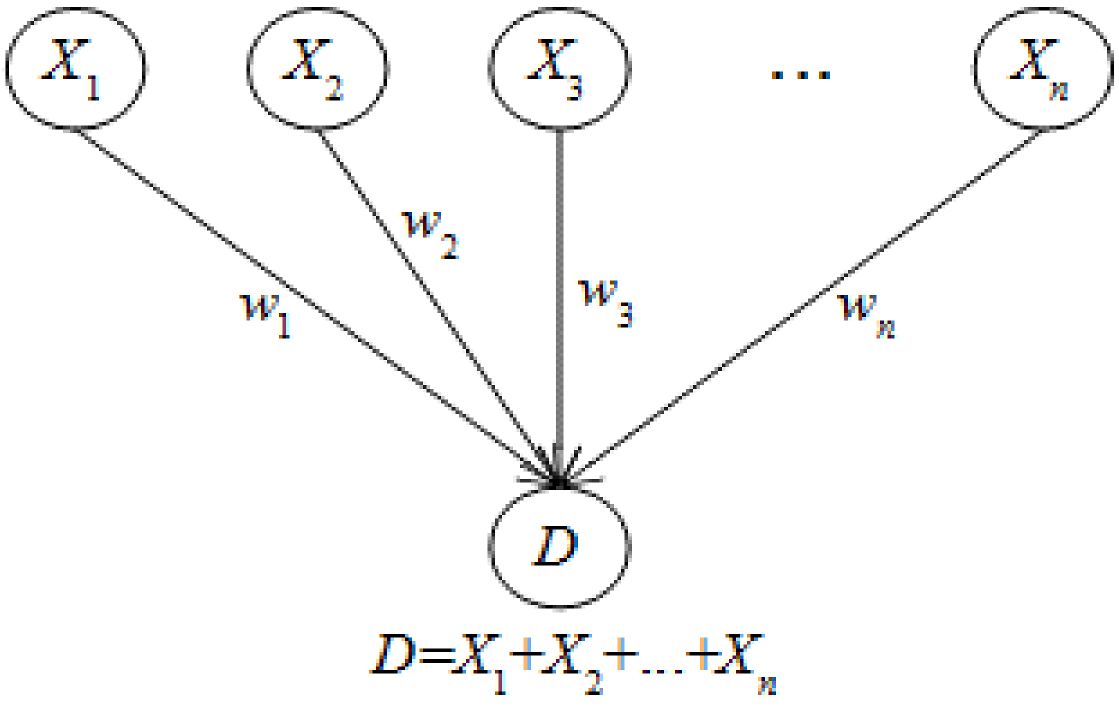
4. Quan hệ chẩn đoán nhiều giả thuyết

- Trong trường hợp SIGMA-gate, biến phụ Y bị khử khỏi mạng X-D và bằng chứng D trở thành biến đích trực tiếp, vậy ta có **mạng SIGMA-D trực tiếp**.
- Công thức 4.7 sau đặc tả CPT của mạng SIGMA-D trực tiếp.

$$P(D|X_1, X_2, \dots, X_n)$$
$$= \sum_{i \in K} \frac{D}{S} w_i + \sum_{j \in L} \frac{M-D}{S} w_j$$

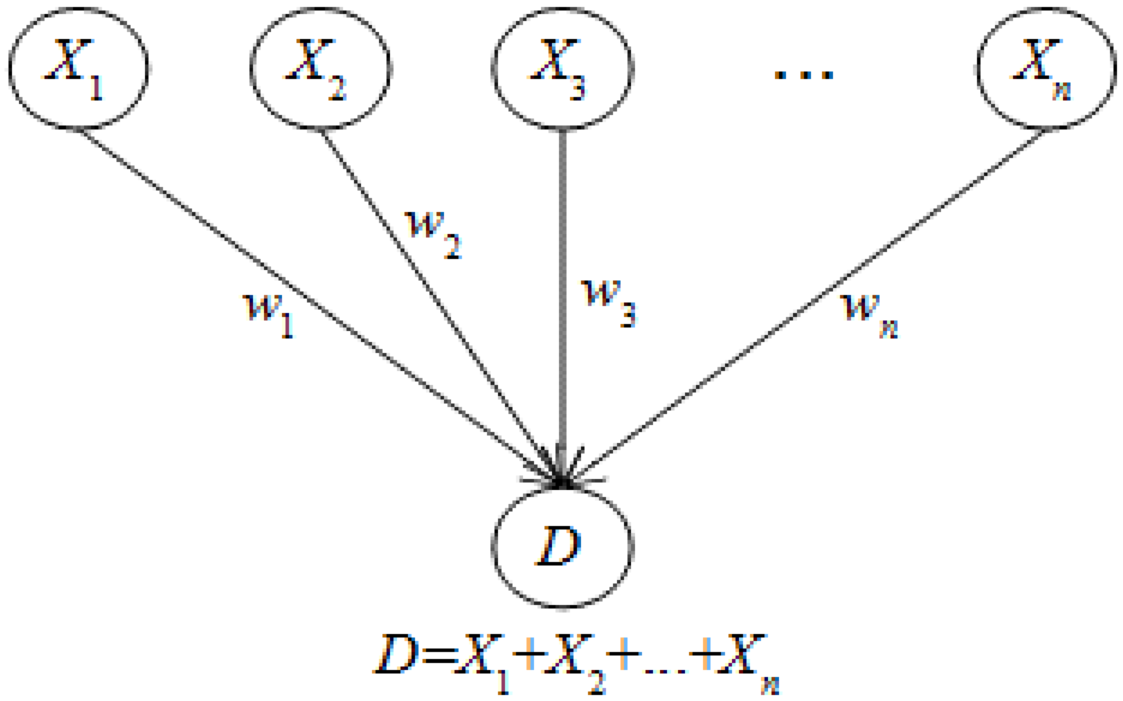
Với tập những X_i đủ và loại trừ lẫn nhau.

$$\sum_{i=1}^{n} w_i = 1 \text{ and } X_i \cap X_j = \emptyset, \forall i \neq j$$



4. Quan hệ chẩn đoán nhiều giả thuyết

- Với công thức 4.6, **mạng SIGMA-D trực tiếp** có cùng những xác suất điều kiện $P(X_i|D)$ và $P(D|X_i)$ với mạng SIGMA-D.
- Mạng SIGMA-D trực tiếp thỏa điều kiện chẩn đoán vì $s(\Omega) = 2^{n-1}$.



4. Quan hệ chẩn đoán nhiều giả thuyết

- Mạng X-D phi tuyến tổng quát nhất là mạng U-D trong khi mạng SIGMA-D thì tuyến tính. Mạng X-D phi tuyến như AND, OR, NAND, NOR, XOR và NXOR là những trường hợp đặc biệt của mạng X-D. **Bây giờ chúng ta kiểm định mạng U-D có thỏa điều kiện chẩn đoán hay không**.
- Suy diễn U-gate với điều kiện bất kỳ trên U là $P(X_1 \uplus X_2 \uplus \ldots \uplus X_n) = \sum_{U \in \mathcal{U}} \left(\prod_{i \in U \cap K} p_i \prod_{i \in U \cap L} (1 - \rho_i)\right) \left(\prod_{i \in \bar{U} \cap K} (1 - p_i) \prod_{i \in \bar{U} \cap L} \rho_i\right)$
- Đặt f là tổng qua tất cả chỉnh hợp của suy diễn U-gate.

$$f(p_i, \rho_i) = \sum_{a(\Omega)} \sum_{U \in \mathcal{U}} \left(\prod_{i \in U \cap K} p_i \prod_{i \in U \cap L} (1 - \rho_i)\right) \left(\prod_{i \in \bar{U} \cap K} (1 - p_i) \prod_{i \in \bar{U} \cap L} \rho_i\right)$$

4. Quan hệ chẩn đoán nhiều giả thuyết

Hàm *f* là tổng của nhiều biểu thức và mỗi biểu thức là tích của bốn tích con (Π) như sau:

$$Expr = \prod_{i \in U \cap K} p_i \prod_{i \in U \cap L} (1 - \rho_i) \prod_{i \in \bar{U} \cap K} (1 - p_i) \prod_{i \in \bar{U} \cap L} \rho_i$$

Trong trường hợp suy biến, luôn tồn tại những biểu thức *Expr* có ít nhất 2 tích con (Π), ví dụ: $Expr = \prod_{i \in U \cap K} p_i \prod_{i \in U \cap L} (1 - \rho_i)$

Luôn tồn tại những *Expr* có ít nhất 5 số hạng liên quan đến p_i và ρ_i nếu $n \geq 5$, ví dụ:

$$Expr = p_1 p_2 p_3 (1 - \rho_4)(1 - \rho_5)$$

Do đó bậc của *f* sẽ lớn hơn hay bằng 5 với $n \geq 5$. Không mất tính tổng quát, mỗi p_i hoặc ρ_i tương ứng là tổng của biến *x* và một biến a_i hoặc b_i. Lưu ý, tất cả p_i, ρ_i, a_i are b_i là biến trừu tượng.

$$p_i = x + a_i$$

$$\rho_i = x + b_i$$

Phương trình $f - 2^{n-1} = 0$ trở thành phương trình $g(x) = 0$ có bậc là $m \geq 5$ nếu $n \geq 5$.

$$g(x) = \pm x^m + C_1 x^{m-1} + \cdots + C_{m-1} x + C_m - 2^{n-1} = 0$$

Những hệ số C_i là hàm của các a_i và b_i. Theo định lý Abel-Ruffini (Wikipedia, Abel-Ruffini theorem, 2016), phương trình $g(x) = 0$ không có nghiệm đại số khi $m \geq 5$. Do đó những biến trừu tượng p_i và ρ_i không bị khử hoàn toàn khỏi $g(x)=0$, vậy **không có đặc tả nào của suy diễn U-gate $\boldsymbol{P(X_1 \mathrm{x} X_2 \mathrm{x} \ldots \mathrm{x} X_n)}$ thỏa điều kiện chẩn đoán**.

4. Quan hệ chẩn đoán nhiều giả thuyết

- Với kết luận không có mạng X-D phi tuyến thỏa điều kiện chẩn đoán, một câu hỏi mới được đặt ra: Có tồn tại mạng X-D tuyến tính tổng quát thỏa điều kiện chẩn đoán hay không?
- Mạng tuyến tính tổng quát này được gọi là **mạng GL-D** và mạng SIGMA-D là trường hợp đặc biệt của mạng GL-D. Xác suất GL-gate phải là tổ hợp tuyến tính của những trọng số.

$$P(X_1 \mathrm{x} X_2 \mathrm{x} \dots \mathrm{x} X_n) = C + \sum_{i=1}^{n} \alpha_i w_i + \sum_{i=1}^{n} \beta_i d_i$$

- Suy diễn GL-gate là đơn nếu α_i và β_i là hàm của chỉ những X_i như sau:

$$P(X_1 \mathrm{x} X_2 \mathrm{x} \dots \mathrm{x} X_n) = C + \sum_{i=1}^{n} h_i(X_i) w_i + \sum_{i=1}^{n} g_i(X_i) d_i$$

4. Quan hệ chẩn đoán nhiều giả thuyết

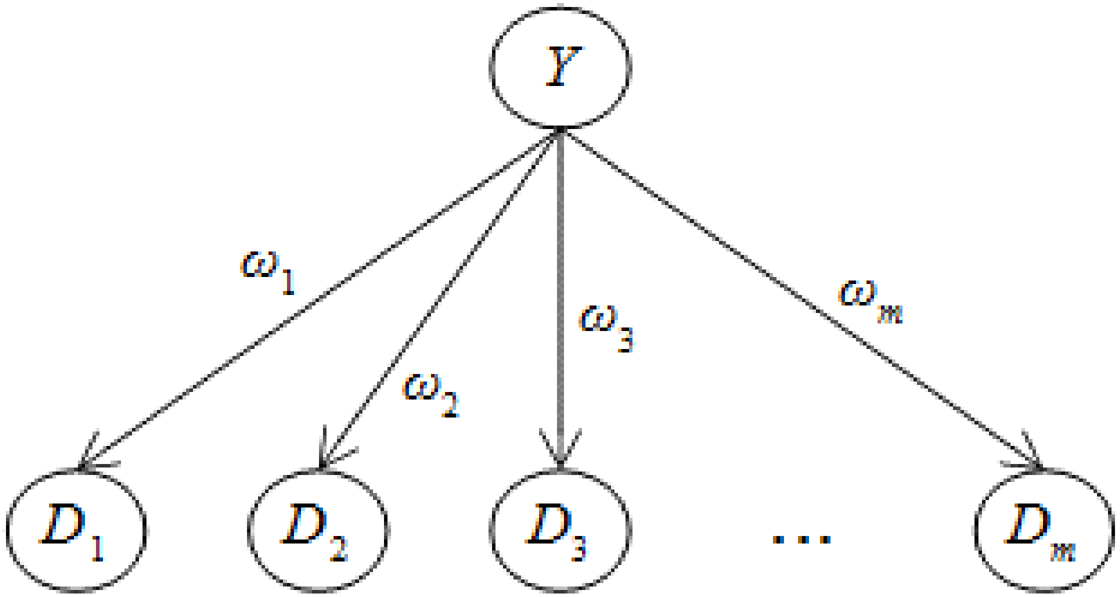
- Giả sử h_i và g_i là những hàm phân bố xác suất của X_i. Với mọi i, ta có: $0 \leq h_i(X_i) \leq 1$, $0 \leq g_i(X_i) \leq 1$, $h_i(X_i = 1) + h_i(X_i = 0) = 1$, $g_i(X_i = 1) + g_i(X_i = 0) = 1$
- **Mạng GL-D** thỏa điều kiện chẩn đoán nếu $s(\Omega) = 2^n C + 2^{n-1} \sum_{i=1}^{n}(w_i + d_i) = 2^{n-1} \Rightarrow 2C + \sum_{i=1}^{n}(w_i + d_i) = 1$
- Giả sử tập những X_i là đủ, ta có $\sum_{i=1}^{n}(w_i + d_i) = 1$. Suy ra C=0.
- Tóm lại, công thức 4.10 đặc tả suy diễn GL-gate đơn sao cho mạng GL-D thỏa điều kiện chẩn đoán.

$$P(X_1 \text{x} X_2 \text{x} \ldots \text{x} X_n) = \sum_{i=1}^{n} h_i(X_i) w_i + \sum_{i=1}^{n} g_i(X_i) d_i$$

Với h_i và g_i là những hàm phân bố xác suất và tập những X_i đủ $\sum_{i=1}^{n}(w_i + d_i) = 1$

4. Quan hệ chẩn đoán nhiều giả thuyết

- Theo những tác giả (Millán & Pérez-de-la-Cruz, 2002), một giả thuyết có thể có nhiều bằng chứng, như trong hình bên cạnh. Đây là *quan hệ chẩn đoán nhiều bằng chứng* được gọi ngắn gọn là **mạng M-E-D**, ngược lại với quan hệ chẩn đoán nhiều giả thuyết.
- Xác suất hợp của mạng M-E-D là $P(Y, D_1, D_2, \ldots, D_m) = P(Y)\prod_{j=1}^{m} P\left(D_j \middle| Y\right) = P(Y)P(D_1, D_2, \ldots, D_m|Y)$
- Hệ số chuyển hóa khả dĩ của mạng M-E-D là $\frac{1}{k} = \prod_{j=1}^{m} P\left(D_j \middle| Y = 1\right) + \prod_{j=1}^{m} P\left(D_j \middle| Y = 0\right)$



4. Quan hệ chẩn đoán nhiều giả thuyết

- Mạng M-E-D sẽ thỏa điều kiện chẩn đoán nếu $k = 1$ vì tất cả giả thuyết và bằng chứng đều nhị phân, dẫn đến công thức 4.11 đặc tả phương trình sau có $2m$ nghiệm thực $P(D_j|Y)$ với mọi $m \geq 2$.

$$\prod_{j=1}^{m} P\left(D_j \middle| Y = 1\right) + \prod_{j=1}^{m} P\left(D_j \middle| Y = 0\right) = 1 \quad (4.11)$$

- Giả sử phương trình 4.11 có 4 nghiệm thực: $a_1 = P(D_1 = 1|Y = 1)$, $a_2 = P(D_2 = 1|Y = 1)$, $b_1 = P(D_1 = 1|Y = 0)$, $b_2 = P(D_2 = 1|Y = 0)$

- Từ phương trình 4.11, suy ra $\begin{cases} a_1 = a_2 = 0 \\ b_1 = b_2 \\ a_1^2 + b_1^2 = 1 \\ b_1 = 2 \end{cases}$ or $\begin{cases} a_1 = a_2 = 0.5 \\ b_1 = b_2 \\ a_1^2 + b_1^2 = 1 \\ b_1 = 1.5 \end{cases}$ dẫn đến mâu thuẫn (b_1=2 hoặc b_1=1.5) và vì vậy, **không thể áp dụng điều kiện chẩn đoán cho mạng M-E-D**.

4. Quan hệ chẩn đoán nhiều giả thuyết

- **Chưa thể mô hình hóa mạng M-E-D bằng những cổng X-gate**. Giải pháp khả dĩ là nhóm nhiều bằng chứng D_1, D_2,…, D_m thành một bằng chứng đại diện *D* kế tiếp phụ thuộc vào giả thuyết *Y* nhưng giải pháp này sẽ không chính xác khi đặc tả những xác suất điều kiện do sự thiếu nhất quán của chiều quan hệ (từ D_j đến *D* và từ *Y* đến *D*) ngoại trừ tất cả D_j bị khử và *D* trở thành véc-tơ. Tuy nhiên véc-tơ bằng chứng là giải pháp phức tạp khi gây ra vấn đề mới từ vấn đề cũ.
- Giải pháp khác là đảo chiều quan hệ, khi đó giả thuyết phụ thuộc vào bằng chứng để tận dụng suy diễn X-gate. Tuy nhiên phương pháp đảo chiều làm thay đổi quan điểm rằng quan hệ chẩn đoán phải từ giả thuyết đến bằng chứng.

4. Quan hệ chẩn đoán nhiều giả thuyết

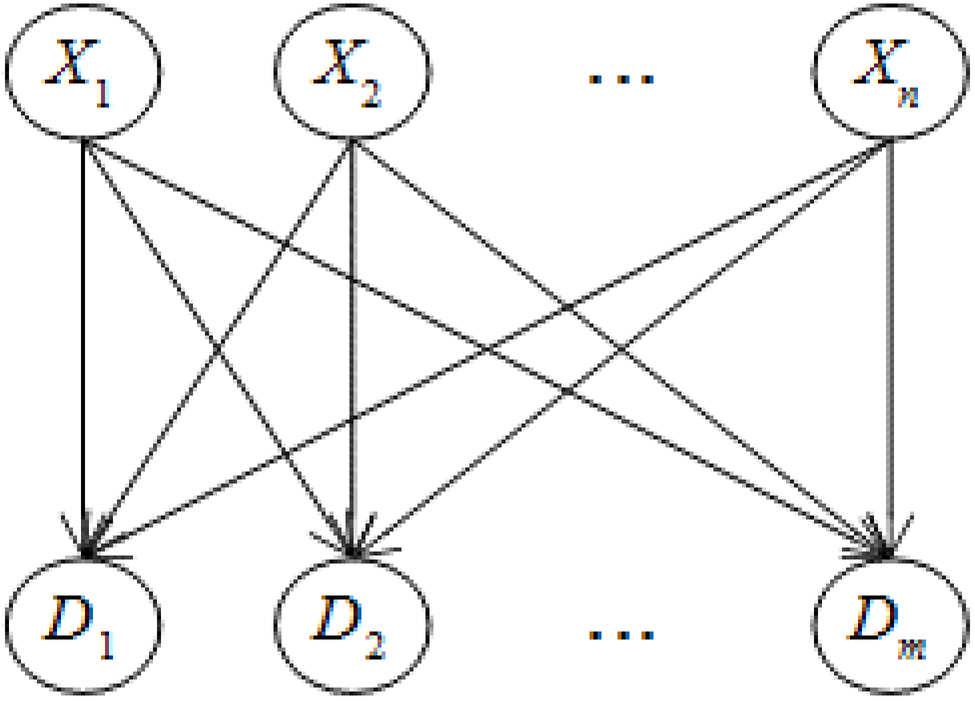
- Giải pháp khác là xây dựng **mạng M-E-D** với X-gate bằng *điều kiện chẩn đoán từng phần*, là sự nới lỏng mạng M-E-D, được định nghĩa như sau:

$$P\left(Y\middle|D_j\right) = kP\left(D_j\middle|Y\right)$$

- Xác suất hợp của mạng M-E-D là: $P(Y, D_1, D_2, \dots, D_m) = P(Y)\prod_{j=1}^{m} P\left(D_j\middle|Y\right)$
- Mạng M-E-D thỏa điều kiện chẩn đoán từng phần vì $P\left(Y\middle|D_j\right) = \frac{1}{2}P\left(D_j\middle|Y\right)$
- Điều kiện chẩn đoán từng phần thể hiện một quan điểm khác biệt nhưng không là giải pháp tối ưu do với một ví dụ rằng chúng ta không thể chẩn đoán bệnh khi chỉ dựa vào một triệu chứng mà bỏ qua những triệu chứng rõ ràng khác.

4. Quan hệ chẩn đoán nhiều giả thuyết

- Nếu thành công trong việc đặc tả những xác suất điều kiện của mạng M-E-D, chúng ta có thể xây dựng mạng mở rộng với *n* giả thuyết $X_1, X_2, \ldots, X_n$ và *m* bằng chứng $D_1, D_2, \ldots, D_m$. Mạng mở rộng này được gọi là **mạng M-HE-D**, biểu diễn *quan hệ chẩn đoán nhiều giả thuyết nhiều bằng chứng*.
- Mạng M-HE-D là trường hợp tổng quát nhất của mạng chẩn đoán, đã được nghiên cứu bởi các tác giả (Millán & Pérez-de-la-Cruz, 2002, p. 297). Chúng ta có thể xây dựng bất cứ BN lớn nào từ nhiều mạng M-HE-D và vì vậy, nghiên cứu này vẫn còn mở.



5. Kết luận

- Nhìn chung, chuyển hóa quan hệ là xác định những xác suất điều kiện dựa trên cổng lô-gic gắn chặt với ngữ nghĩa của quan hệ. Nhược điểm của cổng lô-gic là đòi hỏi mọi biến phải nhị phân.
- Để giảm nhẹ nhược điểm, tôi dùng bằng chứng có giá trị số nhằm mở rộng khả năng của BN đơn giản. Tuy nhiên, sẽ gây ra lỗi khi kết hợp giả thuyết nhị phân và bằng chứng số. Do đó tôi đề xuất *điều kiện chẩn đoán* để xác thực tính đúng đắn của bằng chứng số trong suy diễn BN.
- Nhiều BN đơn giản tạo thành BN lớn và sự suy diễn trong BN lớn rất phức tạp. Trong tương lai, tôi sẽ nghiên cứu phương pháp suy diễn hiệu quả cho những BN được tạo từ các cổng X-gate.
- Tôi sẽ cố gắng nghiên cứu sâu mạng M-E-D và mạng M-HE-D, đây là những vấn đề tôi chưa giải quyết được.

5. Kết luận

- Hai tài liệu chính tôi nghiên cứu là cuốn sách “Learning Bayesian Networks” của tác giả (Neapolitan, 2003) và bài báo “A Bayesian Diagnostic Algorithm for Student Modeling and its Evaluation” của hai tác giả (Millán & Pérez-de-la-Cruz, 2002).
- Đặc biệt, suy diễn SIGMA-gate được kế thừa và dựa trên công trình của hai tác giả Eva Millán and José Luis Pérez-de-la-Cruz.
- Nghiên cứu này xuất phát từ công trình PhD của tôi “A User Modeling System for Adaptive Learning” (Nguyen, 2014).
- Những nghiên cứu có liên quan là mô hình hóa người dùng, mô hình chồng và mạng Bayesian từ những công trình của các tác giả (Fröschl, 2005), (De Bra, Smits, & Stash, 2006), (Murphy, 1998) và (Heckerman, 1995).

Cảm ơn đã lắng nghe

Tham khảo

1. De Bra, P., Smits, D., & Stash, N. (2006). The Design of AHA! In U. K. Wiil, P. J. Nürnberg, & J. Rubart (Ed.), *Proceedings of the seventeenth ACM Hypertext Conference on Hypertext and hypermedia (Hypertext '06)* (pp. 171-195). Odense, Denmark: ACM Press.
2. Díez, F. J., & Druzdzel, M. J. (2007). *Canonical Probabilistic Models*. National University for Distance Education, Department of Inteligencia Artificial. Madrid: Research Centre on Intelligent Decision-Support Systems. Retrieved May 9, 2016, from http://www.cisiad.uned.es/techreports/canonical.pdf
3. Fröschl, C. (2005). *User Modeling and User Profiling in Adaptive E-learning Systems*. Master Thesis, Graz University of Technology, Austria.
4. Heckerman, D. (1995). *A Tutorial on Learning With Bayesian Networks*. Microsoft Corporation, Microsoft Research. Redmond: Microsoft Research. Retrieved from ftp://ftp.research.microsoft.com/pub/dtg/david/tutorial.ps
5. Kschischang, F. R., Frey, B. J., & Loeliger, H.-A. (2001, February). Factor Graphs and the Sum-Product Algorithm. *IEEE Transactions on Information Theory, 47*(2), 498-519. doi:10.1109/18.910572
6. Millán, E., & Pérez-de-la-Cruz, J. L. (2002, June). A Bayesian Diagnostic Algorithm for Student Modeling and its Evaluation. (A. Kobsa, Ed.) *User Modeling and User-Adapted Interaction, 12*(2-3), 281-330. doi:10.1023/A:1015027822614
7. Millán, E., Loboda, T., & Pérez-de-la-Cruz, J. L. (2010, July 29). Bayesian networks for student model engineering. (R. S. Heller, J. D. Underwood, & C.-C. Tsai, Eds.) *Computers & Education, 55*(4), 1663-1683. doi:10.1016/j.compedu.2010.07.010
8. Murphy, K. P. (1998). *A Brief Introduction to Graphical Models and Bayesian Networks*. Retrieved 2008, from Kevin P. Murphy's home page: http://www.cs.ubc.ca/~murphyk/Bayes/bnintro.html
9. Neapolitan, R. E. (2003). *Learning Bayesian Networks*. Upper Saddle River, New Jersey, USA: Prentice Hall.
10. Nguyen, L. (2014, April). *A User Modeling System for Adaptive Learning*. University of Science, Ho Chi Minh city, Vietnam. Abuja, Nigeria: Standard Research Journals. Retrieved from http://standresjournals.org/journals/SSRE/Abstract/2014/april/Loc.html
11. Nguyen, L. (2016, March 28). Theorem of SIGMA-gate Inference in Bayesian Network. (V. S. Franz, Ed.) *Wulfenia Journal, 23*(3), 280-289.
12. Pearl, J. (1986, September). Fusion, propagation, and structuring in belief networks. *Artificial Intelligence, 29*(3), 241-288. doi:10.1016/0004-3702(86)90072-X
13. Wikipedia. (2014, October 10). *Set (mathematics)*. (A. Rubin, Editor, & Wikimedia Foundation) Retrieved October 11, 2014, from Wikipedia website: http://en.wikipedia.org/wiki/Set_(mathematics)
14. Wikipedia. (2015, November 22). *Factor graph*. (Wikimedia Foundation) Retrieved February 8, 2017, from Wikipedia website: https://en.wikipedia.org/wiki/Factor_graph
15. Wikipedia. (2016, June 10). *Abel-Ruffini theorem*. (Wikimedia Foundation) Retrieved June 26, 2016, from Wikipedia website: https://en.wikipedia.org/wiki/Abel%E2%80%93Ruffini_theorem
16. Wikipedia. (2016, June 2). *Logic gate*. (Wikimedia Foundation) Retrieved June 4, 2016, from Wikipedia website: https://en.wikipedia.org/wiki/Logic_gate