1. **Câu 1**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | BFS | Greedy  f(n) = h(n) | A\*  f(m) =  g(m)+h(m)min | UCS | IDS | IDA\*  (DFS + A\*) | DFS |
| Xử lý lặp | Không | Không | Nếu f tốt hơn  - Đưa lại nút đã mở rộng  - Cập nhật nút ở biên | Nếu f tốt hơn  - Đưa lại nút đã mở rộng  - Cập nhật nút ở biên | Luôn đưa nút lặp quay lại | Luôn đưa nút lặp quay lại | - Không đưa quay lại nếu đã xét  - Đưa lại nếu nằm ở stack |
| Thêm vào O | Cuối | Anywhere | Anywhere | Anywhere | Đầu  (Chỉ thêm nếu depth(n) <= c, c++) | Đầu  (Chỉ thêm nếu f(m) <= i, i+=α) | Đầu |
| Lấy n ra khỏi O | Đầu | f tốt nhất (ưu tiên đứng trước trong alphabet) | f tốt nhất (ưu tiên đứng trước trong alphabet) | f tốt nhất (ưu tiên đứng trước trong alphabet) | Đầu | Đầu | Đầu |

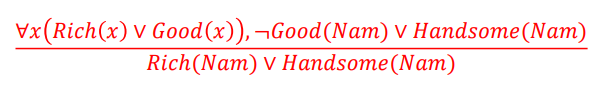
1. **Câu 2**
2. **Thủ tục suy diễn**

- Logic vị từ

* Lượng từ với mọi (ký hiệu ∀) được dùng với kéo theo chứ không dùng với “và”.
  + Chẳng hạn, để nói rằng mọi sinh viên đều chăm học
  + ∀x Sinh\_viên(x) ⇒ Chăm\_học(x) là đúng
  + ∀x Sinh\_viên(x) ∧ Chăm\_học(x) là sai (mọi người đều là sinh viên và đều chăm học)
* Lượng tử tồn tại (ký hiệu ∃) được dùng với “và” chứ không dùng với “kéo theo”.
  + Một số sinh viên chăm học
  + ∃x Sinh\_viên(x) ∧ Chăm\_học(x) là đúng trong khi
  + ∃x Sinh\_viên(x) ⇒ Chăm\_học(x) là sai (do đúng khi có một người x nào đó không phải là sinh viên)

- Suy diễn với logic vị từ

* *Phép thế:* SUBST ({x/Nam, y/An}, Like(x,y)) = Like(Nam, An)
* *Phép loại trừ với mọi:*∀x Like(x, Icecream),{x/Nam}: Like(Nam, Icecream)
* *Loại trừ tồn tại:* ∃x GoodAtMath(x), {x/C}: GoodAtMath(C)
* *Nhập đề tồn tại:* Like(Nam, Icecream),{Nam/x}: ∃Like(x, Icecream)
* *Phép hợp nhất*
* *GMP*
* *Phép giải*
  1. **Suy diễn tiến**
* Hợp nhất 1 phần vế trái, thêm vế phải còn lại vào KB và suy diễn tiếp
* Sử dụng GMP cho phép xây dựng thuật toán suy diễn tự động, suy diễn tiến và suy diễn lùi
* GMP: ∀x (Cat(x) => Like (x, Fish)), Cat(Tom): Like(Tom, Fish)
  1. **Suy diễn lùi**
* Hợp nhất vế phải, cố c/m phần tử vế trái bằng suy diễn lùi
  1. **Phép giải**



* 1. **Phép giải + Phản chứng**
* KB: ├ Q?
* Thêm ¬ Q vào KB, sau đó chứng minh tồn tại một tập con của KB mới có giá trị False

(KB ├ Q) <=> (KB ∧ ¬ Q ├ False)

1. **CF, CNF**

* CNF - dạng chuẩn hội: hội của các câu tuyển
* CF(clause form): Câu tuyển

Bước 1: P⇔Q ≡ (P=>Q)∧(Q=> P).

Bước 2: P⇒Q ≡ ¬ P ∨ Q

Bước 3: De Morgan

Bước 4: Chuẩn hóa tên biến | mỗi lượng tử có biến riêng

Bước 5: Loại bỏ các lượng tử ∃ bằng cách sử dụng hằng Skolem/hàm Skolem

* ∃x P(x) được biến đổi thành P(C), trong đó C là hằng Skolem chưa xuất hiện trong KB
* ∀x (∃y (P(x,y)) thành ∀x (P(x,f(x))

Bước 6: Loại bỏ ∀

* ∀x ( P (x,y) ∨ Q (x) ) => P (x,y) ∨ Q (x)

Bước 7: Sắp xếp và ra ngoài hoặc => CNF

* (P ∧ Q) ∨ R ≡ ( P ∨ R ) ∧( Q ∨ R )

Bước 8: Loại bỏ các phép và để tạo thành các clause riêng => CF

* ( P ∨ R ∨ S ) ∧ ( Q ∨ ¬ R ) => P ∨ R∨ S

Q ∨¬ R

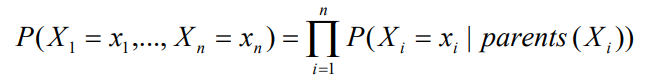
Bước 9: Chuẩn hóa tên biến sao cho mỗi câu có biến riêng của mình

1. **Câu 3**
2. **Một số công thức**

* P(A∨B) = P(A) + P(B) - P(A∧B)
* P(A) + P(¬ A) = 1
* P(A) = P(A,B) + P(A,¬ B)
* P(A|B) + P(¬A|B) = 1
* P(A|B).P(B) = P(B|A).P(A) = P(A,B) (Bayes)
* P(A, B, C, D) = P(A|B,C,D).P(B|C,D).P(C|D).P(D)
* B,C đlxs biết A: P(B,C|A) = P(B|A).P(C|A)
* B,C đlxs: P(B,C) = P(B).P(C)

1. **Mạng Bayes**

* *Độc lập xác suất:* Mỗi nút V độc lập với tất cả các nút không phải là hậu duệ của V, nếu biết giá trị các nút bố mẹ của V
* *Tính xác suất đồng thời (Tổng quát)*



1. **Suy diễn mạng Bayes**

* *Xác suất nhân quả:* 
  + Suy diễn nhân quả (trên xuống): cần tính P(Q|E) khi E là nút cha của Q
  + Các bước:

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

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

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

* *Xác suất chẩn đoán:* 
  + Suy diễn chẩn đoán (dưới lên): cần tính P(Q|E) khi E là nút con của Q
  + Các bước:

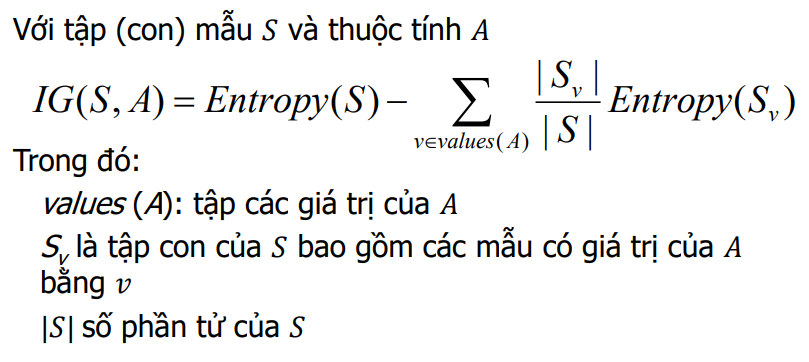
Bước 1: Biến đổi về suy diễn nhân quả sử dụng quy tắc bayes

Bước 2: Thực hiện giống suy diễn nhân quả

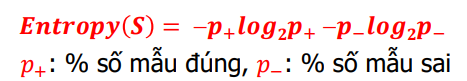
1. **Câu 4**
2. **Nút gốc ID3**

Các thuộc tính có độ tăng thông tin lớn nằm gần gốc => Chọn IG lớn nhất

* *Độ tăng thông tin IG*



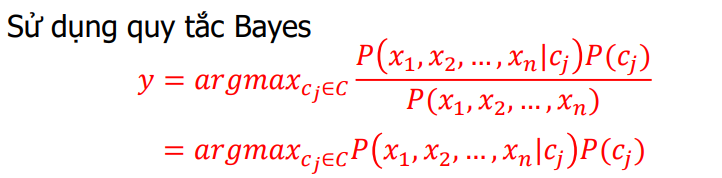
* *Entropy*

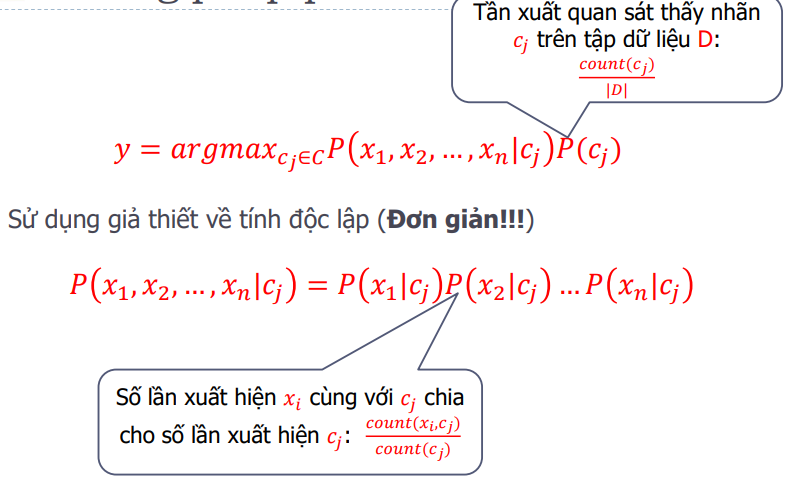


* H(S) [9+, 5-] = -(9/14)log2(9/14) - (5/14)log2(5/14)
* IG(S, Gió) = H(S) - (8/14)H(SGió = yếu) - (6/14)H(SGió = mạnh)
* H(SGió = yếu) [6+, 2-] = -(6/8)log2(6/8) - (2/8)log2(2/8)
* …

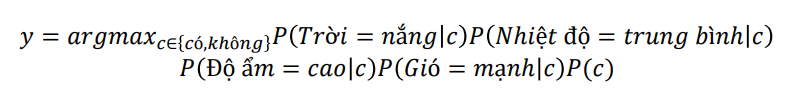
1. **Phân loại Bayes**

* *Sử dụng quy tắc Bayes*





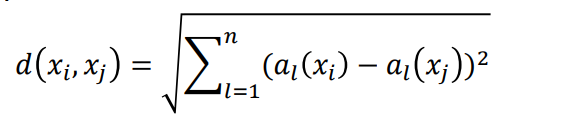
* Ví dụ xác định nhãn



* + cj = “có”: (2/9).(4/9).(3/9).(9/14) = 0.007
  + cj = “không”: (2/9).(4/9).(3/9).(9/14) = 0.041
  + Vậy cj = “không”

1. **K-NN**

* *Khoảng cách Euclidean*



* *Thuật toán:* Với mẫu x cần phân loại

B1: Tính d(x, xi) hoặc d2(x, xi) (xi: tất cả các mẫu trong CSDL)

B2: Tìm k mẫu d(x, xi) min

B3: Xác định phân loại nhãn chiếm đa số trong tập