

# Cơ sở trí tuệ nhân tạo

quocpt@sgu.edu.vn

## Nội dung lý thuyết (30 tiết)

- Tuần 1,2,3: Thuật toán heuristic/Thuật toán Greedy/Thuật toán A<sup>KT</sup>.
- Tuần 4,5,6: Biểu diễn tri thức + Đăng ký tiểu luận
- Tuần 7,8,9: Mở đầu về máy học
- Tuần 10: Bài tập tổng hợp và Kiểm tra lần 1
- Tuần 11,12,13,14: Một số thuật toán metaheuristic
- Tuần 15: Ôn thi và Kiểm tra lần 2.

## Tài liệu tham khảo chính

- MIT Open Courseware, “*Artificial Intelligence*”, Spring, 2005

## Đánh giá học phần

- Điểm thi kết thúc HP: 50% (lý thuyết, 90 phút, tự luận, không sử dụng tài liệu, trường tổ chức thi).
- Điểm quá trình: 50% (Lý thuyết 25% và thực hành 25%).  
(Đánh giá phần lý thuyết)
  - ✓ Chuyên cần & Thảo luận giờ lý thuyết: 5%
  - ✓ Kiểm tra lý thuyết lần 1: 10%
  - ✓ Kiểm tra lý thuyết lần 2: 10%
  - ✓ Điểm +: Đối với các SV hoàn thành tốt các bài tập & thảo luận.
  - ✓ Sinh viên có thể chọn làm tiểu luận (cá nhân, 10%) để thay thế bài kiểm tra lý thuyết lần 2.

## Nội dung tiểu luận

Mỗi sinh viên chọn thực hiện một trong các đề tài sau:

1. Cài đặt 01 trong số 4 trò chơi sau: Trò chơi NIM ( $n=3$ ), trò chơi Caro ( $d=5$ ), Taci (n bất kỳ), tháp Hà Nội (n bất kỳ)
2. Cài đặt 2 thuật toán (Quinlan/ hoặc ILA) và thuật toán phân cụm Kmean (thuật toán, chương trình, thực nghiệm)
3. Xây dựng 1 thuật toán (dạng metaheuristic) và chương trình giải bài toán cây steiner nhỏ nhất
4. Xây dựng 1 thuật toán (dạng metaheuristic) và chương trình giải bài toán max clique.

## Một số nội dung sinh viên tìm hiểu

- Trí tuệ nhân tạo
- Trí thông minh nhân tạo
- Mục tiêu nghiên cứu của trí tuệ nhân tạo
- Các lĩnh vực nghiên cứu của trí tuệ nhân tạo
- Các bài toán điển hình áp dụng trí tuệ nhân tạo
- Một số hệ thống (máy) có trí tuệ nhân tạo
- Dự báo tương lai của trí tuệ nhân tạo

❖ Bài giảng này giới thiệu một số thuật toán cơ sở của một số lĩnh vực cụ thể của ngành trí tuệ nhân tạo

# **GREEDY ALGORITHM**

## **References:**

- Robert Sedgewick, Kevin Wayne, "Algorithms", Fourth edition, Addison-Wesley, 2011.

## **NỘI DUNG**

1. Giới thiệu
2. Một số ví dụ điển hình
3. Bài tập

# GIỚI THIỆU

## GIỚI THIỆU VỀ THUẬT TOÁN THAM LAM

- Tham lam (greedy) là một trong số các phương pháp giải bài toán tối ưu;
- Các thuật toán tham lam dựa vào sự đánh giá tối ưu cục bộ địa phương (local optimum) để đưa ra quyết định tức thì tại mỗi bước lựa chọn, với hy vọng cuối cùng sẽ tìm ra được phương án tối ưu tổng thể (global optimum).

## GIỚI THIỆU VỀ THUẬT TOÁN THAM LAM (...)

- Thuật toán tham lam có trường hợp luôn tìm ra đúng phương án tối ưu, có trường hợp không.
- Trong trường hợp thuật toán tham lam không tìm ra đúng phương án tối ưu, chúng ta thường thu được một phương án khả dĩ chấp nhận được.
- Với một bài toán có nhiều thuật toán để giải quyết, thông thường thuật toán tham lam có tốc độ tốt hơn hẳn so với các thuật toán tối ưu khác như quy hoạch động, nhánh cận, chia để trị,...

## SƠ ĐỒ THUẬT TOÁN THAM LAM

Giả sử nghiệm của bài toán có thể biểu diễn dưới dạng một vector  $(x_1, x_2, \dots, x_n)$ , mỗi thành phần  $x_i$  ( $i = 1, 2, \dots, n$ ) được chọn ra từ tập  $S_i$ . Mỗi nghiệm của bài toán  $X = (x_1, x_2, \dots, x_n)$ , được xác định “**độ tốt**” bằng một hàm  $f(X)$  và mục tiêu cần tìm nghiệm có giá trị  $f(X)$  càng lớn càng tốt (hoặc càng nhỏ càng tốt).

Tư tưởng của phương pháp tham ăn như sau: Ta xây dựng vector nghiệm X dần từng bước, bắt đầu từ vector không (). Giả sử đã xây dựng được  $(k-1)$  thành phần  $(x_1, x_2, \dots, x_{k-1})$  của nghiệm và khi mở rộng nghiệm ta sẽ chọn  $x_k$  “**tốt nhất**” trong các ứng cử viên trong tập  $S_k$  để được  $(x_1, x_2, \dots, x_k)$ . Việc lựa chọn như thế được thực hiện bởi một hàm chọn. Cứ tiếp tục xây dựng, cho đến khi xây dựng xong hết thành phần của nghiệm.

Lược đồ tổng quát của phương pháp tham ăn.

```
procedure Greedy;
begin
X:=∅;
i:=0;
while (chưa xây dựng xong hết thành phần của nghiệm) do
begin
    i:=i+1;
    <Xác định  $S_i$ >;
     $x \leftarrow \text{select}(S_i)$ ; // chọn ứng cử viên tốt nhất trong tập  $S_i$ 
end;
end;
```

Trong lược đồ tổng quát trên, Select là hàm chọn, để chọn ra từ tập các ứng cử viên  $S_i$  một ứng cử viên được xem là tốt nhất, nhiều hứa hẹn nhất.

Cần nhấn mạnh rằng, thuật toán tham ăn trong một số bài toán, nếu xây dựng được hàm thích hợp có thể cho nghiệm tối ưu. Trong nhiều bài toán, thuật toán tham ăn chỉ tìm được nghiệm gần đúng với nghiệm tối ưu.

### Lược đồ tổng quát thuật toán tham lam (ref)

```
void Greedy(A,S)//A là tập các ứng cử viên, S là tập nghiệm
{
S=∅;
while (A ≠ ∅)
{
    x=select(A); //chọn phần tử tốt nhất trong A
    A=A - {x};
    if (S ∪ {x} chấp nhận được)
        S= S ∪ {x};
}
}
```

- Cụm từ “thuật toán tham lam” có thể được hiểu là một dạng của “thuật toán heuristic” (sẽ được chúng tôi trình bày chi tiết hơn trong phần các thuật toán metaheuristic).

## MỘT SỐ DẠNG BÀI TOÁN THAM LAM ĐIỂN HÌNH

## BÀI 1.

Bài toán Tìm tập các đoạn thẳng không giao nhau

**Đầu vào:** Cho họ các đoạn thẳng  $C = \{(a_1, b_1), (a_2, b_2), \dots, (a_n, b_n)\}$ ; các điểm nằm trên trực hoành.

**Đầu ra:** Tập các đoạn thẳng không giao nhau có lực lượng lớn nhất.

### Greedy 1: Bắt đầu sớm thì chọn trước

- Sắp xếp các đoạn thẳng theo thứ tự tăng dần của đầu mút trái;
- Bắt đầu từ tập  $S$  là tập rỗng, ta lần lượt bổ sung các đoạn trong danh sách theo thứ tự đã sắp xếp vào  $S$  nếu nó không có điểm chung với bất cứ đoạn nào trong  $S$ ;
- Độ phức tạp của thuật toán là  $O(n \log n)$ .

## **Greedy 2: Đoạn thẳng ngắn nhất thì được lựa chọn trước**

- Sắp xếp các đoạn thẳng theo thứ tự tăng dần của độ dài;
- Bắt đầu từ tập  $S$  là tập rỗng, ta lần lượt bổ sung các đoạn thẳng theo thứ tự đã sắp vào  $S$  nếu nó không có điểm chung với bất cứ đoạn nào trong  $S$ ;

## **Greedy 3. Kết thúc sớm thì lựa chọn trước**

- Sắp xếp các đoạn thẳng theo thứ tự tăng dần của đầu mút phải;
- Bắt đầu từ tập  $S$  là tập rỗng, ta lần lượt xét các đoạn trong danh sách theo thứ tự đã sắp vào  $S$  nếu nó không có điểm chung với bất cứ đoạn nào trong  $S$ .
- Các thuật toán tham lam trên có tìm được lời giải tối ưu không ?

## Giải mỗi bài toán sau bằng 3 greedy trên

Test 1:

INPUT

12  
3 7;  
3 4;  
5 7;  
4 5;  
-5 -1;  
-2 0;  
1 5;  
-9 -6;  
-8 -7;  
2 6;  
1 7;  
-9 0;

Test 2:

INPUT  
7  
1 3;  
2 6;  
4 5;  
2 4;  
1 5;  
1 2;  
-6 1;

Test 3:

14  
(-2,8); (9,11); (0,8); (-5,-1);  
(-1,0); (6,10); (0,2);  
(-5,-4); (1,8); (0,4); (6,8);  
(4,10); (-6,-5); (-6,-3);

## BÀI 2. Bài toán ba lô

Có  $n$  món hàng; món hàng  $i$  có trọng lượng  $a_i$  và giá trị  $c_i$ ,  $i=1..n$ .

Cần chọn các món hàng này vào ba lô có trọng lượng là  $M$  sao cho tổng trọng lượng của các món hàng không quá  $M$ , đồng thời tổng giá trị của chúng là lớn nhất (các giá trị này có thể là số thực).

### Greedy 1:Sắp xếp các món hàng theo thứ tự giảm dần của giá trị

Lần lượt xét các món hàng theo thứ tự đã sắp, và chất món hàng đang xét vào túi nếu như trọng lượng còn lại ba lô đủ chứa nó (tức là tổng trọng lượng của các món hàng đã xếp vào túi và trọng lượng của món hàng đang xét là không vượt quá  $M$ ).

### Greedy 2:Sắp xếp các món hàng theo thứ tự tăng dần của trọng lượng.

Lần lượt xét các món hàng theo thứ tự đã sắp, và chất món hàng đang xét vào túi nếu như trọng lượng còn lại ba lô đủ chứa nó (tức là tổng trọng lượng của các món hàng đã xếp vào túi và trọng lượng của món hàng đang xét là không vượt quá  $M$ ).

**Greedy 3:**Sắp xếp các món hàng theo thứ tự giảm dần  
của giá trị một đơn vị trọng lượng ( $c_i/w_i$ ).

Lần lượt xét các món hàng theo thứ tự đã sắp, và chất  
món hàng đang xét vào túi nếu như trọng lượng còn lại  
ba lô đủ chứa nó (tức là tổng trọng lượng của các món  
hang đã xếp vào túi và trọng lượng của món hàng đang  
xét là không vượt quá  $M$ ).

### Giải bài toán sau bằng 3 greedy trên

Sử dụng chiến lược tham lam nêu trên, hãy giải quyết bài toán cụ thể sau:

$i$	1	2	3	4	5	6
$a[i]$	5.1	1.4	6.2	4.3	9.1	5.6
$c[i]$	7.4	5.7	3.2	6.6	14.7	8.3

Cho biết số thứ tự của các món hàng được chọn và tổng giá trị các món hàng được chọn khi  
 $w=12.5$ .

### BÀI 3.

Có  $n$  công việc được đánh số từ 1 đến  $n$  và có một máy để thực hiện. Biết  $p_i$  là **thời gian** cần thiết để hoàn thành công việc  $i$ ,  $d_i$  là **thời điểm** cần hoàn thành công việc  $i$ .

Máy bắt đầu hoạt động từ thời điểm 0. Mỗi công việc cần được thực hiện liên tục từ lúc bắt đầu cho đến khi kết thúc, không được phép ngắt quãng. Giả sử  $c_i$  là thời điểm hoàn thành công việc  $i$ . Khi đó, nếu  $c_i > d_i$  ta nói công việc  $i$  hoàn thành trễ hạn, còn nếu  $c_i \leq d_i$  thì ta nói công việc  $i$  hoàn thành đúng hạn.

Hãy tìm trình tự các công việc sao cho số công việc hoàn thành trễ hạn là ít nhất (hay số công việc hoàn thành đúng hạn là nhiều nhất).

### VÍ DỤ

Ví dụ: giả sử có 5 công việc với thời gian thực hiện và thời gian hoàn thành như sau:

$i$	1	2	3	4	5
$p_i$	6	3	5	7	2

$d_i$	8	4	15	20	3
-------	---	---	----	----	---

Nếu thực hiện theo thứ tự 1, 2, 3, 4, 5 thì sẽ có 3 công việc bị trễ hạn là công việc 2, 4 và 5. Còn nếu thực hiện theo thứ tự 5, 1, 3, 4, 2 thì chỉ có 1 công việc bị trễ hạn là công việc 2, đây là thứ tự thực hiện mà số công việc bị trễ hạn ít nhất (nghiệm tối ưu).

Thuật toán: Chọn các công việc theo thứ tự  $p_i$  tăng dần.

Ví dụ:  $(p_i, d_i)$

3, 10

2, 8

3, 4

1, 1

5, 9

2, 3

6, 24

10, 24

7, 15

9, 23

#### Bài 4. XẾP LỊCH THỰC HIỆN CÁC NHIỆM VỤ

Chúng ta có rất nhiều nhiệm vụ trong ngày. Giả sử có  $n$  nhiệm vụ và nhiệm vụ thứ  $i$  phải bắt đầu ngay sau thời điểm  $s_i$ , thực hiện liên tục và kết thúc tại thời điểm  $f_i$ . Có thể coi mỗi nhiệm vụ tương ứng với một khoảng thời gian thực hiện  $(s_i, f_i)$ .

Hãy chọn ra nhiều nhất các nhiệm vụ để làm, sao cho không có thời điểm nào chúng ta phải làm hai nhiệm vụ cùng lúc, hay nói cách khác, khoảng thời gian thực hiện hai nhiệm vụ bất kỳ là không giao nhau.

### Input

- Dòng 1 chứa số nguyên dương  $n \leq 10^5$
- $n$  dòng tiếp theo, dòng thứ  $i$  chứa hai số nguyên  $s_i, f_i$   
 $(0 \leq s_i < f_i \leq 10^9)$

### Output

Phương án chọn ra nhiều nhiệm vụ nhất để thực hiện

#### Sample Input

```
5
7 9
6 8
1 3
0 6
3 7
```

#### Sample Output

```
Task 3: (1, 3]
Task 5: (3, 7]
Task 1: (7, 9]
Number of selected tasks: 3
```

Thuật toán:

- Sắp xếp các nhiệm vụ theo thứ tự tăng dần của thời điểm kết thúc  $f[i]$ .
- Khởi tạo thời điểm  $FinishTime := 0$
- Duyệt các nhiệm vụ theo danh sách đã sắp xếp (nhiệm vụ kết thúc sớm sẽ được xét trước nhiệm vụ kết thúc muộn), nếu xét đến nhiệm vụ  $i$  có  $s[i] \geq FinishTime$  thì chọn ngay nhiệm vụ  $i$  vào phương án tối ưu và cập nhật  $FinishTime$  thành thời điểm kết thúc nhiệm vụ  $i$ :  
 $FinishTime := f[i]$

## BÀI 5. BÀI TOÁN PHÂN VIỆC

Giả sử có  $m$  máy như nhau được ký hiệu từ  $P_1, \dots, P_m$ .

Có  $n$  công việc  $J_1, \dots, J_n$  cần được thực hiện.

Các công việc có thể được thực hiện đồng thời và bất kỳ công việc nào cũng có thể chạy trên một máy nào đó.

Mỗi lần máy được cho thực hiện một công việc nó sẽ làm cho tới khi hoàn chỉnh. Công việc  $J_i$  có thời gian thực hiện là  $T_i$ .

Hãy tìm cách phân công sao cho tất cả các công việc được hoàn thành trong thời gian sớm nhất.

### Thuật toán 1

Giả sử  $L$  là một thứ tự các công việc cần được thực hiện.

Lặp lại các công việc sau cho đến khi tất cả các công việc đều được phân công: “Nếu có máy nào rảnh thì nạp công việc kế tiếp trong danh sách  $L$  vào (nếu có 2 hay nhiều máy cùng rảnh tại một thời điểm thì máy với chỉ số thấp sẽ được phân cho công việc)”.

### **Ví dụ 1:**

Giả sử có 3 máy  $P_1, P_2, P_3$  và 6 công việc  $J_1, J_2, J_3, J_4, J_5, J_6$ , với

$$T_i = (2, 5, 8, 1, 5, 1)$$

$$L = (J_2, J_5, J_1, J_4, J_6, J_3)$$

Thì phân công theo phương án này sẽ không tối ưu (thời gian hoàn thành các công việc là 12).

### **Thuật toán 2**

Gọi  $L^*$  là phương án mà các công việc được sắp theo thứ tự thời gian giảm dần.

Áp dụng như thuật toán 1 thì thời gian hoàn thành là 8.

## Ví dụ 2:

Cho hai máy  $P_1, P_2$  và 5 công việc  $J_1, J_2, J_3, J_4, J_5$ . Thời gian thực hiện các công việc đã được sắp giảm là 3,3,2,2,2.

Khi đó cách phân công công việc là:

$P_1: 3 \ 2 \ 2$

$P_2: 3 \ 2$

Thời gian hoàn thành là 7 (không tối ưu) do thời gian hoàn thành tối ưu là 6:

3 3

2 2 2

## Ví dụ bài toán phân việc

Một dịch vụ in ấn luận văn tốt nghiệp, có 3 nhân viên đánh máy và một quản lý. Dịch vụ nhận được yêu cầu đánh máy luận văn của sinh viên làm luận văn tốt nghiệp như sau:

Luận văn	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12
Số trang	240	180	100	140	100	160	60	140	140	200	80	100

Giả sử trong một giờ thì một nhân viên đánh máy được 10 trang (giả thiết các nhân viên có công suất đánh máy bằng nhau).

- Phân chia các luận văn cho 03 nhân viên đánh máy sao cho thời gian hoàn thành việc đánh máy các luận văn trên là sớm nhất.
- Trong trường hợp người quản lý cũng tham gia đánh máy, nhưng công suất của người quản lý chỉ bằng  $\frac{2}{3}$  công suất của nhân viên. Tìm cách chia các luận văn cho 3 nhân viên và người quản lý, sao cho thời gian hoàn thành việc đánh máy các luận văn là sớm nhất.

### Hướng dẫn giải

a.Theo bài ra, thời gian để các nhân viên đánh máy xong các luận văn tốt nghiệp là:

Luận án	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12
Thời gian	24	18	10	14	10	16	6	14	14	20	8	10

Sắp xếp các luận văn theo thời gian đánh máy giảm

Luận án	L1	L10	L2	L6	L4	L8	L9	L3	L5	L12	L11	L7
Thời gian	24	20	18	16	14	14	14	10	10	10	8	6

Lịch đánh máy các luận văn của nhân viên:

Nhân viên 1:L1(24), L8(14), L5(10),L11(8)

Nhân viên 2:L10(20), L4(14), L9(14), L7(6)

Nhân viên 3:L2(18), L6(16), L3(10), L12(10)

Theo lịch trên, thời gian để các nhân viên 1,2,3 hoàn thành công việc của mình lần lượt là:

56,54,54.

Vậy thời gian hoàn thành việc đánh máy cho các luận văn là 56.

b.Khi người quản lý tham gia, và có công suất đánh máy bằng 2/3 nhân viên. Ta có lịch sau;

Nhân viên 1:L1(24), L9(14), L11(8)

Nhân viên 2:L10(20), L8(14), L12(10)

Nhân viên 3:L2(18), L4(14),L5(10)

Quản lý:L6(24), L3(15), L7( 9)

Thời gian để các nhân viên và người quản lý hoàn thành công việc của mình lần lượt là 46,44,42,48

Vậy thời gian hoàn thành việc đánh máy cho các luận văn khi có người quản lý tham gia là 48.

## BÀI 6. ĐOẠN THẮNG PHỦ

Trên trục số cho  $n$  đoạn:  $[a_1, b_1], [a_2, b_2], \dots, [a_n, b_n]$ . Hãy chọn một số ít nhất trong số  $n$  đoạn để phủ hết đoạn  $[p, q]$ .

### Input

+ Dòng 1 chứa ba số nguyên  $n, p, q$  ( $1 \leq n \leq 10^5, 0 \leq p \leq q \leq 10^9$ )

+  $n$  dòng tiếp theo, dòng thứ  $i$  chứa hai số nguyên  $a_i, b_i$  ( $0 \leq a_i \leq b_i \leq 10^9$ )

### Output

Cách chọn ra ít nhất các đoạn để phủ đoạn  $[p, q]$

Sample Input	Sample Output
7 1 9	Selected Intervals:
6 10	Interval 3: [0, 3]
4 8	Interval 6: [2, 7]
0 3	Interval 1: [6, 10]
1 3	
2 4	
2 7	
5 9	

Thuật toán:

Sắp xếp danh sách các đoạn đã cho theo thứ tự không giảm của cận dưới.  
Đặt chỉ số đầu danh sách  $i:=1$

- Bước 1: Xét phần danh sách bắt đầu từ vị trí  $i$  gồm các đoạn có cận dưới  $\leq p$ . Tìm trong phần này để chọn ra đoạn có cận trên  $RightMost$  lớn nhất vào phương án tối ưu.
- Bước 2: Nếu  $RightMost \geq q$ , thuật toán kết thúc. Nếu không, đặt  $p := RightMost$ , cập nhật  $i$  thành chỉ số đứng sau đoạn vừa xét và lặp lại từ bước 1.

(Sau bước 2, chúng ta có thể thoải mái bỏ qua phần đầu danh sách gồm những đoạn đứng trước vị trí  $i$ , bởi tất cả những đoạn nằm trong phần này đều có cận trên  $\leq RightMost$  mà như đã nhận xét, đoạn tiếp theo cần chọn chắc chắn phải có cận trên  $> RightMost$ )

## BÀI 7. THUẬT TOÁN JOHNSON CHO 2 MÁY

Bài toán lập lịch gia công trên hai máy (*Two-machine flow shop model*): Có  $n$  chi tiết đánh số từ 1 tới  $n$  và hai máy  $A, B$ . Một chi tiết cần gia công trên máy  $A$  trước rồi chuyển sang gia công trên máy  $B$ . Thời gian gia công chi tiết  $i$  trên máy  $A$  và  $B$  lần lượt là  $a_i$  và  $b_i$ . Tại một thời điểm, mỗi máy chỉ có thể gia công một chi tiết. Hãy lập lịch gia công các chi tiết sao cho việc gia công toàn bộ  $n$  chi tiết được hoàn thành trong thời gian sớm nhất.

### Input

- Dòng 1 chứa số nguyên dương  $n \leq 10^5$
- Dòng 2 chứa  $n$  số nguyên dương  $a_1, a_2, \dots, a_n$  ( $\forall i: a_i \leq 100$ )
- Dòng 3 chứa  $n$  số nguyên dương  $b_1, b_2, \dots, b_n$  ( $\forall i: b_i \leq 100$ )

### Output

Lịch gia công tốt nhất có thể

Sample Input	Sample Output
4	Job 2: A[0, 1]; B[1, 3]
3 1 2 4	Job 1: A[1, 4]; B[4, 8]
4 2 1 3	Job 4: A[4, 8]; B[8, 11]
	Job 3: A[8, 10]; B[11, 12]
	Time = 12

Thuật toán:

Chia các chi tiết làm hai nhóm: Nhóm  $S_A$  gồm các chi tiết  $i$  có  $a_i < b_i$  và nhóm  $S_B$  gồm các chi tiết  $j$  có  $a_j \geq b_j$ . Sắp xếp các chi tiết trong nhóm  $S_A$  theo thứ tự tăng dần của các  $a_i$ , sắp xếp các chi tiết trong nhóm  $S_B$  theo thứ tự giảm dần của các  $b_i$  rồi nối hai danh sách đã sắp xếp lại.

## Bài 8. Bài toán người du lịch (TSP - Travelling Salesman Problem)

Có n thành phố (được đánh số từ 1 đến n), một người bán hàng xuất phát từ một thành phố, muốn đi qua tất cả các thành phố khác, mỗi thành phố một lần rồi quay về thành phố xuất phát. Giả thiết biết được chi phí đi từ thành phố i đến thành phố j là  $c_{ij}$ .

Hãy tìm một hành trình cho người bán hàng sao cho tổng chi phí theo hành trình này là thấp nhất.

TSP là bài toán thuộc lớp bài toán NP-khó.

### Thuật toán GTS1 (Greedy Traveling Saleman)

Input: n là số thành phố, u là đỉnh xuất phát u và c ,là ma trận chi phí.

Output: tour (thứ tự các thành phố đi qua; tính cả thành phố xuất phát ở cuối của hành trình).

```
cost – chí phí ứng với tour tìm được  
v=u;  
tour={u};  
cost=0;  
for i=1 to n  
{  
    đặt w là thành phố kè sau thành phố v.  
    tour=tour + {w};  
    cost=cost+c[v,w]  
    v=w;  
}  
tour=tour + {u};  
cost=cost+c[v,u]
```

### **Thuật toán GTS2 (Greedy Traveling Saleman)**

Input: n, c, p, v<sub>i</sub> ( i = 1..p) // v<sub>i</sub> là các thành phố cho trước hoặc cũng có thể được chọn ngẫu nhiên trong tập 1..p

Output: besttour, bestcost

```
bestcost=0;  
besttour={}  
for i=1 to p  
{  
    GTS1(vk); // suy ra được tour(vk) và cost(vk)  
    If cost(vk)<bestcost  
    {  
        bestcost=cost(vk)  
        besttour=tour(vk)  
    }  
}
```

### **Ví dụ**

Cho đồ thị có ma trận chi phí như sau:

$\infty$	20	42	31	6	24
10	$\infty$	17	6	35	18
25	5	$\infty$	27	14	9
12	9	24	$\infty$	30	12
14	7	21	15	$\infty$	38
40	15	16	5	20	$\infty$

Sử dụng thuật toán GTS1 để tìm hành trình bắt đầu tại các đỉnh v<sub>1</sub>=1; v<sub>2</sub>=3; v<sub>3</sub>=4; v<sub>4</sub>=5

Sử dụng thuật toán GTS2 để tìm hành trình tốt nhất với p=4 (v<sub>1</sub>=2; v<sub>2</sub>=3; v<sub>3</sub>=5; v<sub>4</sub>=6; bài toán dạng này có thể không cho cụ thể các v<sub>i</sub>, lúc đó vi có thể được chọn ngẫu nhiên).

**Hướng dẫn giải:**

GTS1(v<sub>1</sub>): 1 → 5 → 2 → 4 → 6 → 3 → 1

$$\text{cost}(v_1) = 6 + 7 + 6 + 12 + 16 + 25 = 72.$$

GTS1(v<sub>2</sub>): 3 → 2 → 4 → 1 → 5 → 6 → 3

$$\text{cost}(v_2) = 5 + 6 + 12 + 6 + 38 + 16 = 83.$$

GTS1(v<sub>3</sub>): 4 → 2 → 1 → 5 → 3 → 6 → 4

$$\text{cost}(v_3) = 9 + 10 + 6 + 21 + 9 + 5 = 60.$$

GTS1(v<sub>4</sub>): 5 → 2 → 4 → 1 → 6 → 3 → 5

$$\text{cost}(v_4) = 7 + 6 + 12 + 24 + 16 + 14 = 79.$$

**Hướng dẫn giải:**

Áp dụng thuật toán GTS1 như trên để tính

GTS1(v<sub>1</sub>): 2 → 4 → 1 → 5 → 3 → 6 → 2

$$\text{cost}(v_1) = 6 + 12 + 6 + 21 + 9 + 15 = 69$$

GTS1(v<sub>2</sub>): 3 → 2 → 4 → 1 → 5 → 6 → 3

$$\text{cost}(v_2) = 5 + 6 + 12 + 6 + 38 + 16 = 83.$$

GTS1(v<sub>3</sub>): 5 → 2 → 4 → 1 → 6 → 3 → 5

$$\text{cost}(v_3) = 7 + 6 + 12 + 24 + 16 + 14 = 79.$$

GTS1(v<sub>4</sub>): 6 → 4 → 2 → 1 → 5 → 3 → 6

$$\text{cost}(v_4) = 5 + 9 + 10 + 6 + 21 + 9 = 60.$$

Kết luận: Hành trình tốt nhất có chi phí là 60 với chi tiết tour như sau:

$$6 \rightarrow 4 \rightarrow 2 \rightarrow 1 \rightarrow 5 \rightarrow 3 \rightarrow 6$$

## BÀI 9. BÀI TOÁN TÔ MÀU

Cho  $n$  thành phố, hãy tô màu các thành phố này sao cho không có bất kỳ hai thành phố nào kề nhau được tô cùng một màu và số màu được tô là ít nhất có thể.

Dữ liệu vào được lưu trên một trận vuông  $c_{ij}$ .

Nếu  $c_{ij} = 1$  thì hai thành phố  $i,j$  là kề nhau,  $c_{ij}=0$  thì hai thành phố  $i,j$  không kề nhau.

TSP là bài toán thuộc lớp bài toán NP-Đầy đủ.

Hãy đề xuất thuật toán tham lam giải bài toán trên.

**Thuật toán tô màu** (thuật toán Welch-Powell)

Bước 1:Sắp xếp các đỉnh theo bậc giảm dần.

Bước 2:Dùng màu thứ nhất tô cho đỉnh có bậc cao nhất và các đỉnh khác có thể tô còn lại.

Bước 3:Dùng màu thứ hai tô cho đỉnh có bậc cao thứ nhất (còn lại) và các đỉnh khác có thể tô còn lại

Bước 4:Và cứ như thế... cho đến khi tất cả các đỉnh được tô màu hết

## Ví dụ (bài toán tô màu)

Có một hội thảo khoa học được tổ chức với 10 chủ đề khác nhau ký hiệu là: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10. Mỗi chủ đề được diễn ra trong một buổi, trong đó các chủ đề sau không được diễn ra đồng thời trong cùng một buổi: (1,2,3); (1,4,5); (3,4,6); (2,5,6); (1,7,8); (7,8,10); (4,7,9); (4,7); (5,8).

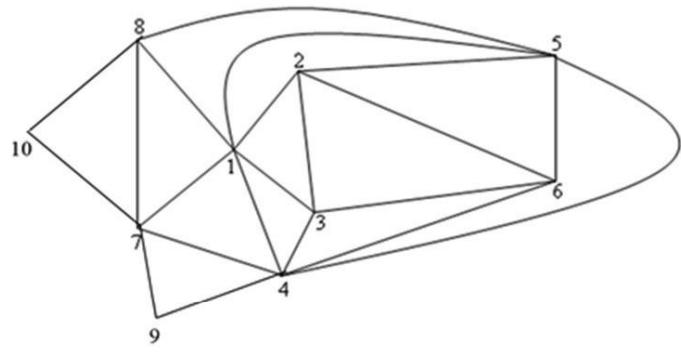
Hãy lập lịch tổ chức hội thảo sao cho số buổi hội thảo diễn ra là ít nhất có thể.

### Bước 1: Xây dựng đồ thị

Gọi mỗi chủ đề là một đỉnh của đồ thị

Nối hai chủ đề không diễn ra đồng thời là một cạnh của đồ thị

Theo bài ra ta có đồ thị sau:



### Bước 2: Tô màu theo thuật toán tham lam

Đỉnh	1	2	3	4	5	6	7	8	9	10
Bậc	6	4	4	5	5	4	3	3	2	2

Sắp xếp các đỉnh theo bậc giảm

Đỉnh	1	4	5	2	3	6	7	8	9	10
Bậc	6	5	5	4	4	4	3	3	2	2

Màu	1	2	3	4	5	6	7	8	9	10
1	x					x			x	x
2		x		x				x		
3			x		x		x			

### Bước 3: Lập lịch

Số buổi cần để tổ chức tất cả các buổi hội thảo là: 3

Buổi 1: Các chủ đề 1,6,9,10

Buổi 2: Các chủ đề 2,4,8

Buổi 3: Các chủ đề 3,5,7

## Ví dụ (bài toán tô màu)

Tại một cửa hàng sách, mới nhập về 12 quyển sách thuộc các loại sau:

Truyện cười: A, C, D, G.

Âm nhac: B. H. K.

Lịch sử: E. J. L.

**E-mail:** Khoa hoc: F, I.

Hãy sắp xếp những quyển sách này vào kệ sao cho số kệ sử dụng là ít nhất mà tuân theo các yêu cầu sau:

- Các quyển sách cùng loại không được để chung một kệ.
  - Quyển A không được để chung với sách khoa học.
  - Quyển L không được để chung với sách âm nhạc.

Bước 1: Lập ma trận kè (cũng có thể vẽ bằng đồ thị)

Bước 2: Tính bậc của từng đỉnh

Đỉnh	A	C	D	G	B	H	K	E	J	L	F	I
Bậc	5	3	3	3	3	3	3	2	2	5	2	2

Bước 3: Tô màu theo nguyên lý tham lam

Đỉnh	A	C	D	G	B	H	K	E	J	L	F	I
màu 1	1	4	4	4	1	4	4	1	4	4	4	4
màu 2	1	2	2	2	1	2	2	1	2	2	2	2
màu 3	1	2	3	3	1	2	3	1	2	3	2	3
màu 4	1	2	3	4	1	2	3	1	2	4	2	3

(Có thể thay – i bằng cách gạch một đường chéo qua i - ý nói ngăn cấm tô màu i)

Bước 4: Kết luận 12 quyển sách trên được xếp vào 4 kệ

Kệ 1: Gồm các quyển sách: A, B, E

Kệ 2: Gồm các quyển sách: C, H, J, F

Kệ 3: Gồm các quyển sách: D, K, I

Kệ 4: Gồm các quyển sách: G, L

## Bài tập (bài toán tô màu)

Tại vòng loại bảng B của một giải vô địch cờ vua gồm 8 kỳ thủ. Các kỳ thủ thi đấu vòng tròn để tính điểm. Biết rằng hiện tại:

Kỳ thủ 1 đã thi đấu với kỳ thủ 3 & 4

Kỳ thủ 4 đã thi đấu với kỳ thủ 2, 3 & 8

Kỳ thủ 5 đã thi đấu với kỳ thủ 6 & 8

Kỳ thủ 7 đã thi đấu với kỳ thủ 1, 4 & 5

Trong một buổi thi mỗi kỳ thủ chỉ thi đấu một trận. Hãy lập lịch thi đấu cho các trận còn lại sao cho số buổi cần thực hiện là ít nhất.

## Bài 10. Tìm cây khung nhỏ nhất

- Phát biểu Bài toán
- Thuật toán Prim
- Thuật toán Kruskal
- Đánh giá độ phức tạp của các thuật toán trên.
- Chứng minh tính đúng của các thuật toán trên.

## BÀI 11.

Cho  $n$  hình chữ nhật được đánh số từ 1 đến  $n$  có kích thước tương ứng  $(c_1, d_1), (c_2, d_2), \dots, (c_n, d_n)$  là các số nguyên dương; mỗi cặp  $(c_i, d_i)$  mô tả chiều dài, chiều rộng của hình chữ nhật thứ  $i$  ( $\text{chiều dài} \geq \text{chiều rộng}$ ). Một hình chữ nhật  $i$  được gọi là chứa hình chữ nhật  $j$  nếu mọi điểm của hình  $j$  đều thuộc hình  $i$ .

Hãy tìm một số  $k$  lớn nhất sao cho trong số  $k$  hình được chọn không có hai hình nào mà hình này chứa hình kia.

Ví dụ:

6,9

5,6

4,7

3,10

2,11

7,8

2,4

3,5

2,3

1,16

# BEST-FIRST SEARCH ALGORITHM

Thuật toán tìm kiếm ưu tiên tối ưu

(mục này chỉ xét phiên bản của BFS đó là thuật toán A<sup>KT</sup> - algorithm knowledgeable for tree)

## Bước 1: Khởi động

- Mọi đỉnh  $n$  là hàm  $f, g, h$  đều ẩn;
- Mở đỉnh đầu tiên  $S_o$ . Gán  $g(S_o)=0$ ;
- Sử dụng tri thức bổ sung ước tính  $h(S_o)$ ;
- Tính  $f(S_o) = g(S_o) + h(S_o)$ ;

## Bước 2: Lượng giá

- Chọn 1 đỉnh mở ứng với hàm  $f$  là min và gọi là đỉnh  $N$ ;
- Nếu  $N$  là đích  $\rightarrow$  dừng (đường đi từ đỉnh ban đầu đến đỉnh  $N$  là ngắn nhất và bằng  $g(N)$ );
- Nếu không tồn tại  $N$  thì cây biểu diễn vẫn đề không có đường đi tới mục tiêu  $\rightarrow$  dừng (bài toán không lời giải);
- Nếu tồn tại nhiều hơn 1 đỉnh  $N$  có cùng hàm  $f_{min}$  thì phải kiểm tra xem trong số đó có đỉnh nào là đích không;
  - + Nếu có  $\rightarrow$  dừng;
  - + Nếu không  $\rightarrow$  chọn ngẫu nhiên 1 trong các đỉnh đó và gọi đó là đỉnh  $N$ ;

## Bước 3: Phát triển

- Đóng đỉnh  $N$  và mở mọi đỉnh sau  $N$ ;
- Mọi đỉnh  $S$  sau  $N$ , tính;  
$$g(S) = g(N) + g(N - S);$$
- Dùng tri thức bổ sung để ước tính hàm  $h(S)$ ;
- Tính  $f(S) = g(S) + h(S)$ ;

## Bước 4: Quay lui

- Quay lại bước 2;

### Bài toán TACI

- Có  $n^2-1$  số mang các giá trị từ 1 tới  $n^2-1$  được sắp xếp vào một lưới các ô vuông kích thước  $n \times n$ . Mỗi số đó được gọi là một quân cờ và lưới ô đó được gọi là bàn cờ. Có một vị trí của bàn cờ bỏ trống. Mỗi lần di chuyển quân, người chơi được phép chuyển một quân ở vị trí ô tiếp giáp cạnh với ô trống vào ô trống.
- *Yêu cầu:* Từ một trạng thái ban đầu (a) (sự sắp xếp ban đầu của các quân trên bàn cờ), hãy thực hiện các nước đi hợp lệ để thu được trạng thái kết thúc (b) (trạng thái đích cần đạt được).

Dùng thuật toán A<sup>KT</sup> giải bài toán TACI với n = 3 như sau:

2	8	3
1	6	4
7		5

Trạng thái ban đầu (a)

1	2	3
8		4
7	6	5

Trạng thái kết thúc (b)

Với độ ước lượng  $H = \sum_{i=1}^{n^2-1} \delta(a_i, b_i)$  Trong đó  $\delta(a_i, b_i) = 0$  nếu  $a_i = b_i$  và  $\delta(a_i, b_i) = 1$  nếu  $a_i$  khác  $b_i$

- + Giả định bài toán của bài toán là đúng So,
- +  $g(S)$  là số bước cần để áp dụng di chuyển từ  $S_0$  đến  $S$ .
- \*  $h_i(S)$ : Giá trị của bước di chuyển thứ  $i$  của ván bài hiện tại sang của bài toán.
- \*  $f(S) = g(S) + h_i(S)$

So	2	8	3	$g=0$
	1	6	4	$h_i=4$
	7		5	$f=4$

$S_1$	$S_2$	$S_3$
$\begin{array}{ c c c } \hline 2 & 8 & 3 \\ \hline 1 & 6 & 4 \\ \hline 2 & 5 \\ \hline \end{array} g=1 f=5 j=6$	$\begin{array}{ c c c } \hline 2 & 8 & 3 \\ \hline 1 & 4 & 9 \\ \hline 2 & 6 \\ \hline \end{array} g=1 h=3 i=4 j=5$	$\begin{array}{ c c c } \hline 2 & 8 & 3 \\ \hline 1 & 6 & 4 \\ \hline 2 & 5 \\ \hline \end{array} g=1 h=5 j=6$
$S_4$	$S_5$	$S_6$
$\begin{array}{ c c c } \hline 2 & 8 & 3 \\ \hline 1 & 4 & 9 \\ \hline 2 & 5 \\ \hline \end{array} g=2 h=4 i=5 j=6$	$\begin{array}{ c c c } \hline 2 & 3 & 8 \\ \hline 1 & 8 & 4 \\ \hline 2 & 6 \\ \hline \end{array} g=2 h=3 i=5 j=5$	$\begin{array}{ c c c } \hline 2 & 8 & 3 \\ \hline 1 & 4 & 9 \\ \hline 2 & 5 \\ \hline \end{array} g=2 h=3 i=5 j=5$

$S_7$	$S_8$
$\begin{array}{ c c c } \hline 2 & 3 & 8 \\ \hline 1 & 8 & 4 \\ \hline 2 & 6 \\ \hline \end{array} g=3 h=2 i=4 j=5$	$\begin{array}{ c c c } \hline 2 & 3 & 8 \\ \hline 1 & 8 & 4 \\ \hline 2 & 6 \\ \hline \end{array} g=3 h=4 i=4 j=5$

$S_9$
$\begin{array}{ c c c } \hline 1 & 2 & 3 \\ \hline 8 & 4 & 9 \\ \hline 2 & 6 \\ \hline \end{array} g=4 h=1 i=5 j=5$

$S_{10}$	$S_{11}$
$\begin{array}{ c c c } \hline 1 & 2 & 3 \\ \hline 2 & 8 & 4 \\ \hline 6 & 5 \\ \hline \end{array} g=5 h=2 i=2 j=7$	$\begin{array}{ c c c } \hline 1 & 2 & 3 \\ \hline 8 & 4 & 9 \\ \hline 2 & 6 \\ \hline \end{array} g=5 h=0 i=0 j=5$

Vậy có  $g=5$  bước chuyển từ từ trạng thái đầu tiên  
trạng thái cuối:  $S_0 \rightarrow S_2 \rightarrow S_5 \rightarrow S_7 \rightarrow S_9 \rightarrow S_{11}$ .

## Bài toán THÁP HÀ NỘI

Giải bài toán tháp Hà Nội trong trường hợp  $n=3$  với cấu hình khởi đầu như sau:



A

B

C

Các trường hợp có thể ở cột C và giá trị h tương ứng:

1							
2	2		1				1
3	3	3	3		1	2	2

0    1    2    3    4    5    6    7

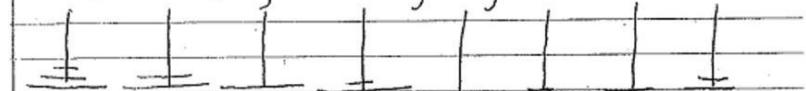
• Khi ban hát bài dân ca bài xoá là Sô.  
 g(S) là số bước di chuyển của đường đi ngắn.  
 (diết) là di chuyển Sô đến đỉnh S.

$$f(S) = g(S) + h(S)$$

• h(S) : Giá trị nhỏ nhất của đường đi từ Sô đến S.  
 (hiệu) là số bước di chuyển của đường đi ngắn.  
 g(S) là số bước di chuyển của đường đi ngắn.

$$f(S) = g(S) + h(S)$$

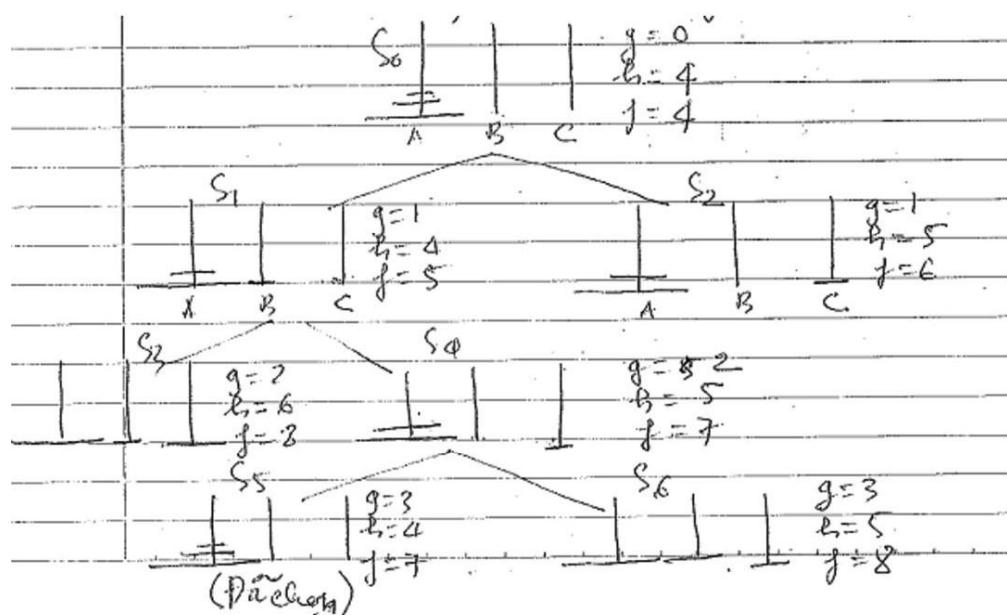
Các bước đi có thể có ở cột C và giá trị của  
 (hiệu), số bước di chuyển là số bước ngắn.

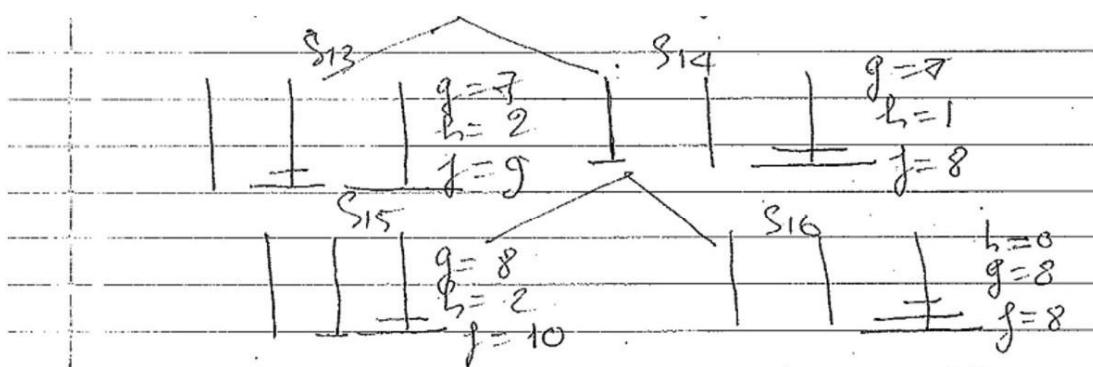
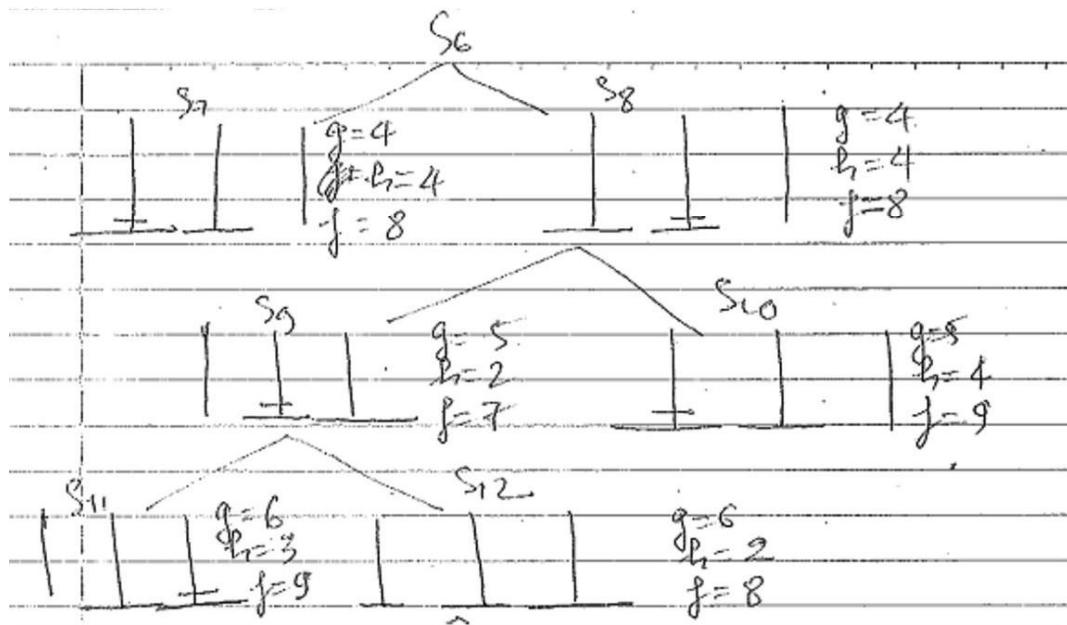


$h=0 \quad h=1 \quad h=2 \quad h=3 \quad h=4 \quad h=5 \quad h=6 \quad h=7$   
 (3 bước) (4 bước) (5 bước) (6 bước) (7 bước) (8 bước) (9 bước)

Theo bài ra, ta có cách biến đổi bài toán như sau:

và sau





Vay ee  $g=8$  bude clugen dē-di-là-hang-hai-bæ  
đâu rẽ hang-hai đičh:

$S_6 \rightarrow S_7 \rightarrow S_8 \rightarrow S_9 \rightarrow S_{12} \rightarrow S_{14} \rightarrow S_{16}$

# LOGIC MỆNH ĐỀ LOGIC VỊ TỪ

## Thuật toán Vương Hạo

- Bước 1: Phát biểu lại giả thiết và kết luận của bài toán dưới dạng chuẩn sau:  
$$GT_1, GT_2, \dots, GT_{n-1}, GT_n \rightarrow KL_1, KL_2, \dots, KL_{m-1}, KL_m$$

Trong đó các  $GT_i$  và  $KL_j$  được xây dựng từ các biến mệnh đề và các phép toán  $\wedge, \vee, \neg$ .
- Bước 2: Chuyển về các giá trị  $GT_i, KL_j$  có dạng phủ định.
- Bước 3: Thay phép toán  $\wedge$  ở  $GT_i$  và phép toán  $\vee$  ở  $KL_j$  bằng dấu “,”.

Bước 4:Nếu dòng hiện hành có một trong hai dạng sau:

Dạng 1:

$GT_1, GT_2, \dots, a \vee b, \dots, GT_{n-1}, GT_n \rightarrow KL_1, KL_2, \dots, KL_{m-1}, KL_m$

Thì thay bằng hai dòng:

$$\left\{ \begin{array}{l} GT_1, GT_2, \dots, a, \dots, GT_{n-1}, GT_n \rightarrow KL_1, KL_2, \dots, KL_{m-1}, KL_m \\ GT_1, GT_2, \dots, b, \dots, GT_{n-1}, GT_n \rightarrow KL_1, KL_2, \dots, KL_{m-1}, KL_m \end{array} \right.$$

Dạng 2:

$GT_1, GT_2, \dots, GT_{n-1}, GT_n \rightarrow KL_1, KL_2, \dots, a \wedge b, \dots, KL_{m-1}, KL_m$

Thì thay bằng hai dòng:

$GT_1, GT_2, \dots, GT_{n-1}, GT_n \rightarrow KL_1, KL_2, \dots, a, \dots, KL_{m-1}, KL_m$

$GT_1, GT_2, \dots, GT_{n-1}, GT_n \rightarrow KL_1, KL_2, \dots, b, \dots, KL_{m-1}, KL_m$

- Bước 5: Một dòng được chứng minh nếu tồn tại chung một mệnh đề ở cả hai vế.

- Bước 6:

- 6.a.Một vấn đề được giải quyết trọn vẹn nếu mọi dòng dẫn xuất biểu diễn ở dạng chuẩn được chứng minh.

- 6.b.Nếu một dòng không còn dấu liên kết  $\wedge$ ,  $\vee$  và cả hai vế không có chung mệnh đề nào thì dòng đó không được chứng minh.

## Thuật toán ROBINSON

Bước 1: Phát biểu lại giả thiết và kết luận bài toán dưới dạng chuẩn sau.

$$GT_1, GT_2, \dots, GT_n \rightarrow KL_1, KL_2, \dots, KL_m$$

Trong đó các  $GT_i$  và  $KL_i$  được xây dựng nhờ các biến mệnh đề và các phép toán  $\vee, \wedge, \neg$

Bước 2: Biến đổi dòng trên thành danh sách các mệnh đề

$$\{GT_1, GT_2, \dots, GT_n, \neg KL_1, \neg KL_2, \dots, \neg KL_m\}$$

Bước 3: Nếu trong danh sách các mệnh đề ở bước 2 có 2 mệnh đề đối ngẫu nhau (dạng  $\{a, \neg a\}$ ) thì vẫn đề được giải quyết xong, còn không thì chuyển sang bước 4.

Bước 4: Xây dựng 1 mệnh đề mới bằng cách tuyển 1 cặp mệnh đề trong danh sách các mệnh đề ở bước 2, nếu mệnh đề mới có các biến mệnh đề đối ngẫu nhau thì những biến đó được loại bỏ.

Bước 5: Bổ sung mệnh đề mới vào danh sách và loại bỏ 2 mệnh đề cũ vừa tạo thành mệnh đề mới ra khỏi danh sách.

Bước 6: Nếu không xây dựng thêm mệnh đề mới nào và trong danh sách các mệnh đề không có 2 mệnh đề đối ngẫu nhau thì vấn đề phát biểu ở dạng chuẩn bước 1 là sai

## Ví dụ

Xét các câu đúng:

$\alpha$ : “Nếu sinh viên giỏi Toán rời rạc thì giỏi Phân tích thiết kế thuật toán”.

$\beta$ : “Nếu sinh viên giỏi Phân tích thiết kế thuật toán thì giỏi Lập trình”.

$\delta$ : “Nếu sinh viên không giỏi Lập trình thì không giỏi Toán rời rạc” .

Bài toán đặt ra là: Từ (hội) các câu *đúng*  $\alpha, \beta$  chứng minh  $\delta$ .

- Xây dựng bài toán trên bằng các biểu thức logic mệnh đề.
- Hãy chứng minh bài toán bằng phương pháp Vương Hạo.
- Hãy chứng minh bài toán bằng phương pháp Robinson.

## Hướng dẫn giải

Đặt các biến mệnh đề:

$p$ : ‘Giỏi Toán rời rạc’

$q$ : ‘Giỏi Phân tích thiết kế thuật toán’

$r$ : ‘Giỏi Lập trình’

Theo bài ra, ta có:

$$\alpha: p \rightarrow q$$

$$\beta: q \rightarrow r$$

$$\delta: \bar{r} \rightarrow \bar{p}$$

a.Xây dựng bài toán bằng biểu thức logic mệnh đề:

$$((p \rightarrow q) \wedge (q \rightarrow r)) \rightarrow (\bar{r} \rightarrow \bar{p})$$

b.Chứng minh bài toán bằng phương pháp Vương Hạo

$$\begin{aligned} & ((p \rightarrow q) \wedge (q \rightarrow r)) \rightarrow (\bar{r} \rightarrow \bar{p}) \\ \Leftrightarrow & ((\bar{p} \vee q) \wedge (\bar{q} \vee r)) \rightarrow (r \vee \bar{p}) \\ \Leftrightarrow & \bar{p} \vee q, \bar{q} \vee r \rightarrow r, \bar{p} \\ \Leftrightarrow & \begin{cases} \bar{p}, \bar{q} \vee r \rightarrow r, \bar{p} & (cm) \\ q, \bar{q} \vee r \rightarrow r, \bar{p} & (1) \end{cases} \\ (1) \Leftrightarrow & \begin{cases} q, r \rightarrow r, \bar{p} & (cm) \\ q, \bar{q} \rightarrow r, \bar{p} & (2) \end{cases} \\ (2) \Leftrightarrow & q \rightarrow r, \bar{p}, q \quad (cm) \end{aligned}$$

Vậy bài toán được chứng minh

c.Chứng minh bài toán bằng phương pháp Robinson

$$\begin{aligned} & ((p \rightarrow q) \wedge (q \rightarrow r)) \rightarrow (\bar{r} \rightarrow \bar{p}) \\ \Leftrightarrow & ((\bar{p} \vee q) \wedge (\bar{q} \vee r)) \rightarrow (r \vee \bar{p}) \\ \Leftrightarrow & \bar{p} \vee q, \bar{q} \vee r \rightarrow r, \bar{p} \\ \Leftrightarrow & \bar{p} \vee q, \bar{q} \vee r, \bar{r}, p \text{ (1)} \end{aligned}$$

Chọn cặp các mệnh đề  $\bar{p} \vee q, \bar{q} \vee r$

$$(1) \Leftrightarrow \bar{p} \vee r, \bar{r}, p \text{ (2)}$$

Chọn cặp các mệnh đề  $\bar{p} \vee r, \bar{r}$

$$(2) \Leftrightarrow \bar{p}, p \text{ đây là cặp mệnh đề đối ngẫu.}$$

Theo thuật toán Robinson, bài toán đã được chứng minh.

## BIỂU DIỄN TRI THỨC BẰNG CÁC LUẬT DẪN

## Trò chơi đong nước

Cho hai bình có dung tích a và b lít, làm thế nào để đong được chính xác c lít ?

Các luật của bài toán đong nước

L1.Nếu bình a lít đầy thì đổ hết nước bình a lít đi.

L2.Nếu bình b lít rỗng thì đổ đầy nước vào bình b lít.

L3.Nếu bình a lít không đầy và bình b lít không rỗng thì đổ nước từ bình b lít sang bình a lít (cho tới khi bình a lít đầy hoặc bình b lít hết nước).

\*Điều kiện của a,b,c để bài toán đong nước là có lời giải ?

## Ví dụ

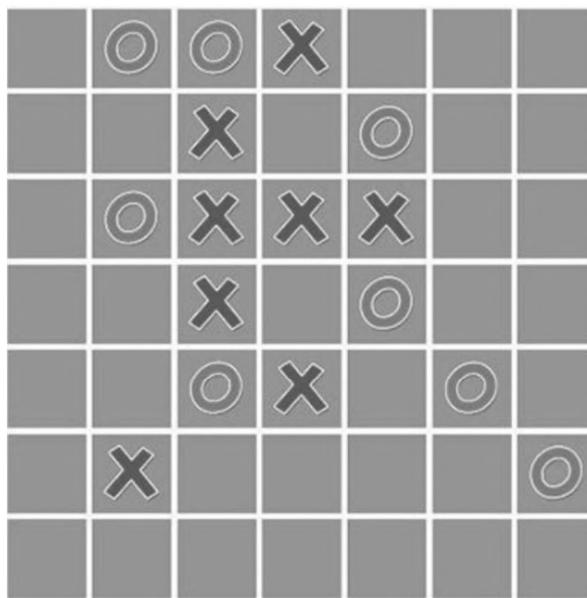
Áp dụng giải bài toán đong nước khi bình 1 đựng 8 lít, bình 2 đựng 10 lít và số nước cần đong là 6 lít. Yêu cầu mô tả số nước ở mỗi bình sau khi áp dụng mỗi luật và hãy ghi rõ đó là luật thứ mấy trong các luật đã nêu ?

Thứ tự	A	b	Luật	Ghi chú
0	0	0		Ban đầu các bình rỗng
1	0	10	2	
2	8	2	3	
3	0	2	1	
4	2	0	3	
5	2	10	2	
6	8	4	3	
7	0	4	1	
8	4	0	3	
9	4	10	2	
10	8	6	3	Dừng

## Trò chơi NIM

Có 3 đống sỏi, mỗi đống sỏi có  $n_1, n_2$  và  $n_3$  viên. Người chơi đến lượt mình được bốc từ một đống bất kỳ một số viên sỏi bất kỳ ( $>0$ ),  $a_i$  không còn gì để bốc là thua, hãy lập trình cho trò chơi NIM.

## Trò chơi caro (x,o)



## Bài tập

**BT1.**Sử dụng thuật toán Vương Hạo và Robinson :

a.Cho cơ sở tri thức KB1 ={ $(a \wedge b) \rightarrow c$ ,  $(b \wedge c) \rightarrow d$  ,  $\neg d$  }

Kiểm tra xem  $a \rightarrow b$  có được suy ra từ cơ sở tri thức trên hay không ?

b.Cho cơ sở tri thức KB2 ={ $(q \wedge p) \rightarrow s$  ,  $(q \wedge s) \rightarrow r$  ,  $q$ ,  $p$ }

Kiểm tra xem  $s$  có được suy ra từ cơ sở tri thức trên hay không ?

**BT2.**

Xét các câu:

- $\alpha$ : “Nếu trời mưa thì Lan mang theo dù”
- $\beta$ : “Nếu Lan mang theo dù thì Lan không bị ướt”
- $\delta$ : “Nếu trời không mưa thì Lan không bị ướt”
- $\lambda$ : “Lan không bị ướt”

Bài toán đặt ra là: Từ các câu *đúng*  $\alpha, \beta, \delta$  chứng minh  $\lambda$ .

- a.Xây dựng bài toán trên bằng các biểu thức logic mệnh đề.
- b.Hãy nêu ngắn gọn một ưu điểm của các thuật toán Vương Hạo và Robinson so với các thuật toán lập bảng chân trị, chứng minh tương đương hoặc chứng minh phản chứng trong việc chứng minh một biểu thức logic..
- c.Chứng minh bài toán bằng các phương pháp Vương Hạo và Robinson

**BT3.**

- a.Có một người nông dân nuôi bò sữa. Một hôm, có một người khách muốn mua đúng 8 lít sữa, nhưng người nông dân chỉ có một can dung tích 16 lít đựng đầy sữa, một can dung tích 7 lít (rỗng) và một can dung tích 3 lít (rỗng). Chỉ được sử dụng 3 can trên của mình, hỏi làm thế nào để người nông dân lấy ra được 8 lít sữa cho khách mua hàng ?

b.Có 3 bình nước: Bình thứ nhất dung tích 12 lít, bình thứ hai dung tích 7 lít, bình thứ ba dung tích 5 lít. Từ bình thứ nhất chứa đầy nước, bình thứ hai và và bình thứ ba trống, không dùng thêm dụng cụ chứa nước nào khác hãy lấy ra hai phần, mỗi phần 6 lít nước.

## MỞ ĐẦU VỀ MÁY HỌC

## Thuật toán Quinlan

- Cho một bảng quan sát (cơ sở dữ liệu) là một tập hợp các mẫu với các thuộc tính nhất định của các đối tượng nào đó.
- Sử dụng một độ đo để định lượng và đề ra một tiêu chuẩn nhằm chọn lựa một thuộc tính mang tính chất phân loại để phân bảng này thành các bảng con nhỏ hơn. Từ các bảng con này dễ dàng phân tích tìm ra qui luật chung.
- Từ đó thiết lập cây quyết định cho thấy thứ tự của thuộc tính đang xét.
- Tìm cây quyết định, xây dựng bộ luật, và đưa ra quyết định một số mẫu cụ thể.

Cho cơ sở dữ liệu gồm các mẫu sau:

Mẫu	Thời tiết	Lá cây	Nhiệt độ	Quyết định (mùa)
1	Mưa	Rụng	Thấp	Đông
2	Nắng	Xanh	Trung bình	Xuân
3	Nắng	Vàng	Trung bình	Thu
4	Nắng	Xanh	Cao	Hè
5	Nắng	Rụng	Thấp	Đông
6	Tuyết	Vàng	Thấp	Đông
7	Mưa	Rụng	Trung bình	Thu
8	Mưa	Xanh	Cao	Hè
9	Tuyết	Xanh	Thấp	Đông
10	Tuyết	Rụng	Thấp	Đông
11	Mưa	Vàng	Trung bình	Thu
12	Mưa	Xanh	Trung bình	Xuân
x	Mưa	Vàng	Cao	?
y	Tuyết	Rụng	Trung bình	?
z	Tuyết	Vàng	Trung bình	?

- a. Từ mẫu 1 đến mẫu 12 hãy rút ra bộ luật cho sự quyết định Mùa theo thuật toán Quinlan.
- b. Áp dụng cho biết kết quả các mẫu x,y,z|

## Hướng dẫn giải

Bước 0: Gọi vecttơ độ đo  $v=(Xuân, Hè, Thu, Đông)$

Bước 1: Tính vecttơ độ đo của các thuộc tính ở CSDL ban đầu

+ Thuộc tính **thời tiết**

$$V_{(\text{thời tiết=nắng})} = (1/4, 1/4, 1/4, 1/4);$$

$$V_{(\text{thời tiết=tuyết})} = (0, 0, 0, 1);$$

$$V_{(\text{thời tiết=mưa})} = (1/5, 1/5, 2/5, 1/5);$$

+ Thuộc tính lá cây

$$V_{(\text{lá cây=Vàng})} = (0, 0, 2/3, 1/3);$$

$$V_{(\text{lá cây=Xanh})} = (2/5, 2/5, 0, 1/5);$$

$$V_{(\text{lá cây=Rụng})} = (0, 0, 1/4, 3/4);$$

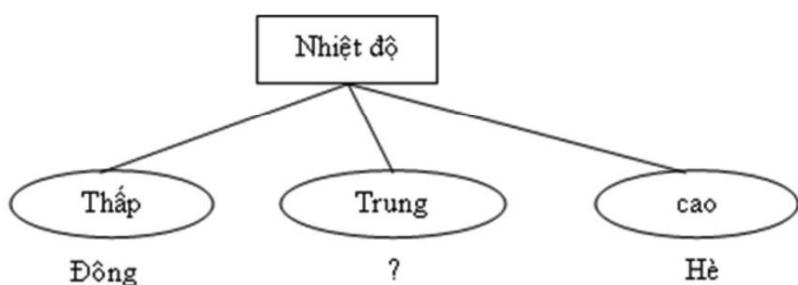
+ Thuộc tính Nhiệt độ

$$V_{(\text{nhiệt độ=Trung bình})} = (2/5, 0, 3/5, 0);$$

$$V_{(\text{nhiệt độ=thấp})} = (0, 0, 0, 1);$$

$$V_{(\text{nhiệt độ=cao})} = (0, 1, 0, 0);$$

Chọn thuộc tính Nhiệt độ làm thuộc tính phân loại:



## Bước 2: CSDL ứng với Nhiệt độ bằng Trung bình

#	Thời tiết	Lá cây	Quyết định (Mùa)
2	Nắng	Xanh	Xuân
3	Nắng	Vàng	Thu
7	Mưa	Rụng	Thu
11	Mưa	Vàng	Thu
12	Mưa	Xanh	Xuân

Tính vector độ đo của các thuộc tính

+ Thuộc tính **thời tiết**

$$V_{(\text{thời tiết}=\text{nắng})} = (1/2, 0, 1/2, 0);$$

$$V_{(\text{thời tiết}=\text{mưa})} = (1/3, 0, 2/3, 0);$$

+ Thuộc tính lá cây

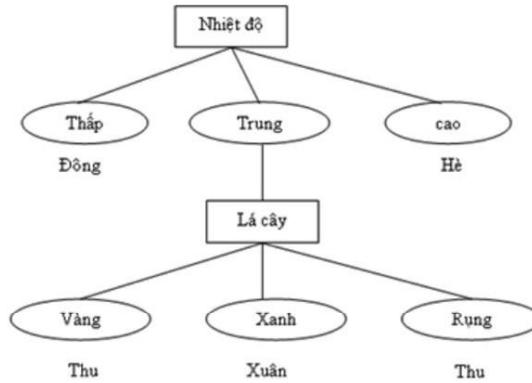
$$V_{(\text{lá cây}=\text{Vàng})} = (0, 0, 1, 0);$$

$$V_{(\text{lá cây}=\text{Xanh})} = (1, 0, 0, 0);$$

$$V_{(\text{lá cây}=\text{Rụng})} = (0, 0, 1, 0);$$

Chọn thuộc tính Lá cây làm thuộc tính phân loại

Từ các bước trên, ta có cây quyết định sau:



**Bước 3: Tập luật** Từ cây quyết định trên, ta có tập luật sau:

- Luật 1: Nếu nhiệt độ thấp thì mùa Đông
- Luật 2: Nếu nhiệt độ cao thì mùa Hè
- Luật 3: Nếu nhiệt độ trung bình và lá cây Xanh thì mùa Xuân
- Luật 4: Nếu nhiệt độ trung bình và lá cây màu Vàng hoặc lá rụng thì mùa Thu

#	Thời tiết	Lá cây	Nhiệt độ	Quyết định (Mùa)	Luật
x	Mưa	Vàng	Cao	Hè	Luật 2
y	Tuyết	Rụng	Trung bình	Thu	Luật 4
z	Tuyết	Vàng	Trung bình	Thu	Luật 4

## Thuật toán ILA (Inductive Learning Algorithm)

- Bước 1: Chia bảng có chứa m mẫu thành n bảng con. Một bảng ứng với một giá trị có thể có của thuộc tính lớp. (Từ bước 2 đến bước 8 sẽ được lặp lại cho mỗi bảng)
- Bước 2: Khởi tạo số lượng thuộc tính kết hợp j với j = 1.
- Bước 3: Với mỗi bảng con đang xét, phân chia các thuộc tính của nó thành một danh sách các thuộc tính kết hợp, mỗi thành phần của danh sách có j thuộc tính phân biệt.
- Bước 4: Với mỗi kết hợp các thuộc tính trong danh sách trên, đếm số lần xuất hiện các giá trị cho các thuộc tính trong kết hợp đó ở các dòng chưa bị khóa của bảng đang xét nhưng nó không được xuất hiện cùng giá trị ở những bảng con khác. Chọn ra một kết hợp trong danh sách sao cho nó có giá trị tương ứng xuất hiện nhiều nhất và được gọi là Max\_combination.

- Bước 5: If max\_combination = 0 thì j = j + 1 quay lại bước 3.
- Bước 6: Khóa các dòng ở bảng con đang xét mà tại đó nó có giá trị bằng với giá trị tạo ra max\_combination.
- Bước 7: Thêm vào R luật mới với giả thiết là max\_combination các thuộc tính và giá trị tương ứng phân biệt và kết nối các bộ này bằng AND, kết luận của luật là giá trị của thuộc tính quyết định tương ứng với bảng con này.
- Bước 8: Nếu tất cả các dòng đều khóa
  - + Nếu còn bảng con thì qua bảng con tiếp theo và quay lại bước 2.
  - + Ngược lại chấm dứt thuật toán
  - + Ngược lại (nghĩa là vẫn còn dòng chưa khóa trong bảng con đang xét) thì quay lại bước 4.

## Thuật toán phân cụm dữ liệu

Trong mặt phẳng tọa độ OXY, cho  $n$  điểm; điểm thứ  $i$  có tọa độ  $(x_i, y_i)$ ,  $i=1..n$ . Cho trước một số  $k$  nguyên dương ( $1 \leq k \leq n$ ). Hãy phân cụm  $n$  điểm trên vào  $k$  cụm sao cho một điểm nằm ở cụm nào thì khoảng cách của nó đến tâm của cụm chứa nó là nhỏ nhất so với đến các tâm của các cụm không chứa nó.

(Giả thiết rằng ban đầu tâm của mỗi cụm trùng với một trong các điểm đã cho).

## (THUẬT TOÁN K-mean)

**Bước 1:** Chọn  $k$  tâm ngẫu nhiên từ  $n$  điểm đã cho ban đầu.  
(ban đầu mỗi cụm có 1 đỉnh duy nhất là tâm của cụm đó)

**Lặp lại bước sau đến khi điều kiện dừng được thỏa:**

**Bước 2:** Duyệt các đỉnh  $i$  với  $i = 1 .. n$   
Đỉnh  $i$  gần với tâm nào nhất thì đưa đỉnh  $i$  vào cụm chứa tâm đó  
**Bước 3:** Cập nhật tâm các cụm theo cách sau:  
Giả sử tâm của cụm  $k$  nào đó  $(x_k, y_k)$  thì  $x_k =$  trung bình cộng  
hoành độ các điểm trong cụm đó; tương tự với  $y_k$

**Điều kiện dừng của thuật toán:**

Khi trong một lần lặp nào đó mà tất cả các cụm đều giữ nguyên  
đỉnh

Bài tập  
(Quinlan, ILA, ID)

Cho bảng quan sát sau:

STT	Quang cảnh	Nhiệt độ	Độ ẩm	Gió	QUYẾT ĐỊNH
1	Mưa	Âm	Cao	Nhẹ	Không
2	Mưa	Lạnh	Thấp	Nhẹ	Đi
3	Nắng	Âm	Thấp	Nhẹ	Đi
4	Mưa	Âm	Thấp	Mạnh	Đi
5	Mưa	Nóng	Cao	Nhẹ	Không
6	Mưa	Nóng	Cao	Mạnh	Không
7	Nhiều mây	Nóng	Cao	Nhẹ	Đi
8	Nắng	Âm	Cao	Nhẹ	Đi
9	Nhiều mây	Âm	Cao	Mạnh	Đi
10	Nắng	Lạnh	Thấp	Nhẹ	Đi
11	Nắng	Lạnh	Thấp	Mạnh	Không
12	Nhiều mây	Lạnh	Thấp	Mạnh	Đi
x	Mưa	Lạnh	Cao	Mạnh	?
y	Nắng	Âm	Thấp	Mạnh	?
z	Nhiều mây	Lạnh	Thấp	Mạnh	?

a. Áp dụng thuật toán QuinLan để xác định thời tiết như thế nào thì **đi / không** cho các mẫu từ 1 đến 12.

b. Áp dụng cho biết kết quả các mẫu x, y và z ?

## Bài tập (thuật toán phân cụm)

a. Trong mặt phẳng tọa độ  $OXY$ , cho 8 điểm  $A(1,4); B(-1,1); C(-1,2); D(-3,2); E(2,6); F(3,5); G(2,-2); H(0,6); I(3,-1)$  và số cụm  $k=3$  (tâm mỗi cụm ban đầu lần lượt tại các điểm  $D, E, G$ ).

Hãy sử dụng thuật toán  $k$ -mean giải bài toán phân cụm dữ liệu trên. Kết quả cần các thông tin: Tâm mỗi cụm và liệt kê các điểm nằm trong mỗi cụm đó.

## Bài tập (thuật toán phân cụm)

b.

10 //n=10

-2 -1

-1 7

1 -1

-1 -4

1 7

11 6

9 4

3 5

7 3

7 4

3 // k=3 cụm

# METAHEURISTIC ALGORITHMS

### References:

1. Xin-She Yang (2010). Engineering optimization. WILEY, pp.21-137.
2. Xin-She Yang (2010). Nature-inspired metaheuristic algorithms. LUNIVER Press, pp.53–62.
3. S.N. Sivanandam, S.N. Deepa (2008). Introduction to genetic algorithms. Springer, pp.15–82.

## NỘI DUNG

1. Bài toán tối ưu
2. Heuristic & Metaheuristic algorithm
3. Genetic algorithm (GA)
4. Bee algorithm (BA)
5. Local search algorithm (LS)
6. Variable neighborhood search algorithm (VNS)
7. Bài tập

## BÀI TOÁN TỐI ƯU

## ĐỊNH NGHĨA

Cho  $(S, f)$  là bài toán tối ưu, trong đó  $S \neq \emptyset$  biểu diễn cho không gian lời giải (còn được gọi là không gian tìm kiếm của bài toán) và  $f$  là hàm mục tiêu:

$$f: S \rightarrow R$$

Theo đó, việc giải bài toán tối ưu là quá trình tìm ra tập chứa giá trị các biến quyết định sao cho lời giải tìm được được biểu diễn bởi những giá trị  $x^* \in S$  thỏa mãn bất đẳng thức:

$$f(x^*) \leq f(x), \forall x \in S. \quad (\text{dạng tìm cực tiểu})$$

## VÍ DỤ 1

### Tìm cực tiểu

$$f(x_1, x_2) = 150x_1 + 175x_2 \rightarrow \min$$

Với điều kiện:

$$7x_1 + 11x_2 \leq 77$$

$$10x_1 + 8x_2 \leq 80$$

$$x_1 \leq 9$$

$$x_2 \leq 6$$

$$x_1, x_2 \geq 0$$

## VÍ DỤ 2

Tìm cực đại

$$f(x_1, x_2) = 2x_1 + x_2 \rightarrow \max$$

Với điều kiện:

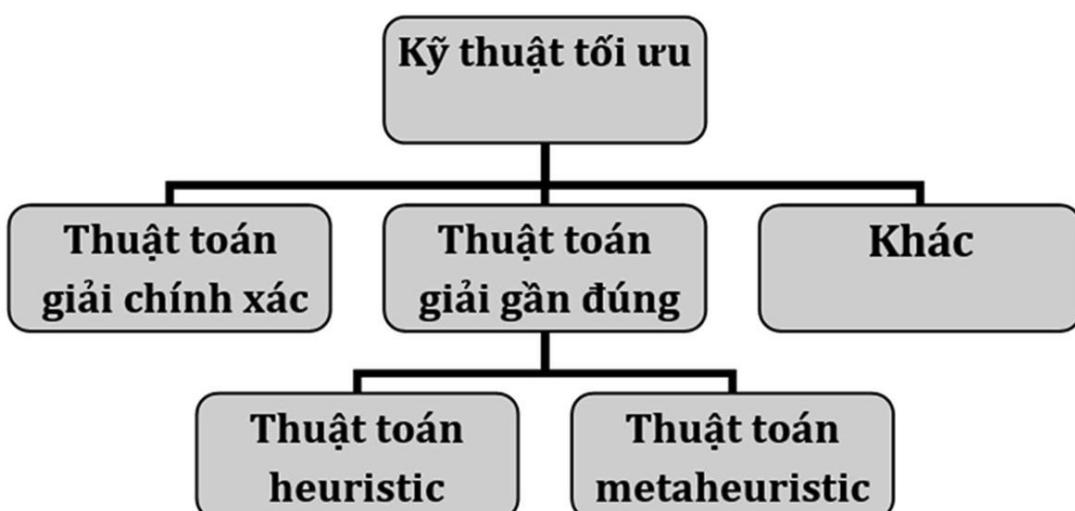
$$x_1 + x_2 \leq 6$$

$$x_1 - x_2 \leq 2$$

$$x_1 + 2x_2 \leq 10$$

$$x_1 - 2x_2 \leq 1$$

## CÁC KỸ THUẬT GIẢI BÀI TOÁN TỐI ƯU



## **HEURISTIC & METAHEURISTIC ALGORITHM**

### **THUẬT TOÁN HEURISTIC**

- Thuật toán heuristic chỉ những kinh nghiệm riêng biệt để tìm kiếm lời giải cho một bài toán tối ưu cụ thể.
- Thuật toán heuristic thường tìm được lời giải có thể chấp nhận được trong thời gian cho phép nhưng không chắc đó là lời giải tối ưu.
- Cấu trúc của một thuật toán heuristic thường là đơn giản và không quá khó để đề xuất.

## THUẬT TOÁN METAHEURISTIC

- Thuật toán metaheuristic sử dụng nhiều heuristic kết hợp với các kỹ thuật phụ trợ nhằm khai phá không gian tìm kiếm;
- Thuật ngữ “metaheuristic” thường đề cập đến các thuật toán dạng cá thể như local search, simulated annealing, tabu search,... hoặc dạng quần thể như genetic algorithm, particle swarm optimization (PSO), ant, bee, bat,...

## CÁC YẾU TỐ CHUNG CỦA MỘT THUẬT TOÁN METAHEURISTIC

- Cách thức tạo lời giải ban đầu,
- Điều kiện dừng,
- Cách thức mã hóa và giải mã lời giải,
- Chiến lược tăng cường hóa,
- Chiến lược đa dạng hóa,...

## SƠ ĐỒ THUẬT TOÁN METAHEURISTIC

- Khởi tạo lời giải ban đầu (lời giải có thể là một cá thể/hoặc một quần thể);
- Định nghĩa các chiến lược tăng cường hóa;
- Định nghĩa các chiến lược đa dạng hóa;
- Tính độ thích nghi cho mỗi lời giải;
- **while** (*điều kiện dừng chưa thỏa{*
  - Thực hiện các chiến lược tăng cường hóa;
  - Thực hiện các chiến lược đa dạng hóa;
  - Cập nhật lời giải tốt nhất cho đến thời điểm hiện tại;*}*)
- Trả về lời giải tốt nhất tìm được;

## KHỞI TẠO LỜI GIẢI BAN ĐẦU

- Khởi tạo ngẫu nhiên;
- Khởi tạo dựa vào một số heuristic riêng của bài toán;

## CHIẾN LƯỢC TĂNG CƯỜNG HÓA LỜI GIẢI

- Tăng cường hóa nhằm mục đích khai thác sâu hơn các vùng không gian lời giải tiềm năng để hy vọng tìm được lời giải có chất lượng tốt hơn,
- Tăng cường hóa thường được thực hiện nhờ áp dụng các tiếp cận tham lam (tìm kiếm được hướng đến phương án lân cận tốt nhất),
- Nếu tính tăng cường quá cao, nó có thể nảy sinh vấn đề hội tụ sớm; các lời giải tối ưu cục bộ tìm được có độ lệch cao hoặc vô nghĩa. Ngược lại, nếu tính tăng cường quá thấp, quá trình hội tụ sẽ chậm,

## CHIẾN LƯỢC TĂNG CƯỜNG HÓA LỜI GIẢI (....)

- Tính tăng cường hóa cho lời giải thường được thực hiện dựa vào các chiến lược tìm kiếm lân cận,
- Tùy theo đặc tính của bài toán mà có thể áp dụng các cách thức tìm kiếm lân cận khác nhau,
- Chiến lược tăng cường hóa được sử dụng xuyên suốt trong quá trình mà thuật toán diễn ra.

## CHIẾN LƯỢC ĐA DẠNG HÓA LỜI GIẢI

- Đa dạng hóa nhằm khai phá những vùng không gian tìm kiếm mới nhằm tránh việc tìm kiếm rơi vào bẫy tối ưu cục bộ,
- Đa dạng hóa thường được thực hiện qua việc sử dụng tính ngẫu nhiên và có thể kết hợp với một số heuristic riêng biệt cho từng bài toán,

## CHIẾN LƯỢC ĐA DẠNG HÓA LỜI GIẢI (....)

- Nếu tính đa dạng quá cao thì sẽ có nhiều vùng không gian lời giải được khai phá ngẫu nhiên, theo đó sẽ làm chậm quá trình hội tụ của thuật toán,
- Nếu tính đa dạng quá thấp thì không gian lời giải được khai phá sẽ bị giới hạn, các lời giải tìm được có khuynh hướng hội tụ cục bộ hoặc vô nghĩa,

## CHIẾN LƯỢC ĐA DẠNG HÓA LỜI GIẢI (....)

- Một lời giải khi nó được cho thực hiện liên tiếp việc tìm kiếm lân cận theo một cách thức nào đó, thì sau một số lần lặp, chất lượng của lời giải đó sẽ không còn cải thiện được nữa (lời giải được gọi là đã khai thác cạn); nếu tiếp tục tìm kiếm lân cận theo hướng đó thì sẽ bế tắc,
- Chiến lược đa dạng hóa thường được sử dụng khi chất lượng của lời giải hiện tại không được cải thiện sau một số lần lặp xác định.

## ĐIỀU KIỆN DỪNG

- Lời giải tốt nhất tìm được của bài toán không được cải thiện sau một số lần lặp định trước,
- Số lần lặp của thuật toán đạt đến một giá trị định trước,
- Thuật toán chạy hết một lượng thời gian định trước,...

## TIÊU CHÍ ĐÁNH GIÁ CHẤT LƯỢNG CÁC THUẬT TOÁN METAHEURISTIC

- Các thuật toán metaheuristic thường được đánh giá dựa trên chất lượng lời giải (độ tốt của lời giải) và thời gian tính tương ứng để có lời giải đó thông qua việc thực nghiệm trên các bộ dữ liệu chuẩn (benchmarks) của bài toán,
- Tiêu chí đánh giá chất lượng lời giải được ghi nhận qua giá trị tốt nhất (và có thể cả giá trị trung bình và giá trị độ lệch chuẩn) sau một số lần chạy đối với mỗi bộ dữ liệu,
- Việc so sánh thời gian tính các thuật toán trên các máy tính có cấu hình khác nhau chỉ mang tính tương đối; tuy nhiên các bảng so sánh thời gian này cho ta một cái nhìn tham khảo về thời gian tính của các thuật toán.

## TÍNH GIÁ TRỊ TRUNG BÌNH VÀ ĐỘ LỆCH CHUẨN (ôn tập kiến thức cơ bản)

Một biến số phản ảnh tình trạng của một bệnh trong hai nhóm bệnh nhân (nhóm A gồm 6 bệnh nhân, và nhóm B gồm 4 bệnh nhân) như sau:

Nhóm A: 6, 7, 8, 4, 5, 6

Nhóm B: 10, 2, 3, 9

Có thể dễ dàng thấy rằng số trung bình của nhóm A là 6, bằng với số trung bình của nhóm B.

chỉ số để phản ánh sự khác biệt giữa các bệnh nhân

$$\text{Nhóm A: } D_A = (6-6) + (7-6) + (8-6) + (4-6) + (5-6) + (6-6) = 0$$

$$\text{Nhóm B: } D_B = (10-6) + (2-6) + (3-6) + (9-6) = 0$$

Lấy bình phương của từng cá nhân và cộng số bình phương lại với nhau:

$$\text{Nhóm A: } D_A^2 = (6-6)^2 + (7-6)^2 + (8-6)^2 + (4-6)^2 + (5-6)^2 + (6-6)^2 = 10$$

$$\text{Nhóm B: } D_B^2 = (10-6)^2 + (2-6)^2 + (3-6)^2 + (9-6)^2 = 50$$

Chia  $D^2$  cho số cỡ mẫu-1

$$\text{Nhóm A: } s_A^2 = \frac{10}{6-1} = 2$$

$$\text{Nhóm B: } s_B^2 = \frac{50}{4-1} = 16.7$$

Chỉ số  $s^2$  ở đây chính là *phuong sai*.

*độ lệch chuẩn* (kí hiệu  $s$ ).

$$\text{Nhóm A: } s_A = \sqrt{2} = 1.41$$

$$\text{Nhóm B: } s_B = \sqrt{16.7} = 4.08$$

## VÍ DỤ 1

- Tìm giá trị trung bình và độ lệch chuẩn của dãy mẫu sau  
20; 20; 23; 26; 31; 20; 28; 17; 15; 19
- Cho dãy giá trị của  $n$  lần thực nghiệm một bộ dữ liệu của thuật toán A. Hãy tìm *giá trị trung bình* và *độ lệch chuẩn* bằng Excel
- Hãy cài đặt thuật toán tìm *giá trị trung bình* và *độ lệch chuẩn*.

## GENETIC ALGORITHM

## SƠ ĐỒ THUẬT TOÁN DI TRUYỀN

(chi tiết từng bước xem trong file bài giảng kèm theo)

### Genetic Algorithm

1. Khởi tạo quần thể ban đầu với các lời giải ngẫu nhiên
2. Đánh giá độ thích nghi của các thê của quần thê
3. **while** (chưa thỏa điều kiện dừng)
4. Chọn ngẫu nhiên một cặp cá thê cha-mẹ từ quần thê
5. Lai ghép (Crossover: Nếu xảy ra xác suất lai ghép  $p_{crossover}$  thì lai ghép 2 cá thê cha mẹ để tạo thành các cá thê con; ngược lại,các cá thê con giống cá thê cha-mẹ.
6. Đột biến (Mutation): Nếu xảy ra xác suất đột biến  $p_{mutation}$  thì đột biến bằng cách thay đổi nhỏ trong cá thê con.
7. Đánh giá lại độ thích nghi của các cá thê con.
8. Chọn lọc (Selection): Chọn các cá thê có độ thích nghi cao hơn cho thế hệ sau.
9. **end while**

## BEE ALGORITHM

- Thuật toán *bầy ong* sử dụng ba chiến lược tìm kiếm trọng tâm của các thuật toán metaheuristic nói chung là tìm kiếm lân cận, tìm kiếm ngẫu nhiên và tìm kiếm sâu hơn ở những vùng tiềm năng; trong đó việc tìm kiếm sâu hơn ở những vùng tiềm năng là một chiến lược đặc trưng của thuật toán bầy ong,
- Các thuật toán bầy ong được cho là có khả năng thoát khỏi bầy tối ưu cục bộ và do đó nó được hy vọng có thể tìm được lời giải tối ưu toàn cục,
- Thuật toán bầy ong được đánh giá là một trong những thuật toán metaheuristic thích hợp để giải các bài toán tối ưu tổ hợp khó.

## SƠ ĐỒ THUẬT TOÁN BEES

(chi tiết từng bước xem trong file bài giảng kèm theo)

### Basic\_Bees\_Algorithm()

1. Khởi tạo quần thể với các lời giải ngẫu nhiên
2. Đánh giá độ thích nghi của quần thể
3. Trong khi <điều kiện dừng chưa thoả>
4. Chọn vùng để tìm kiếm lân cận (neighbourhood search).
5. Tuyển thêm ong cho các vùng được chọn  
(vùng e sẽ có nhiều ong hơn) và đánh giá độ thích nghi.
6. Chọn ong có độ thích nghi cao nhất trong mỗi vùng.
7. Phân công những ong còn lại tìm kiếm ngẫu nhiên và đánh giá độ thích nghi.
8. Quay lại bước 3

Thuật toán Bees bao gồm một tập các tham số:

- + n: số ong do thám (tương ứng số vùng)
- + m: số vùng được chọn trong n vùng được thăm
- + e: số vùng tốt nhất trong m vùng
- + nep: số ong được cử đến e vùng tốt nhất
- + nsp: số ong được cử đến các vùng còn lại (m-e)
- + ngh: Kích thước của bụi hoa và lân cận của nó.

## LOCAL SEARCH ALGORITHM

- Từ một giải pháp hiện tại được chọn là  $s$ , thuật toán local search - LS sẽ tìm một giải pháp lân cận  $s'$  của  $s$ ;
- Nếu  $s'$  tốt hơn  $s$  thì  $s'$  sẽ trở thành giải pháp hiện tại,
- Quá trình này sẽ dừng khi thuật toán lặp đến một số lần định trước hoặc khi giải pháp tốt nhất không được cải thiện qua một số lần lặp xác định,
- LS là nền tảng của nhiều thuật toán metaheuristic khác.

## VARIABLE NEIGHBORHOOD SEARCH ALGORITHM (VNS)

- Ý tưởng của thuật toán lân cận biến đổi (Variable neighborhood search-VNS) là thực hiện lần lượt từng thuật toán lân cận, hết thuật toán lân cận này đến thuật toán lân cận khác,
- Trong quá trình thực hiện thuật toán VNS, ta luôn ghi nhận lời giải tốt nhất (kỉ lục),
- Khi thực hiện một thuật toán lân cận, nếu tìm được kỉ lục mới thì ta quay trở lại thực hiện thuật toán VNS từ đầu. Ngược lại, ta chuyển sang thuật toán lân cận tiếp theo,
- Quá trình được tiếp tục cho đến khi thực hiện hết tất cả các thuật toán lân cận mà lời giải tốt nhất không được cải thiện.

## BÀI TẬP 1

- Đề xuất một thuật toán metaheuristic giải bài toán cây steiner nhỏ nhất,
- Cài đặt thực nghiệm và đánh giá chất lượng thuật toán đề xuất.

## CÂY STEINER NHỎ NHẤT

### Định nghĩa 1. Cây Steiner

Cho  $G = (V(G), E(G))$  là một đồ thị vô hướng liên thông và có trọng số không âm trên cạnh; trong đó  $V(G)$  là tập gồm  $n$  đỉnh,  $E(G)$  là tập gồm  $m$  cạnh,  $w(e)$  là trọng số của cạnh  $e$ ,  $e \in E(G)$ . Cho  $L$  là tập con các đỉnh của  $V(G)$ , cây  $T$  đi qua tất cả các đỉnh trong  $L$  được gọi là cây Steiner của  $L$ .

Tập  $L$  được gọi là tập *terminal*, các đỉnh thuộc tập  $L$  được gọi là đỉnh *terminal*, đỉnh thuộc cây  $T$  mà không thuộc tập  $L$  được gọi là đỉnh Steiner. Khác với các bài toán cây khung thường gấp, cây Steiner chỉ cần đi qua tất cả các đỉnh thuộc tập *terminal*  $L$  và có thể thêm một số đỉnh khác nữa thuộc tập  $V(G)$ .

### Định nghĩa 2. Chi phí cây Steiner

Cho  $T = (V(T), E(T))$  là một cây Steiner của đồ thị  $G$ , chi phí của cây  $T$ , ký hiệu là  $C(T)$ , là tổng trọng số của các cạnh thuộc cây  $T$ , tức là  $C(T) = \sum_{e \in E(T)} w(e)$ .

### Định nghĩa 3. Cây Steiner nhỏ nhất

Cho đồ thị  $G$  được mô tả như trên, bài toán tìm cây Steiner có chi phí nhỏ nhất được gọi là bài toán cây Steiner nhỏ nhất (Steiner Minimal Trees problem – SMT); hoặc cũng có thể được gọi ngắn gọn hơn đó là bài toán cây Steiner Steiner Trees problem).

## BÀI TẬP 2

- Cho đồ thị vô hướng liên thông  $G$ . Hãy tìm một đồ thị con đầy đủ có nhiều đỉnh nhất.
- Đề xuất một thuật toán metaheuristic giải bài toán max clique lớn nhất.
- Cài đặt thực nghiệm và đánh giá chất lượng thuật toán đề xuất.

## 1.1. BÀI TOÁN TỐI UU

**Định nghĩa 6.1 (tối ưu):** Cho  $(S, f)$  là bài toán tối ưu, trong đó  $S \neq \emptyset$  biểu diễn cho không gian lời giải (còn được gọi là không gian tìm kiếm của bài toán) và  $f$  là hàm mục tiêu:

$$f: S \rightarrow R$$

Theo đó, việc giải bài toán tối ưu là quá trình tìm ra tập chứa giá trị các biến quyết định sao cho lời giải tìm được được biểu diễn bởi những giá trị  $x^* \in S$  thỏa mãn bất đẳng thức:

$$f(x^*) \leq f(x), \forall x \in S.$$

Hệ thức tương đương giữa bài toán tối ưu cận trên và bài toán tối ưu cận dưới như sau:

$$\max \{f(x)/x \in S\} \Leftrightarrow \min \{-f(x)/x \in S\}.$$

Dựa trên miền giá trị của  $S$  mà ta có tên gọi các bài toán tối ưu khác nhau: tối ưu nhị phân  $S \subseteq B$ , tối ưu rời rạc  $S \subseteq N$ , tối ưu liên tục  $S \subseteq R$ , tối ưu không đồng nhất (tối ưu hỗn tạp)  $S \subseteq B \cup N \cup R, \dots$

Hai lời giải  $x$  và  $y$  được gọi là lân cận của nhau nếu chúng cùng một lân cận trong không gian tìm kiếm.

**Định nghĩa 6.2 (lân cận):** Cho  $(S, f)$  là bài toán tối ưu, một cấu trúc lân cận trong  $S$  được phát biểu như sau:

$$N: S \rightarrow S,$$

Sao cho với mỗi lời giải  $x \in S$  thì tập  $S_x \subseteq S$  là đã được xác định. Theo đó, nếu  $x$  cùng thuộc lân cận với  $y$  thì  $y$  cũng thuộc lân cận với  $x$ , tức là  $y \in S_x \Leftrightarrow x \in S_y$ .

Nói chung, với bài toán tối ưu phức tạp, hàm mục tiêu thường biểu diễn một lời giải tối ưu là tối ưu chỉ trong một lân cận của nó, nhưng sẽ không là tối ưu nếu xét trên toàn bộ không gian tìm kiếm. Vì vậy, phương pháp tìm kiếm toàn cục có thể dễ dàng bị kẹt tại một giá trị tối ưu trong lân cận của nó; trường hợp này gọi là tối ưu cục bộ.

**Định nghĩa 6.3 (tối ưu cục bộ):** Cho  $(S, f)$  là bài toán tối ưu, và  $S_{x'} \subseteq S$  là lân cận của lời giải  $x' \in S_{x'}$ ,  $x'$  được xem là lời giải tối ưu cục bộ nếu thỏa mãn bất đẳng thức sau:

$$f(x') \leq f(x), \forall x \in S_{x'}$$

Trong các bài toán tối ưu thực tế, ta thường giải quyết các ràng buộc của bài toán. Theo đó, vùng lời giải khả thi của  $S$  sẽ bị giới hạn trong phạm vi thỏa mãn tất cả các ràng buộc đó.

**Định nghĩa 6.4 (tối ưu thỏa mãn ràng buộc):** Cho  $(S, f)$  là bài toán tối ưu, định nghĩa  $M = \{x \in S / g_k(x) \geq 0, \forall k \in [1, 2, \dots, q]\}$  là vùng chứa các lời giải khả thi trên hàm mục tiêu  $f: S \rightarrow R$  và các hàm  $g_k: S \rightarrow R$  gọi là các ràng buộc.

Mỗi  $g_k$  có tên gọi khác nhau dựa trên giá trị của chúng ứng với  $x \in S$ .

$$\text{Ràng buộc thỏa (satisfied):} \Leftrightarrow g_k(x) \geq 0,$$

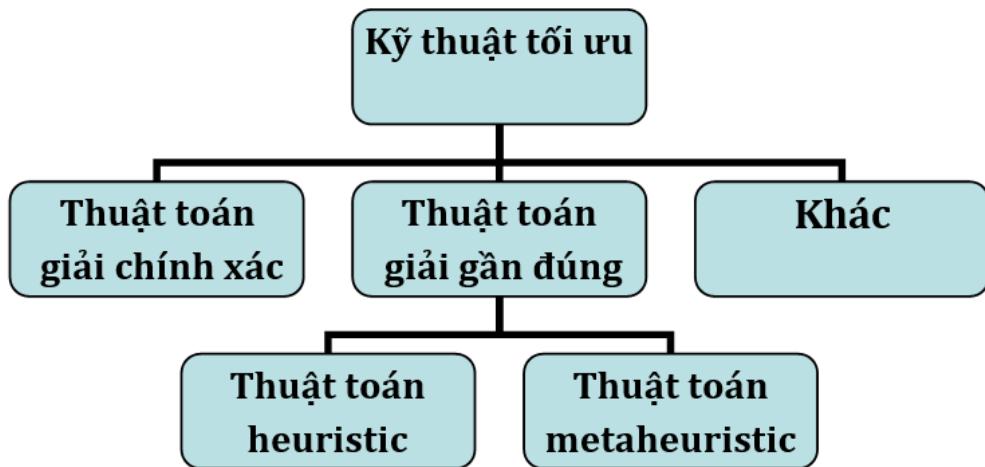
$$\text{Ràng buộc động (active):} \Leftrightarrow g_k(x) = 0,$$

$$\text{Ràng buộc tĩnh (inactive) :} \Leftrightarrow g_k(x) > 0,$$

$$\text{Ràng buộc vi phạm (violated):} \Leftrightarrow g_k(x) < 0,$$

Bài toán tối ưu toàn cục là thuật ngữ được sử dụng khi không có ràng buộc nào tồn tại; nghĩa là khi và chỉ khi  $M = S$ ; trong trường hợp ngược lại, bài toán được gọi là bài toán tối ưu thỏa mãn ràng buộc.

Hình 6.1 sau là một cách phân lớp các kỹ thuật tối ưu phổ biến hiện nay.



**Hình 6.1.** Phân lớp các kỹ thuật tối ưu

Một bài toán tối ưu thường có thể tiếp cận theo nhiều thuật toán khác nhau; chẳng hạn như: các thuật toán giải đúng (chính xác), thuật toán giải gần đúng cận tỉ lệ, thuật toán heuristic, thuật toán metaheuristic,...

Hướng tiếp cận giải đúng thường được phát triển dựa trên thuật toán quy hoạch động, thuật toán nhánh cạn hoặc thuật toán nhánh cạn. Hướng tiếp cận giải đúng chỉ có thể giải được các bài toán có kích thước nhỏ; do đó tính ứng dụng thực tiễn của nó không cao.

Hướng tiếp cận gần đúng cạn tỉ lệ có ưu điểm là có sự đảm bảo về mặt toán học theo nghĩa lời giải tìm được gần đúng theo một cạn tỉ lệ  $\alpha$  nào đó so với lời giải tối ưu. Tuy nhiên các cạn tỉ lệ tìm được của nhiều thuật toán giải các bài toán tối ưu trong thực tế thường là khá lỏng.

Hướng tiếp cận heuristic thường tìm được lời giải có chất lượng chấp nhận được ở một số loại dữ liệu nhất định đối với mỗi bài toán tối ưu. Thường thì hướng tiếp cận này có thời gian thực hiện nhanh hơn nhiều so với các thuật toán metaheuristic.

Hướng tiếp cận metaheuristic thường tìm được lời giải có chất lượng tốt hơn các thuật toán giải gần đúng còn lại. Tuy nhiên hiệu quả của các thuật toán metaheuristic chủ yếu được đánh giá thông qua thực nghiệm mà chưa chứng minh được về mặt toán học.

## **1.2. THUẬT TOÁN HEURISTIC VÀ THUẬT TOÁN METAHEURISTIC**

### **1.2.1. Heuristic**

Thuật toán heuristic chỉ những kinh nghiệm riêng biệt để tìm kiếm lời giải cho một bài toán tối ưu cụ thể. Thuật toán heuristic thường tìm được lời giải có thể chấp nhận được trong thời gian cho phép nhưng không chắc chắn là lời giải tốt nhất; thậm chí các thuật toán heuristic không chắc hiệu quả trên mọi loại dữ liệu đối với một bài toán cụ thể. Cấu trúc của một thuật toán heuristic thường là đơn giản.

### **1.2.2. Metaheuristic**

Thuật toán metaheuristic sử dụng nhiều heuristic kết hợp với các kỹ thuật phụ trợ nhằm khai phá không gian tìm kiếm; metaheuristic thuộc lớp các thuật toán tìm kiếm tối ưu.

Thuật ngữ “metaheuristic” thường đề cập đến các thuật toán dạng cá thể như local search, simulated annealing, tabu search,... hoặc dạng quần thể như genetic algorithm, particle swarm optimization, harmony search, ant, bee, bat,...

Khó để xác định thuật toán metaheuristic dạng cá thể hay dạng quần thể dạng nào là hiệu quả hơn; vì cả hai dạng này đều có hiệu năng mang lại như nhau dưới điều kiện thích hợp. Các thuật toán dạng quần thể có thể hiệu quả hơn trên các bài toán tối ưu đa mục tiêu vì nhiều phép

tìm kiếm được thực hiện cùng lúc; điều này có thể đúng theo quan sát thực nghiệm, nhưng chưa đủ cơ sở lý thuyết để khẳng định.

Hai yếu tố quan trọng nhất của một thuật toán metaheuristic là tính tăng cường (intensification) và tính đa dạng (diversification). Một thuật toán metaheuristic được đánh giá là hiệu quả nếu thuật toán đó có khả năng sinh ra một dãy các lời giải đa dạng nhằm khai phá hiệu quả không gian tìm kiếm. Song song đó nó cũng cần có khả năng tăng cường tìm kiếm các lân cận xung quanh các lời giải tốt một cách hiệu quả. Tính tăng cường và tính đa dạng là hai thành phần có phần đối lập nhau, việc tận dụng tốt tổ hợp giữa chúng là chìa khóa thành công của mọi thuật toán metaheuristic.

Chất lượng của các thuật toán metaheuristic phụ thuộc vào nhiều yếu tố như: cách thức tạo lời giải ban đầu, điều kiện dừng, cách thức mã hóa và giải mã lời giải,... nhưng quan trọng nhất là chiến lược tăng cường hóa và đa dạng hóa. Chiến lược tăng cường hóa được sử dụng xuyên suốt trong quá trình mà thuật toán diễn ra, còn chiến lược đa dạng hóa thường được sử dụng khi chất lượng của lời giải hiện tại không được cải thiện sau một số lần lặp xác định.

### 1.3. PHÂN TÍCH CÁC THÀNH PHẦN CỦA MỘT THUẬT TOÁN METAHEURISTIC

#### 1.3.1. Sơ đồ thuật toán metaheuristic

Sơ đồ chung của các thuật toán metaheuristic giải các bài toán tối ưu gồm các bước sau:

- Khởi tạo lời giải ban đầu (lời giải có thể là một cá thể/hoặc một quần thể);

- Định nghĩa các chiến lược tăng cường hóa;

- Định nghĩa các chiến lược đa dạng hóa;

- Tính độ thích nghi cho mỗi lời giải;

- **while** (*điều kiện dừng chưa thỏa*)

{

- Thực hiện các chiến lược tăng cường hóa;

- Thực hiện các chiến lược đa dạng hóa;

- Cập nhật lời giải tốt nhất cho đến thời điểm hiện tại;

}

- Trả về lời giải tốt nhất được;

### **1.3.2. Khởi tạo lời giải ban đầu**

Trong các thuật toán metaheuristic, lời giải ban đầu được khởi tạo ngẫu nhiên hoặc theo một hoặc một số heuristic riêng của bài toán; và thường thì lời giải ban đầu được khởi tạo ngẫu nhiên, nhất là khi số lần thực nghiệm trên mỗi bộ dữ liệu là đủ lớn để kết quả thực nghiệm là khách quan.

### **1.3.3. Chiến lược tăng cường hóa lời giải**

Tăng cường hóa nhằm mục đích khai thác sâu hơn các vùng không gian lời giải tiềm năng để hy vọng tìm được lời giải có chất lượng tốt hơn. Tăng cường hóa thường được thực hiện nhờ áp dụng các tiếp cận tham lam (tìm kiếm được hướng đến phương án lân cận tốt nhất). Nếu tính tăng cường quá cao, nó có thể nảy sinh vấn đề hội tụ sớm; các lời giải tối ưu cục bộ tìm được có độ lệch cao hoặc vô nghĩa. Ngược lại, nếu tính tăng cường quá thấp, quá trình hội tụ sẽ chậm.

Tính tăng cường hóa cho lời giải thường được thực hiện dựa vào các chiến lược tìm kiếm lân cận. Tùy theo đặc tính của bài toán mà có thể áp dụng các cách thức tìm kiếm lân cận khác nhau. Trong đó thường là:

- Thay một gen ngẫu nhiên trong lời giải bằng một gen ngẫu nhiên khác, hoặc
- Thay một gen ngẫu nhiên trong lời giải bằng một gen tốt nhất trong tập (hoặc tập con) các gen ứng viên (xem gen là thành phần cấu tạo nên lời giải).

### **1.3.4. Chiến lược đa dạng hóa lời giải**

Đa dạng hóa nhằm khai phá những vùng không gian tìm kiếm mới nhằm tránh việc tìm kiếm rơi vào bẫy tối ưu cục bộ. Đa dạng hóa thường được thực hiện qua việc sử dụng tính ngẫu nhiên và có thể kết hợp với một số heuristic riêng biệt cho từng bài toán. Nếu tính đa dạng quá cao thì sẽ có nhiều vùng không gian lời giải được khai phá ngẫu nhiên, theo đó sẽ làm chậm quá trình hội tụ của thuật toán. Nếu tính đa dạng quá thấp thì không gian lời giải được khai phá sẽ bị giới hạn, các lời giải tìm được có khuynh hướng hội tụ cục bộ hoặc vô nghĩa.

Một lời giải khi nó được cho thực hiện liên tiếp việc tìm kiếm lân cận theo một cách thức nào đó, thì sau một số lần lặp, chất lượng của lời giải đó sẽ không còn cải thiện được nữa (lời giải được gọi là đã khai thác cạn); nếu tiếp tục tìm kiếm lân cận theo hướng đó thì sẽ bế tắc. Có hai chiến lược đa dạng hóa lời giải thường được sử dụng:

- Thay thế ngẫu nhiên  $k$  gen của lời giải bằng  $k$  gen ngẫu nhiên khác mà không quan tâm đến chất lượng lời giải sau phép thay thế, hoặc

- Tạo lời giải hoàn toàn mới dựa vào xác suất của mỗi gen (phụ thuộc vào đặc tính riêng của bài toán) để thay thế cho lời giải đã được khai thác cạn.

### **1.3.5. Điều kiện dừng của các thuật toán metaheuristic**

Có nhiều điều kiện dừng đã được áp dụng trong các thuật toán metaheuristic, sau đây là ba điều kiện dừng thường được sử dụng:

- Lời giải tốt nhất tìm được của bài toán không được cải thiện sau *một số lần lặp định trước*, hoặc
- Số lần lặp của thuật toán đạt đến một *giá trị định trước*.
- Thuật toán chạy hết một lượng *thời gian định trước*.
- (các tham số này được đề xuất từ quá trình thực nghiệm tham số).

### **1.3.6. Tiêu chí đánh giá chất lượng các thuật toán metaheuristic**

Các thuật toán metaheuristic thường được đánh giá dựa trên chất lượng lời giải (độ tốt của lời giải) và thời gian tính tương ứng để có lời giải đó thông qua việc thực nghiệm trên các bộ dữ liệu chuẩn (benchmarks) của bài toán.

Tiêu chí đánh giá chất lượng lời giải được ghi nhận qua giá trị tốt nhất (và có thể cả giá trị trung bình và giá trị độ lệch chuẩn) sau một số lần chạy đối với mỗi bộ dữ liệu.

Việc so sánh thời gian tính các thuật toán trên các máy tính có cấu hình khác nhau chỉ mang tính tương đối; tuy nhiên các bảng so sánh thời gian này cho ta một cái nhìn tham khảo về thời gian tính của các thuật toán.

## 2. SƠ ĐỒ MỘT SỐ THUẬT TOÁN METAHEURISTIC

### 2.1. THUẬT TOÁN DI TRUYỀN

#### 2.1.1. Giới thiệu

Thuật toán di truyền (Genetic Algorithm - GA) thuộc lớp các phương pháp tìm kiếm thông dụng có sự kết hợp với các yếu tố của kỹ thuật tìm kiếm ngẫu nhiên (stochastic) có định hướng giúp tạo nên sự cân bằng tương đối giữa tìm kiếm và khai phá trong không gian lời giải. Khi các cá thể có độ thích nghi cao dần có mặt, phép lai tạo cơ hội khai phá những lời giải lân cận quanh chúng. Hay nói khác đi, cách mà phép lai thực hiện tìm kiếm (tìm hoặc khai phá) được xác định bởi yếu tố môi trường di truyền (hay còn gọi là tính đa dạng quần thể).

Nhìn chung, một thuật toán di truyền gồm 5 thành phần cơ bản, như Michalewicz đề cập:

- Biểu diễn di truyền cho các cá thể của bài toán.
- Cách để tạo quần thể (tập ban đầu gồm các lời giải tiềm năng).
- Một hàm lượng giá để đánh giá các lời giải theo độ thích nghi.
- Các toán tử di truyền để thay đổi các thành phần thuộc gen của cá thể con (phép lai, đột biến, chọn lọc, ...).
- Các giá trị tham số mà GA sử dụng (kích thước quần thể, xác suất thực hiện các toán tử di truyền lên quần thể, ...).

Thuật toán di truyền được mô phỏng dựa theo hoạt động di truyền của sinh vật trong tự nhiên; đó là lai ghép (crossover), đột biến (mutation) và chọn lọc (selection).

GA được hình thành dựa trên quan niệm cho rằng, quá trình tiến hoá tự nhiên là quá trình hoàn hảo nhất, hợp lý nhất và tự nó đã mang tính tối ưu. Quan niệm này có thể được xem như một tiên đề đúng mà không cần chứng minh, nó phù hợp với thực tế khách quan. Quá trình tiến hoá thể hiện tính tối ưu ở chỗ, thế hệ sau luôn tốt hơn (phát triển hơn, hoàn thiện hơn) thế hệ trước. Trong suốt quá trình tiến hoá tự nhiên, các thế hệ mới luôn được sinh ra để bổ sung và thay thế thế hệ cũ. Cá thể nào phát triển hơn, thích nghi hơn với môi trường sẽ tồn tại. Ngược lại, cá thể nào không thích ứng được với môi trường sẽ bị đào thải. Sự thay đổi môi trường là động lực thúc đẩy quá trình tiến hoá. Ngoài ra, tiến hoá cũng tác động trở lại góp phần làm thay đổi môi trường. Các cá thể mới sinh ra trong quá trình tiến hoá nhờ sự kết hợp (lai ghép) ở thế hệ cha-mẹ. Một cá thể mới có thể mang những tính trạng của cha-mẹ (di truyền), cũng có thể mang những tính trạng hoàn toàn mới (đột biến). Di truyền và đột biến là hai cơ chế có vai trò

quan trọng như nhau trong tiến trình tiến hoá, dù rằng đột biến xảy ra với xác suất nhỏ hơn nhiều so với hiện tượng di truyền. GA có bốn quá trình (phép toán) căn bản: lai ghép, đột biến, sinh sản và chọn lọc tự nhiên.

Khi sử dụng GA để giải quyết bài toán tối ưu thì có nhiều khía cạnh mâu chốt cần phải giải quyết: *Thứ nhất* là cách biểu diễn di truyền của một lời giải hay nói cách khác là tìm một cấu trúc dữ liệu phù hợp biểu diễn cho các cá thể (từ đây trở về sau khái niệm lời giải/cá thể/gen được hiểu là như nhau), *thứ hai* là cách thức khởi tạo quần thể ban đầu, *thứ ba* là xác định một hàm thích nghi đóng vai trò môi trường (tính thích nghi), *thứ tư* là định nghĩa các phép toán di truyền phù hợp, *thứ năm* là xác định rõ các tham số di truyền chẳng hạn như kích thước của quần thể, xác suất áp dụng cho từng phép toán di truyền, số thế hệ cần lặp lại,...

### 2.1.2. Khởi tạo quần thể

Nhìn chung, có 2 cách để tạo quần thể bao đầu: khởi tạo heuristic hoặc khởi tạo ngẫu nhiên trong khi còn thỏa các điều kiện biên hoặc các ràng buộc của bài toán. Mặc dù độ thích nghi trung bình của các cá thể khởi tạo theo phương pháp heuristic là tương đối cao nhằm giúp GA tìm kiếm lời giải nhanh hơn, nhưng trong phần lớn các bài toán có kích thước lớn, ví dụ, những bài toán thiết kế mạng, cách tiếp cận theo hướng heuristic chỉ có thể khai phá được một lượng nhỏ trong không gian lời giải và khó có thể tìm ra lời giải tối ưu toàn cục vì thiếu đi tính đa dạng của quần thể. Thường thì ta phải thiết kế riêng một thủ tục mã hóa dựa trên các nhiễm sắc thể để khởi tạo quần thể ban đầu từ đó.

### 2.1.3. Lượng giá độ thích nghi

Việc lượng giá độ thích nghi là để kiểm tra giá trị hàm mục tiêu của lời giải đối với các ràng buộc của bài toán. Nói chung, hàm mục tiêu cung cấp một cơ sở để lượng giá mỗi cá thể. Tuy nhiên, giá trị của nó biến đổi theo từng bài toán. Để duy trì tính đồng nhất với các dạng toán. Ta có thể sử dụng hàm thích nghi để bình thường hóa giá trị của hàm mục tiêu trong khoảng 0 đến 1. Giá trị được bình thường hóa của hàm mục tiêu chính là giá trị thích nghi của cá thể, và cơ chế chọn lọc được sử dụng để lượng giá các cá thể trong quần thể.

Trong quá trình tìm kiếm, quần thể trong GA trải qua giai đoạn tiến hóa theo độ thích nghi, và do đó hình thành nên quần thể mới. Ở mỗi bước lặp, các lời giải tốt được tái sử dụng và các lời giải xấu sẽ được đào thải, mục đích là để cá thể con gồm những lời giải tốt được tái sinh. Để phân biệt giữa các lời giải, một hàm lượng giá (hay còn gọi là hàm thích nghi) đóng một vai trò quan trọng trong môi trường tiến hóa, và các cơ chế điều hướng cũng thực sự cần thiết để

áp dụng vào hàm mục tiêu cho các hàm thích nghi. Khi lượng giá một hàm thích nghi cho nhiễm sắc thể, ta cần phải thiết kế riêng một thủ tục giải mã tương ứng trên nhiễm sắc thể đó.

Độ tốt của một cá thể trong quần thể chỉ là một trong số những cơ sở để xác định tính thích nghi của cá thể đó đối với môi trường. Rõ ràng một cá thể tốt nhất ở thế hệ hiện tại vẫn có khả năng bị kẹt lại trong các thế hệ sau và ngược lại, một cá thể chưa tốt ở thế hệ hiện tại vẫn có khả năng tiềm tàng dẫn đến cá thể tốt hơn ở thế hệ sau. Mặc dù vậy, thường thì cá thể tốt ở thế hệ hiện tại vẫn sẽ có xác suất dẫn đến cá thể tối ưu cao hơn những cá thể xấu hơn. Do đó, chúng ta vẫn lấy độ tốt của cá thể làm một yếu tố căn bản nhất để xác định tính thích nghi của cá thể đó. Thông thường, độ thích nghi của cá thể cũng chính là xác suất để cá thể đó được chọn lọc hoặc lai ghép hoặc đột biến khi tiến hành sinh ra thế hệ kế tiếp.

Sau đây là hai phương pháp thường được sử dụng để xác định tính thích nghi của cá thể.

### **Độ thích nghi tiêu chuẩn**

Hàm mục tiêu của bài toán thường là hàm dùng để đánh giá mức độ tốt của một lời giải (hàm mục tiêu được sử dụng để lượng giá độ tốt của các cá thể và có thể xem như là hàm lượng giá).

- Gọi *Popsize* là kích thước của quần thể. Mỗi cá thể *i* được ký hiệu là *t<sub>i</sub>*. Độ thích nghi *f(t<sub>i</sub>)* của mỗi cá thể *t<sub>i</sub>* (*i*=1..*Popsize*), ở đây *f(t<sub>i</sub>)* là hàm mục tiêu; với *i*=1..*n*-1 và *t<sub>i</sub>* là các cạnh của cây khung.

- Tính tổng giá trị thích nghi của cả quần thể:  $F = \sum_{i=1}^{Popsize} f(t_i)$
- Tính xác suất chọn cho từng cá thể *t<sub>i</sub>*:  $p_i = f(t_i)/F$
- Tính vị trí xác suất *q<sub>i</sub>* cho từng cá thể *t<sub>i</sub>*:  $q_i = \sum_{j=1}^i p_j$ , (*i* = 1..*Popsize*)

### **Độ thích nghi xếp hạng**

Cách tính độ thích nghi tiêu chuẩn như trên chỉ thực sự hiệu quả đối với những quần thể mà các cá thể có độ tốt tương đối đồng đều. Nếu vì một lý do nào đó mà có một cá thể có độ tốt quá cao hoặc quá thấp, tách biệt hẳn các cá thể còn lại thì các cá thể của thế hệ sau sẽ bị “hút” về phía cá thể đặc biệt đó tạo nên hiện tượng di truyền cục bộ. Phương pháp xác định độ thích nghi xếp hạng sẽ loại bỏ hiện tượng di truyền cục bộ này; phương pháp này không làm việc trên giá trị độ lớn của hàm mục tiêu *f* mà chỉ làm việc dựa trên thứ tự của các cá thể trên quần thể sau khi đã sắp xếp các cá thể theo giá trị của hàm mục tiêu *f*.

## 2.1.4. Các phép toán di truyền

### Lai ghép (Crossover)

Phép lai là quá trình hình thành nhiễm sắc thể mới trên cơ sở các nhiễm sắc thể cha-mẹ, bằng cách ghép một hay nhiều đoạn gen của hai (hay nhiều) nhiễm sắc thể cha-mẹ với nhau. Phép lai xảy ra với xác suất  $p_{crossover}$ . Có nhiều chiến lược lai ghép như one-point crossover (lai ghép một điểm), two-point crossover (lai ghép 2 điểm), uniform crossover (lai ghép đồng nhất). Sau đây là mô phỏng đơn giản về one-point crossover như sau:

- Chọn ngẫu nhiên hai cá thể bất kỳ trong quần thể; giả sử các nhiễm sắc thể của cha-mẹ đều có  $m$  gen.
- Tạo một số ngẫu nhiên trong khoảng từ 1 đến  $m-1$  (gọi là điểm lai). Điểm lai chia các chuỗi cha-mẹ dài  $m$  thành hai nhóm chuỗi con dài  $m_1$  và  $m_2$ . Hai chuỗi nhiễm sắc thể con mới sẽ là  $m_{11} + m_{22}$  và  $m_{21} + m_{12}$ .
- Đưa hai cá thể mới này vào quần thể để tham gia các quá trình tiến hóa tiếp theo.

### Bàn luận thêm về phép lai

Phép lai là toán tử di truyền chính. Nó được thực hiện cùng lúc trên 2 nhiễm sắc thể và tạo ra cá thể con bằng cách tổ hợp lại các đặc điểm của cả 2 cá thể lai (cá thể cha). Cách đơn giản để thực hiện phép lai là chọn ra ngẫu nhiên một điểm cắt (cut-point) và tạo ra cá thể con bằng cách kết hợp phân đoạn bên trái điểm cắt của cá thể cha thứ 1 và phân đoạn bên phải điểm cắt của cá thể lai còn lại. Phương pháp này hoạt động tốt trên cách biểu diễn chuỗi các bit. Hiệu năng của GA lẻ thuộc rất nhiều vào hiệu năng mà toán tử lai được sử dụng.

Xác suất lai (ký hiệu là  $pc$ ) được định nghĩa là tỉ lệ số lượng cá thể con được sinh ra ở mỗi thế hệ so với kích thước quần thể ( $popSize$ ). Xác suất lai kiểm soát số lượng nhiễm sắc thể được thực hiện phép lai  $pC \times popSize$ . Xác suất lai càng cao cho phép khai phá nhiều hơn trong không gian tìm kiếm, và giảm thiểu cơ hội hội tụ về lời giải tối ưu sai (false optimum); nhưng nếu xác suất lai quá cao, kết quả sẽ làm tốn nhiều thời gian tính toán trong việc khai phá những vùng không tìm nang trong không gian tìm kiếm.

Cho đến hiện tại, nhiều phép lai đã được đề xuất dành cho phương pháp mã hóa số thực, chúng được chia làm 4 lớp: truyền thống, số học, có hướng, và ngẫu nhiên.

Phép lai đơn giản: 1 điểm cắt, 2 điểm cắt và nhiều điểm cắt hoặc đồng nhất

Phép lai ngẫu nhiên: phép lai phẳng (flat crossover), phép lai cong (blend crossover)

Phép lai số học: được thực hiện thông qua khái niệm tổ hợp tuyến tính của các vector trong mảng lý thuyết tập lồi (convex set theory). Hoạt động trên cách biểu diễn di truyền dạng số thực có chấm động (floating point), các phép lai số học như phép lai lồi (convex), hữu hạn (affine), tuyến tính (linear), trung bình(average), trung gian(intermediate), dạng mở rộng của phép lai trung gian, thường là phô biến.

Phép lai có hướng (direction-based): được thực hiện bằng cách đưa độ lệch hướng vào các toán tử di truyền. Phép lai có hướng sử dụng giá trị của hàm mục tiêu để xác định hướng cho tìm kiếm di truyền.

Phép lai ngẫu nhiên: tạo ra cá thể con bằng cách tinh chỉnh các cá thể cha thông qua việc phân bố lại ngẫu nhiên cấu trúc của cá thể.

### **Đột biến (Mutation)**

Đột biến là hiện tượng cá thể con mang một số tính trạng không có trong gen di truyền của cha-mẹ. Phép đột biến xảy ra với xác suất  $p_{mutation}$ , nhỏ hơn nhiều so với xác suất lai  $p_{crossover}$ . Có nhiều chiến lược đột biến; sau đây là phép đột biến đảo:

- Chọn ngẫu nhiên một cá thể bất kỳ cha-mẹ trong quần thể.
- Tạo một số ngẫu nhiên  $k$  trong khoảng từ 1 đến  $m$ ,  $1 \leq k \leq m$ .
- Thay đổi gen thứ  $k$  và trả cá thể này về quần thể để tham gia quá trình tiến hoá tiếp theo.

### **Bàn luận thêm về phép đột biến**

Đột biến là toán tử cơ bản giúp sinh ra những thay đổi cùng lúc trên nhiều nhiễm sắc thể khác nhau. Một cách đơn giản để tạo đột biến là thay đổi một hoặc nhiều gen. Trong GA, đột biến đóng vai trò then chốt trong quá trình thay thế các gen đã mất trong quá trình chọn lọc tự nhiên của quần thể nhằm giúp chúng có cơ hội được thử trong nhiều biến cố khác nhau hoặc cung cấp các gen chưa từng hiện diện trong quần thể ban đầu.

Xác suất đột biến (ký hiệu là  $pM$ ) được định nghĩa là % trong tổng số gen của quần thể tham gia đột biến. Xác suất đột biến kiểm soát xác suất mà các gen mới được tạo để thử trong quần thể. Nếu xác suất đột biến quá thấp, nhiều gen vốn hữu dụng sẽ không có cơ hội được thử, trong khi nếu xác suất đột biến quá cao, thì sẽ có gây ra nhiều thay đổi hỗn loạn trong quần thể, cá thể con sẽ không còn giống với cha mẹ chúng, và thuật toán mất đi khả năng lưu vết trong khi thực hiện tìm kiếm.

Cho đến nay, nhiều toán tử đột biến được đề xuất cho phương pháp mã hóa số thực, có thể chia phép đột biến thành 4 nhóm như phép lai. Phép đột biến ngẫu nhiên gồm đột biến đồng nhất, đột biến cận và đột biến đơn giản, là thuộc về các phép đột biến truyền thống, các phép đột biến mà chỉ đơn giản là thay thế một gen ở một vùng định rõ bằng một số thực ngẫu nhiên. Đột biến động (dynamic mutation) hay (non-uniform mutation) dành riêng cho việc tinh chỉnh với độ chính xác cao, được xếp trong hàng toán tử đột biến số học. Đột biến có hướng (directional mutation) là dạng đột biến dựa trên hướng, dùng để gia tăng độ lệch của giá trị hàm thích nghi. Hướng có thể là ngẫu nhiên như vô hướng để tránh việc các nhiễm sắc thể bị kẹt vào 1 góc. Nếu nhiễm sắc thể gần cận, thì đột biến có hướng kèm theo vài điều kiện có thể điều hướng nó về biên gần nó, và hiện tượng tắc nghẽn (jamming) có thể xảy ra. Nhiều toán tử đột biến dành cho mã hóa nguyên (integer encoding) đã được đề xuất.

**Đột biến đảo ngược (Inversion mutation):** chọn 2 điểm ngẫu nhiên trong nhiễm sắc thể và đảo chuỗi con giữa chúng.

**Đột biến chèn (Insertion mutation):** chọn một gen ngẫu nhiên và chèn chúng vào một điểm bất kỳ

**Đột biến thay thế (Displacement mutation):** chọn ngẫu nhiên một chuỗi con các gen và chèn chúng vào một vị trí ngẫu nhiên. Do đó, đột biến chèn có thể xem là một trường hợp đặc biệt của đột biến thay thế.

**Đột biến hoán đổi nghịch đảo (Reciprocal exchange mutation):** chọn ngẫu nhiên 2 vị trí và sau đó hoán đổi các gen ở các vị trí.

### **Chọn lọc (Selection)**

Phép chọn lọc là quá trình loại bỏ các cá thể xấu trong quần thể để chỉ giữ lại trong quần thể các cá thể tốt. Có nhiều chiến lược chọn lọc như: Ranking selection, Tournament selection.

+ Ranking Selection:Sắp xếp các cá thể theo giá trị giảm dần độ thích nghi tức là tăng dần hàm mục tiêu của lời giải và chỉ chọn những cá thể nào có độ thích nghi cao hơn.

+ Tournament Selection:Chọn ngẫu nhiên từng cặp cá thể, so sánh độ thích nghi, và chỉ chọn cá thể có độ thích nghi cao hơn.

### **Bàn luận thêm về phép chọn lọc**

Chọn lọc giúp điều hướng trong GA. Nếu áp lực chọn lọc quá cao, thì sẽ làm chậm quá trình tìm kiếm. Thông thường, áp suất chọn lọc thấp sẽ giúp thuật toán mở rộng việc khai phá trong không gian lời giải, trong khi áp suất trong lọc cao sẽ làm thu hẹp đi không gian tìm kiếm.

Chọn lọc tự nhiên điều hướng thuật toán về những vùng tiềm năng trong không gian tìm kiếm. Trong 2 thập kỷ qua, nhiều phương pháp chọn lọc đã được đề xuất, thực nghiệm và so sánh. Có các phương pháp chọn lọc phổ biến như sau:

Chọn lọc vòng quay Roulette (Roulette wheel)

Chọn lọc ( $\lambda + \mu$ )

Chọn lọc phiên (tournament selection)

Chọn lọc cắt xén (truncation selection)

Chọn lọc ưu tú (elitist selection)

Xếp hạng và định cỡ (ranking and scanning)

Chọn lọc phân chia (sharing selection)

Chọn lọc vòng quay Roulette: được đề xuất bởi Holland, là phương pháp chọn lọc nổi tiếng, ý tưởng chính của phương pháp này là đưa ra một xác suất chọn lọc hay còn gọi là xác suất tồn tại, tỉ lệ với độ thích nghi cho mỗi nhiễm sắc thể. Sau đó, dùng mô hình vòng quay roulette để biểu diễn xác suất này. Quá trình chọn lọc dựa trên việc quay vòng quay với số lần quay đúng bằng kích thước quần thể, mỗi lần chọn ra một nhiễm sắc thể.

Không giống chọn lọc theo tỷ lệ, chọn lọc ( $\lambda + \mu$ ) là các thủ tục tất định sẽ chọn ra các nhiễm sắc thể tốt nhất từ các cá thể cha và con. Chú ý rằng cả 2 cách chọn lọc trên đều không chọn những nhiễm sắc thể giống nhau trong quần thể. Nhiều nghiên cứu chỉ ra chọn lọc ( $\lambda + \mu$ ) được dùng trong dạng toán tối ưu tổ hợp.

Chọn lọc phiên (tournament selection): cho chạy phiên trong số một ít các cá thể được chọn ngẫu nhiên trong quần thể và chọn ra cá thể trội (cá thể có độ thích nghi cao nhất). Áp lực chọn lọc có thể dễ dàng được tinh chỉnh bằng cách điều chỉnh kích thước phiên chạy. Nếu kích thước phiên chạy quá lớn, các cá thể yếu sẽ có ít cơ hội được chọn.

Chọn lọc cắt xén (truncation selection): cũng là một thủ tục tất định dùng để xếp hạng mọi cá thể theo độ thích nghi của chính và chọn ra các cá thể tốt nhất làm các cá thể cha. Chọn lọc ưu tú thông thường được dùng như là phương pháp bổ sung phụ trợ cho quá trình chọn lọc theo tỉ lệ.

Các cơ chế xếp hạng và lấy theo tỷ lệ được đề xuất để hỗ trợ cho những bài toán này, phương pháp lấy tỉ lệ ánh xạ những giá trị hàm mục tiêu thành những số thực dương, và xác suất tồn tại của mỗi nhiễm sắc thể được xác định dựa trên giá trị thực của chúng. Tỷ lệ thích nghi có 2

điều cần lưu ý: (1) là duy trì độ khác biệt hợp lý giữa tỉ lệ thích nghi tương đối của các nhiễm sắc thể, và (2) là để ngăn chặn trường hợp mà nhiều nhiễm sắc thể trội luôn thỏa điều làm cho việc chọn lọc bị giới hạn sớm đi nhưng sẽ dùng nó sau.

Phương pháp chọn lọc chia được dùng để duy trì tính đa dạng của quần thể đối với bài toán tối ưu hóa hàm đa hình (multi-model function optimization). Việc chia được dùng để duy trì tính đa dạng của quần thể. Chức năng chia là cách để xác định bước giảm (degradation) trong độ thích nghi của một cá thể do ảnh hưởng bởi cá thể lân cận ở một khoảng cách nào đó. Nhờ vào bước giảm, xác suất lai tạo của cá thể ưu tú được kiềm hãm đi để tạo cơ hội cho các cá thể khác được phép sinh sản cá thể con.

### 2.1.5. Sơ đồ thuật toán di truyền

#### Genetic Algorithm

1. Khởi tạo quần thể ban đầu với các lời giải ngẫu nhiên
2. Đánh giá độ thích nghi của các thê của quần thể
3. **while** (chưa thỏa điều kiện dừng)
4. Chọn ngẫu nhiên một cặp cá thể cha-mẹ từ quần thể
5. Lai ghép (Crossover: Nếu xảy ra xác suất lai ghép  $p_{crossover}$  thì lai ghép 2 cá thể cha mẹ để tạo thành các cá thể con; ngược lại, các cá thể con giống cá thể cha-mẹ.)
6. Đột biến (Mutation): Nếu xảy ra xác suất đột biến  $p_{mutation}$  thì đột biến bằng cách thay đổi nhở trong cá thể con.
7. Đánh giá lại độ thích nghi của các cá thể con.
8. Chọn lọc (Selection): Chọn các cá thể có độ thích nghi cao hơn cho thế hệ sau.
9. **end while**

Ban đầu, ta sẽ phát sinh một số lượng lớn, giới hạn các cá thể có gen ngẫu nhiên. Tập các cá thể này được gọi là quần thể ban đầu (initial population). Sau đó, dựa trên một hàm nào đó, ta sẽ xác định được một giá trị gọi là độ thích nghi (fitness). Vì phát sinh ngẫu nhiên nên độ "tốt" của lời giải hay tính thích nghi của các cá thể trong quần thể ban đầu là không xác định.

Để cải thiện độ thích nghi của quần thể, người ta tìm cách tạo ra quần thể mới. Có hai thao tác thực hiện trên thế hệ hiện tại để tạo ra một thế hệ khác với độ thích nghi tốt hơn.

Thao tác đầu tiên là sao chép nguyên mẫu một nhóm các cá thể tốt từ thế hệ trước rồi đưa vào thế hệ sau (selection). Thao tác này đảm bảo độ thích nghi của thế hệ sau luôn được giữ ở một mức độ hợp lý. Các cá thể được chọn thông thường là các cá thể có độ thích nghi cao nhất.

Thao tác thứ hai là tạo ra các cá thể mới bằng cách thực hiện các thao tác sinh sản trên một số cá thể được chọn từ thế hệ trước, thông thường cũng là những cá thể có độ thích nghi cao. Có hai loại thao tác sinh sản: Lai ghép (crossover) và đột biến (mutation). Trong thao tác lai ghép, từ gen của hai cá thể được chọn trong thế hệ trước sẽ được phối hợp với nhau (theo một số quy tắc nào đó) để tạo thành hai gen mới.

Thao tác chọn lọc và lai tạo giúp tạo ra thế hệ sau. Tuy nhiên, một số trường hợp thế hệ khởi tạo ban đầu có đặc tính phong phú và chưa phù hợp nên các cá thể không rải đều được hết không gian của bài toán. Từ đó, khó có thể tìm ra lời giải tối ưu cho bài toán. Thao tác đột biến sẽ giúp giải quyết được vấn đề này. Đó là sự biến đổi ngẫu nhiên một hoặc nhiều thành phần gen của một cá thể ở thế hệ trước tạo ra một cá thể hoàn toàn mới ở thế hệ sau. Nhưng thao tác này chỉ được phép xảy ra với tần suất rất thấp, vì thao tác này có thể gây xáo trộn và làm mất đi những cá thể đã chọn lọc và lai ghép có tính độ thích nghi cao, dẫn đến thuật toán không còn hiệu quả.

Thế hệ mới được tạo ra lại được xử lý như thế hệ trước. Tức là xác định độ thích nghi và tạo ra thế hệ mới cho đến khi có một cá thể đạt được giải pháp mong muốn hoặc đạt đến thời gian giới hạn.

## 2.2. THUẬT TOÁN BEES

Thuật toán Bees là thuật toán tối ưu dựa trên bầy ong và được ứng dụng trong nhiều bài toán thuộc các lĩnh vực khác nhau: Lập lịch làm việc cho một máy sản xuất, Giải các bài toán liên tục và tối ưu hoá trong kỹ thuật, Tối ưu hoá đa mục tiêu, Huấn luyện mạng nơron cho nhận dạng mẫu, Xếp thời khoá biểu,...

Thuật toán *bầy ong* sử dụng ba chiến lược tìm kiếm trọng tâm của các thuật toán metaheuristic nói chung là tìm kiếm lân cận, tìm kiếm ngẫu nhiên và tìm kiếm sâu hơn ở những vùng tiềm năng; trong đó việc tìm kiếm sâu hơn ở những vùng tiềm năng là một chiến lược đặc trưng của thuật toán bầy ong. Các thuật toán bầy ong được cho là có khả năng thoát khỏi bầy tối ưu cục bộ và do đó nó được hy vọng có thể tìm được lời giải tối ưu toàn cục. Thuật toán bầy ong được đánh giá là một trong những thuật toán metaheuristic thích hợp để giải các bài toán tối ưu tổ hợp khó.

### **2.2.1. Bees trong tự nhiên**

Một bầy ong mật có thể mở rộng khoảng cách tìm kiếm trên 10 km và theo nhiều hướng cùng lúc khác nhau để khai thác số lượng lớn nguồn thức ăn. Một bầy ong gọi là “phát đạt” khi đem được nhiều ong đến những vùng có nhiều thức ăn. Về nguyên tắc cơ bản, những bụi hoa (flower patches) mà có nhiều mật hoa hoặc phấn hoa mà có thể được lấy với ít công sức nhất thì sẽ có nhiều ong đến, ngược lại bụi hoa với ít mật hoa hoặc phấn hoa thì sẽ nhận được ít ong hơn.

Quá trình tìm kiếm thức ăn của bầy ong được bắt đầu bởi các ong do thám (scout bees) được cử đi để tìm kiếm các bụi hoa đầy hứa hẹn. Các ong do thám di chuyển ngẫu nhiên từ bụi hoa này sang bụi hoa khác. Trong suốt mùa thu hoạch, bầy ong tiếp tục khám phá và chọn ra một phần của quần thể làm ong do thám. Khi bầy ong trở về tổ thì những ong do thám nào mà tìm thấy bụi hoa với tỷ lệ nằm trên ngưỡng chất lượng (được đo lường bằng cách kết hợp các yếu tố chẳng hạn như lượng đường) về mật hoa hoặc phấn hoa và đi đến “sàn nhảy” (dance floor) để biểu diễn “điệu nhảy lúc lắc” (waggle dance). Điệu nhảy kỳ lạ này thì cần thiết đối với quần thể ong và chứa 3 phần thông tin quan trọng đến bụi hoa gồm: Hướng (direction) được tìm thấy, khoảng cách (distance) từ tổ đến bụi hoa và chất lượng (hoặc độ thích nghi). Thông tin này giúp cho quần thể gửi những ong đến các bụi hoa một cách chính xác mà không cần những chỉ dẫn hoặc bản đồ. Kiến thức về môi trường bên ngoài của mỗi cá thể được thu lượm từ “điệu nhảy lúc lắc”. Điệu nhảy này giúp cho bầy ong đánh giá sự khác biệt giữa các bụi hoa về chất lượng thức ăn cũng như số năng lượng cần thiết để thu hoạch bụi hoa đó. Sau “điệu nhảy lúc lắc” trên “sàn nhảy”, các ong tham gia nhảy (chẳng hạn ong do thám) trở về bụi hoa với các ong đi cùng (follower bees) chờ sẵn ở ngoài tổ. Các bụi hoa nhiều hứa hẹn sẽ có nhiều ong đi cùng hơn. Điều này cho phép bầy ong thu hoạch thức ăn một cách nhanh chóng và hiệu quả. Trong khi thu hoạch từ các bụi hoa, các ong sẽ giám sát lượng thức ăn để quyết định dựa trong “điệu nhảy lúc lắc” kế tiếp khi chúng về tổ. Nếu các bụi hoa mà vẫn còn nhiều thức ăn tốt, thì chúng sẽ quảng cáo trong “điệu nhảy lúc lắc” và sẽ có nhiều ong khác đến các bụi hoa này.

### **2.2.2. Sơ đồ thuật toán Bees**

Thuật toán Bees là thuật toán tối ưu hoá được mô phỏng dựa trên quá trình tìm kiếm thức ăn của ong mật trong tự nhiên để tìm ra lời giải tối ưu.

#### **Basic\_Bees\_Algorithm( )**

1. Khởi tạo quần thể với các lời giải ngẫu nhiên
2. Đánh giá độ thích nghi của quần thể
3. Trong khi <điều kiện dừng chưa thoả>

4. Chọn vùng để tìm kiếm lân cận (neighbourhood search).
5. Tuyển thêm ong cho các vùng được chọn  
(vùng e sẽ có nhiều ong hơn) và đánh giá độ thích nghi.
6. Chọn ong có độ thích nghi cao nhất trong mỗi vùng.
7. Phân công những ong còn lại tìm kiếm ngẫu nhiên và đánh giá độ thích nghi.
8. Quay lại bước 3

Thuật toán Bees bao gồm một tập các tham số:

- + n: số ong do thám (tương ứng số vùng)
- + m: số vùng được chọn trong n vùng được thăm
- + e: số vùng tốt nhất trong m vùng.
- + nep: số ong được cử đến e vùng tốt nhất.
- + nsp: số ong được cử đến các vùng còn lại (m-e)
- + ngh: Kích thước của bụi hoa và lân cận của nó.

Thuật toán bắt đầu với n ong do thám được khởi tạo ngẫu nhiên trong không gian tìm kiếm. Độ thích nghi của các vùng được thăm bởi các ong do thám được đánh giá trong Bước 2. Trong Bước 4, các ong mà có độ tương thích cao nhất thì được xem như “ong được chọn” và vùng được thăm bởi chúng thì được chọn cho tìm kiếm lân cận (neighbourhood search). Sau đó, trong Bước 5 và 6, thuật toán tiến hành tìm kiếm trong lân cận của các vùng được chọn, phân công nhiều ong hơn để tìm kiếm gần với e vùng tốt nhất. Các ong có thể được chọn trực tiếp dựa trên độ thích nghi kết hợp với các vùng mà chúng đã thăm. Ngoài ra, giá trị độ thích nghi được sử dụng để xác định xác suất các ong được chọn. Những tìm kiếm trong lân cận của các vùng e tốt nhất đại diện cho những lời giải hứa hẹn hơn được thực hiện chi tiết hơn bằng cách tuyển thêm ong theo chúng hơn những ong được chọn khác. Kết hợp với việc do thám, việc tuyển thêm ong khác nhau này là thao tác chính của thuật toán Bees. Tuy nhiên, ở bước 6, trong mỗi khu vực (patch) chỉ có một ong với độ thích nghi cao nhất sẽ được chọn để tham gia vào quần thể ong tiếp theo. Trong tự nhiên, thì không có giới hạn. Giới hạn này được giới thiệu ở đây để giảm số điểm cần khám phá. Trong bước 7, các ong còn lại trong quần thể được phân công một cách ngẫu nhiên do thám quanh không gian tìm kiếm để cho ra những lời giải tiềm năng mới. Các bước này được lặp đi lặp lại cho đến khi thoả điều kiện dừng. Tại cuối mỗi bước lặp, quần thể ong sẽ có gồm 2 phần cho quần thể mới đại diện cho mỗi khu vực được chọn và những ong do thám khác được phân công để tìm kiếm ngẫu nhiên.

### **2.3. THUẬT TOÁN TÌM KIẾM CỤC BỘ**

Thuật toán tìm kiếm cục bộ (Local Search -LS) có ý tưởng chính như sau:

Từ một giải pháp hiện tại được chọn là  $s$ , thuật toán local search - LS sẽ tìm một giải pháp lân cận  $s'$  của  $s$ ; nếu  $s'$  tốt hơn  $s$  thì  $s'$  sẽ trở thành giải pháp hiện tại. Quá trình này sẽ dừng khi thuật toán lặp đến một số lần định trước hoặc khi giải pháp tốt nhất không được cải thiện qua một số lần lặp xác định. LS là nền tảng của nhiều thuật toán metaheuristic khác.

### **2.4. THUẬT TOÁN LÂN CẬN BIẾN ĐỔI**

Ý tưởng của thuật toán lân cận biến đổi (Variable neighborhood search-VNS) là thực hiện lần lượt từng thuật toán lân cận, hết thuật toán lân cận này đến thuật toán lân cận khác. Trong quá trình thực hiện thuật toán VNS, ta luôn ghi nhận lời giải tốt nhất (kỉ lục). Khi thực hiện một thuật toán lân cận, nếu tìm được kỉ lục mới thì ta quay trở lại thực hiện thuật toán VNS từ đầu. Ngược lại, ta chuyển sang thuật toán lân cận tiếp theo. Quá trình được tiếp tục cho đến khi thực hiện hết tất cả các thuật toán lân cận mà lời giải tốt nhất không được cải thiện.

Đề xuất các thuật toán heuristic/meta heuristic giải các bài toán:

**BT7.1.** Bài toán cây steiner nhỏ nhất

**BT7.2.** Bài toán max clique

**ĐỀ SỐ 1****Câu 1 (2 điểm)**

Một dịch vụ in án luận văn tốt nghiệp, có 3 nhân viên đánh máy và một quản lý. Dịch vụ nhận được yêu cầu đánh máy luận văn của sinh viên làm luận văn tốt nghiệp như sau:

Luận văn	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12
Số trang	220	160	80	120	80	140	40	120	120	180	60	80

Giả sử trong một giờ thì một nhân viên đánh máy được 10 trang (giả thiết các nhân viên có công suất đánh máy bằng nhau).

a. Phân chia các luận văn cho 03 nhân viên đánh máy sao cho thời gian hoàn thành việc đánh máy các luận văn trên là sớm nhất.

b. Trong trường hợp người quản lý cũng tham gia đánh máy, nhưng công suất của người quản lý chỉ bằng  $\frac{2}{3}$  công suất của nhân viên. Tìm cách chia các luận văn cho 3 nhân viên và người quản lý, sao cho thời gian hoàn thành việc đánh máy các luận văn là sớm nhất.

**Câu 2 (2 điểm)**

Tại một cửa hàng sách, mới nhập về 12 quyển sách thuộc các loại sau:

Truyện cười: A, C, D, G. Âm nhạc: B, H, K.

Lịch sử: E, J, L. Khoa học: F, I.

Hãy sử dụng thuật toán tô màu tham lam để sắp những quyển sách này vào kệ sao cho số kệ sử dụng là ít nhất mà tuân theo các yêu cầu sau:

- Các quyển sách cùng loại không được để chung một kệ.

- Quyển A không được để chung với sách khoa học.

- Quyển L không được để chung với sách âm nhạc.

**Câu 3 (2 điểm)**

Xét các câu đúng:

α: “Nếu sinh viên giỏi Toán rời rạc thì giỏi Phân tích thiết kế thuật toán”.

β: “Nếu sinh viên giỏi Phân tích thiết kế thuật toán thì giỏi Lập trình”.

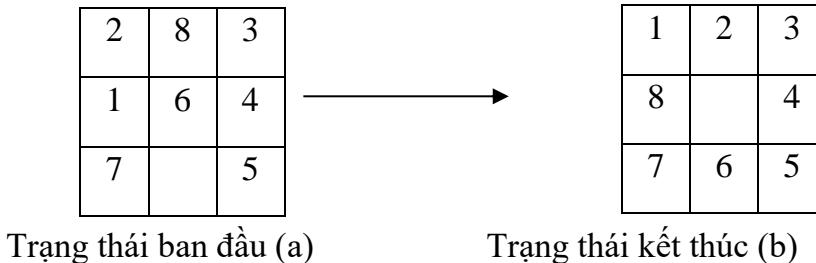
δ: “Nếu sinh viên không giỏi Lập trình thì không giỏi Toán rời rạc” .

Bài toán đặt ra là: Từ (hội) các câu *đúng* α, β chứng minh δ.

- a. Xây dựng bài toán trên trên bằng các biểu thức logic mệnh đề.  
 b. Hãy chứng minh bài toán bằng phương pháp Vương Hạo.

#### Câu 4 (2 điểm)

Dùng thuật toán A<sup>KT</sup> giải bài toán TACI với n = 3 như sau:



Với độ ước lượng  $H = \sum_{i=1}^{n^2-1} \delta(a_i, b_i)$  Trong đó  $\delta(a_i, b_i) = 0$  nếu  $a_i = b_i$  và  $\delta(a_i, b_i) = 1$  nếu  $a_i$  khác  $b_i$

#### Câu 5 (2 điểm)

Cho cơ sở dữ liệu sau:

Mẫu	Quang cảnh	Nhiệt độ	Độ ẩm	Gió	Quyết định
1	Mưa	Âm	Cao	Nhẹ	Không
2	Nhiều mây	Nóng	Cao	Nhẹ	Đi
3	Nhiều mây	Âm	Cao	Mạnh	Đi
4	Nắng	Âm	Thấp	Nhẹ	Đi
5	Nắng	Lạnh	Thấp	Mạnh	Không
6	Mưa	Nóng	Cao	Mạnh	Không
7	Nắng	Âm	Cao	Nhẹ	Đi
8	Nắng	Lạnh	Thấp	Nhẹ	Đi
9	Nhiều mây	Lạnh	Thấp	Mạnh	Đi
10	Mưa	Lạnh	Thấp	Nhẹ	Đi
11	Mưa	Nóng	Cao	Nhẹ	Không
12	Mưa	Âm	Thấp	Mạnh	Đi

Từ CSDL trên hãy rút ra bộ luật cho sự quyết định đi/không đi theo thuật toán

Quinlan

Hết

## ĐỀ SỐ 2

### Câu 1 (2 điểm)

Xét các câu đúng:

- α: “Nếu sinh viên giỏi Toán rời rạc thì giỏi Phân tích thiết kế thuật toán”.
- β: “Nếu sinh viên giỏi Phân tích thiết kế thuật toán thì giỏi Lập trình”.
- δ: “Nếu sinh viên không giỏi Lập trình thì không giỏi Toán rời rạc” .

Bài toán đặt ra là: Từ (hội) các câu *đúng* α, β chứng minh δ.

- a. Xây dựng bài toán trên bằng các biểu thức logic mệnh đề.
- b. Hãy chứng minh bài toán bằng phương pháp Robinson

### Câu 2 (2 điểm)

*Bài toán đong nước:* Cho hai bình có dung tích a và b lít, làm thế nào để đong được chính xác c lít ?

- a. Hãy nêu các luật để giải bài toán đong nước.
- b. Hãy nêu một điều kiện của a,b,c để bài toán đong nước là có lời giải ?
- c. Áp dụng giải bài toán đong nước khi bình 1 đựng 6 lít, bình 2 đựng 8 lít và số nước cần đong là 4 lít. Yêu cầu mô tả số nước ở mỗi bình sau khi áp dụng mỗi luật và hãy ghi rõ đó là luật thứ mấy trong các luật đã nêu ?

### Câu 3 (2 điểm)

Có 6 đội bóng đá GIALAI, ĐÀ NẴNG, LONG AN, BÌNH DƯƠNG, HẢI PHÒNG, ĐỒNG THÁP thi đấu vòng tròn tranh giải vô địch. Trong một tuần, mỗi đội chỉ thi đấu với một đội khác đúng một trận. Biết rằng đội GIA LAI đã thi đấu với đội ĐÀ NẴNG và BÌNH DƯƠNG; đội ĐÀ NẴNG đã thi đấu với đội LONG AN và ĐỒNG THÁP; đội HẢI PHÒNG đã thi đấu với đội LONG AN và ĐỒNG THÁP.

Hãy sử dụng thuật toán tô màu để lập một lịch thi đấu sao cho các trận còn lại sẽ được thực hiện trong một số tuần là ít nhất có thể.

### Câu 4 (2 điểm)

Cho đồ thị có ma trận chi phí như sau:

$\infty$	19	27	25	1	20
7	$\infty$	11	2	28	14
8	4	$\infty$	23	9	4
7	4	19	$\infty$	26	10
9	5	16	10	$\infty$	40
31	10	11	2	82	$\infty$

Sử dụng thuật toán GTS2 để tìm hành trình tốt nhất với  $p=4$  ( $v_1=1$ ;  $v_2=3$ ;  $v_3=4$ ;  $v_4=5$ )

### Câu 5 (2 điểm)

Khi nghiên cứu về sở thích của trẻ con khi chọn đồ chơi, người ta có được các dữ liệu cho bởi bảng sau:

Stt	Kích cỡ	Màu sắc	Hình dáng	Chọn
1	Trung bình	Xanh	Khối gạch	Có
2	Nhỏ	Đỏ	Hình nêm	Không
3	Nhỏ	Đỏ	Hình cầu	Có
4	Lớn	Đỏ	Hình nêm	Không
5	Lớn	Xanh	Hình trụ	Có
6	Lớn	Đỏ	Hình trụ	Không
7	Lớn	Xanh	Hình cầu	Có

- a. Tìm cây quyết định bằng thuật toán Quinlan.
- b. Dựa vào cây quyết định đã tìm ra, hãy đưa ra các luật của việc chọn đồ chơi của trẻ con.
- c. Hãy cho biết kết quả của các mẫu dưới đây:

Stt	Kích cỡ	Màu sắc	Hình dáng	Chọn
8	Trung bình	Đỏ	Khối gạch	?
9	Lớn	Đỏ	Hình trụ	?

Hết

## ĐỀ SỐ 3

### Câu 1 (2 điểm)

Giả sử có m máy như nhau được ký hiệu từ  $P_1, \dots, P_m$ . Có n công việc  $J_1, \dots, J_n$  cần được thực hiện. Các công việc có thể được thực hiện đồng thời và bất kỳ công việc nào cũng có thể chạy trên một máy nào đó. Mỗi lần máy được cho thực hiện một công việc nó sẽ làm cho tới khi hoàn chỉnh. Công việc  $J_i$  có thời gian thực hiện là  $T_i$ . Mục đích của chúng ta là tổ chức cách phân công các công việc được hoàn thành với thời gian sớm nhất có thể.

- Hãy nêu thuật toán giải quyết bài toán trên.
- Giả sử có 3 máy  $P_1, P_2, P_3$  và 6 công việc  $J_1, J_2, J_3, J_4, J_5, J_6$  với  $T_i = (7, 10, 13, 6, 9, 6)$ . Hãy tìm một phương án tốt để sắp các công việc vào các máy.

### Câu 2 (2 điểm)

Có một hội thảo khoa học được tổ chức với 8 chủ đề khác nhau ký hiệu là: 1, 2, 3, 4, 5, 6, 7, 8. Mỗi chủ đề được diễn ra trong một buổi, trong đó các chủ đề sau không được diễn ra đồng thời trong cùng một buổi:

(1,2,3); (1,4,5); (3,4,6); (2,5,6); (1,7,8); (4,7); (5,8).

Hãy lập lịch tổ chức hội thảo sao cho số buổi hội thảo diễn ra là ít nhất có thể.

### Câu 3 (2 điểm)

a. Cho cơ sở tri thức KB1 =  $\{(q \wedge p) \rightarrow s, (q \wedge s) \rightarrow r, q, p\}$

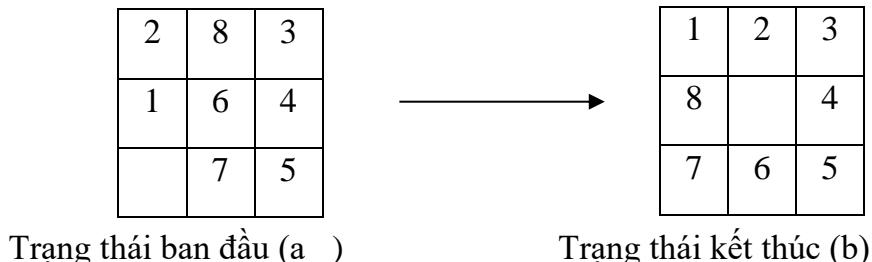
Hãy sử dụng thuật toán Vương Hạo để kiểm tra xem  $s$  có được suy ra từ cơ sở tri thức trên hay không?

b. Cho cơ sở tri thức KB2 =  $\{(q \rightarrow s) \wedge (p \rightarrow s)\}$

Hãy sử dụng thuật toán Robinson để kiểm tra xem  $(q \vee p) \rightarrow s$  có được suy ra từ cơ sở tri thức trên hay không?

#### Câu 4 (2 điểm)

Dùng thuật toán A<sup>KT</sup> giải bài toán TACI với n = 3 như sau:



Với hàm heuristic  $H = \sum_{i=1}^{n^2-1} \delta(a_i, b_i)$  Trong đó  $\delta(a_i, b_i)$  là số bước dịch chuyển (theo chiều ngang và chiều dọc) để đẩy ô  $a_i$  về đúng vị trí ô  $b_i$

#### Câu 5 (2 điểm)

Cho cơ sở dữ liệu sau:

#	Trời	Áp suất	Gió	Quyết định
1	Trong	Cao	Bắc	Không mưa
2	Mây	Cao	Nam	Mưa
3	Mây	Trung Bình	Bắc	Mưa
4	Trong	Thấp	Bắc	Không mưa
5	Mây	Thấp	Bắc	Mưa
6	Mây	Cao	Bắc	Mưa
7	Mây	Thấp	Nam	Không mưa
8	Trong	Cao	Nam	Không mưa
9	Trong	Trung Bình	Bắc	?
10	Mây	Trung Bình	Nam	?

- a. Từ mẫu 1 đến mẫu 8 hãy rút ra bộ luật cho sự quyết định.
- b. Cho biết kết quả của mẫu #9 và #10

Hết

## ĐỀ SỐ 4

### Câu 1 (2 điểm)

a. Ứng dụng thuật toán Robinson để chứng minh tính đúng đắn của phát biểu sau:

$$(p \rightarrow q, r \rightarrow s, (q \vee s) \rightarrow t, \neg t) \rightarrow (\neg r \wedge \neg p)$$

b. Cho cơ sở tri thức KB = { $(a \wedge b) \rightarrow c, (b \wedge c) \rightarrow d, \neg d$  }

Hãy sử dụng thuật toán Vương Hạo để kiểm tra xem  $a \rightarrow b$  có được suy ra từ cơ sở tri thức trên hay không ?

### Câu 2 (2 điểm)

Có một hội thảo khoa học được tổ chức với 9 chủ đề khác nhau ký hiệu là: A,B,C,D,E,F,G,H,I. Mỗi chủ đề được diễn ra trong một buổi, trong đó có các chủ đề sau không được diễn ra đồng thời trong cùng một buổi: {B,H,I}, {D,F,I}, {D,H,I}, {F,G,H}, {A,E}, {B,C}, {C,D}, {E,D}, {A,B,D}, {A,H,I}.

Hãy sử dụng thuật toán tô màu tham lam đã nêu trên để bố trí các chủ đề trên vào các buổi sao cho số buổi hội thảo diễn ra là ít nhất có thể.

### Câu 3 (2 điểm)

Hãy sử dụng thuật toán A<sup>KT</sup> – tìm kiếm với tri thức bổ sung (Algorithm knowledgeable For Tree) để giải bài toán tháp Hà Nội trong trường hợp n=3 biết:



### Câu 4 (2 điểm)

Có 12 chi tiết máy  $D_1, D_2, \dots, D_{12}$  phải được lần lượt gia công trên 2 máy  $M_1, M_2$ . Thời gian gia công chi tiết  $D_i$  trên máy  $M_1$  là  $\{4, 16, 7, 13, 9, 2, 14, 15, 7, 11, 3, 8\}$ , trên máy  $M_2$  là  $\{5, 7, 13, 9, 12, 6, 9, 12, 4, 17, 3, 14\}$ . Hãy tìm lịch (trình tự gia công) các chi tiết trên hai máy sao cho việc hoàn thành gia công tất cả các chi tiết là sớm nhất có thể được. Giả thiết rằng, trình tự gia công các chi tiết trên hai máy là như nhau và các chi tiết được làm trên máy  $M_1$  rồi đến máy  $M_2$ .

### **Thuật toán Johnson :**

'Mỗi một chi tiết  $D_1, D_2, \dots, D_n$  cần phải được lần lượt gia công trên 2 máy A,B. Thời gian gia công chi tiết  $D_i$  trên máy A là  $a_i$  trên máy B là  $b_i$  ( $i=1,2..n$ ). Hãy tìm lịch (trình tự gia công) các chi tiết trên hai máy sao cho việc hoàn thành gia công tất cả các chi tiết là sớm nhất'.

Chia các chi tiết thành 2 nhóm: nhóm N1, gồm các chi tiết  $D_i$  thoả mãn  $a_i < b_i$ , tức là  $\min(a_i, b_i) = a_i$  và nhóm N2 gồm các chi tiết  $D_i$  thoả mãn  $a_i > b_i$  tức là  $\min(a_i, b_i) = b_i$ . Các chi tiết  $D_i$  thoả mãn  $a_i = b_i$  xếp vào nhóm nào cũng được. Sắp xếp các chi tiết trong N1 theo chiều tăng của các  $a_i$  và sắp xếp các chi tiết trong N2 theo chiều giảm của các  $b_i$ . Nối N2 vào đuôi N1, dãy thu được (đọc từ trái sang phải) sẽ là lịch gia công.

### **Câu 5 (2 điểm)**

Cho cơ sở dữ liệu sau:

Mẫu	Thời tiết	Lá cây	Nhiệt độ	Quyết định (mùa)
1	Mưa	Rụng	Thấp	Đông
2	Nắng	Xanh	Trung bình	Xuân
3	Nắng	Vàng	Trung bình	Thu
4	Nắng	Xanh	Cao	Hè
5	Nắng	Rụng	Thấp	Đông
6	Tuyết	Vàng	Thấp	Đông
7	Mưa	Rụng	Trung bình	Thu
8	Mưa	Xanh	Cao	Hè
9	Tuyết	Xanh	Thấp	Đông
10	Tuyết	Rụng	Thấp	Đông
11	Mưa	Vàng	Trung bình	Thu
12	Mưa	Xanh	Trung bình	Xuân
13	Mưa	Vàng	Thấp	?
14	Tuyết	Rụng	Trung bình	?

a.Từ mẫu 1 đến mẫu 12 hãy rút ra bộ luật cho sự quyết định Mùa.

b. Áp dụng cho biết kết quả các mẫu 13 và 14.

Hết

## ĐỀ SỐ 5

### CÂU 1: (1.5 điểm)

Một dịch vụ in án luận văn tốt nghiệp, có 3 nhân viên đánh máy và một quản lý. Dịch vụ nhận được yêu cầu đánh máy luận văn của sinh viên làm luận văn tốt nghiệp như sau:

Luận văn	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12
Số trang	200	140	70	100	60	120	50	80	100	150	40	60

Giả sử trong một giờ thì một nhân viên đánh máy được 10 trang

a. Phân chia các luận văn cho 03 nhân viên đánh máy sao cho thời gian hoàn thành việc đánh máy luận văn là sớm nhất.

b. Trong trường hợp người quản lý cũng tham gia đánh máy, nhưng công suất của người quản lý chỉ bằng  $\frac{1}{2}$  công suất của một nhân viên. Tìm cách chia các luận văn cho 3 nhân viên và người quản lý, sao cho thời gian hoàn thành việc đánh máy luận văn là sớm nhất.

### CÂU 2: (2 điểm)

Có một hội thảo khoa học được tổ chức với 10 chủ đề khác nhau ký hiệu là: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10. Mỗi chủ đề được diễn ra trong một buổi, trong đó các chủ đề sau không được diễn ra đồng thời trong cùng một buổi: (1,2,3); (1,4,5); (3,4,6); (2,5,6); (1,7,8); (7,8,10); (4,7,9); (4,7); (5,8).

Hãy lập lịch tổ chức hội thảo sao cho số buổi hội thảo diễn ra là ít nhất có thể.

### CÂU 3: (2 điểm)

Xét các câu:

$\alpha$ : “Nếu trời mưa thì Lan mang theo dù”

$\beta$ : “Nếu Lan mang theo dù thì Lan không bị ướt”

$\delta$ : “Nếu trời không mưa thì Lan không bị ướt”

$\lambda$ : “Lan không bị ướt”

Bài toán đặt ra là: Từ các câu  $\alpha, \beta, \delta$  chứng minh  $\lambda$ .

a. Xây dựng bài toán trên bằng các biểu thức logic mệnh đề.

b. Hãy chứng minh bài toán bằng phương pháp Vương Hạo và phương pháp Robinson.

### CÂU 4: (1.5 điểm)

Có một người nông dân nuôi bò sữa. Một hôm, có một người khách muốn mua đúng 8 lít sữa (người khách có một can rỗng dung tích 8 lít), nhưng người nông dân chỉ có một can dung tích 16 lít đựng đầy sữa, một can dung tích 7 lít (rỗng) và một can dung tích 3 lít (rỗng). Chỉ được sử dụng 3 can trên của mình, hỏi làm thế nào để người nông dân lấy ra được 8 lít sữa cho khách mua hàng ?

### CÂU 5: (3 điểm)

Cho cơ sở dữ liệu gồm các mẫu sau:

Mẫu	Thời tiết	Lá cây	Nhiệt độ	Quyết định (mùa)
1	Mưa	Rung	Thấp	Đông
2	Nắng	Xanh	Trung bình	Xuân
3	Nắng	Vàng	Trung bình	Thu
4	Nắng	Xanh	Cao	Hè
5	Nắng	Rụng	Thấp	Đông
6	Tuyết	Vàng	Thấp	Đông
7	Mưa	Rụng	Trung bình	Thu
8	Mưa	Xanh	Cao	Hè
9	Tuyết	Xanh	Thấp	Đông
10	Tuyết	Rung	Thấp	Đông
11	Mưa	Vàng	Trung bình	Thu
12	Mưa	Xanh	Trung bình	Xuân

$x$	Mưa	Vàng	Cao	?
$y$	Tuyết	Rụng	Trung bình	?
$z$	Tuyết	Vàng	Trung bình	?

- a. Từ mẫu 1 đến mẫu 12 hãy rút ra bộ luật cho sự quyết định Mùa theo thuật toán Quinlan.  
 b. Áp dụng cho biết kết quả các mẫu  $x, y, z$ .

Hết

## ĐỀ SỐ 6

### CÂU 1: (4 điểm)

- Hãy trình bày khái niệm thuật toán heuristic. So sánh thuật toán heuristic với các thuật toán giải đúng trong việc giải bài toán tối ưu.
- Có 3 đồng sỏi lần lượt có  $n_1, n_2$  và  $n_3$  viên, có hai người tham gia trò chơi này; luôn phiền đến lượt mình được bốc từ một đồng sỏi bất kỳ một số ( $>0$ ) viên sỏi bất kỳ. Người chơi nào đến lượt mình mà không còn sỏi để bốc là người thua. Hãy trình bày thuật toán giải quyết trò chơi trên.
- Cho một cơ sở dữ liệu gồm  $n$  thuộc tính  $A_1, A_2, \dots, A_n$  và một thuộc tính quyết định  $D$ ; cơ sở dữ liệu này có  $m$  mẫu tin. Hãy trình bày thuật toán Quinlan để tìm bộ luật cho quyết định  $D$  dựa vào cơ sở dữ liệu trên.

### CÂU 2: (3 điểm)

Có  $n$  công việc được đánh số từ 1 đến  $n$  và một máy để thực hiện. Biết  $t_i$  là lượng thời gian cần thiết để hoàn thành công việc thứ  $i$  và  $f_i$  là thời điểm cần hoàn thành công việc thứ  $i$ . Giả thiết máy bắt đầu hoạt động từ thời điểm 0; mỗi công việc cần được thực hiện liên tục từ lúc bắt đầu cho đến khi kết thúc, không được phép ngắt quãng. Giả sử  $c_i$  là thời điểm hoàn thành công việc thứ  $i$ . Khi đó, nếu  $c_i > f_i$  thì xem như công việc thứ  $i$  hoàn thành trễ hạn, nếu  $c_i \leq f_i$  thì công việc thứ  $i$  hoàn thành đúng hạn. Cần tìm một trình tự các công việc sao cho số công việc hoàn thành trễ hạn là ít nhất có thể.

- Hãy trình bày một thuật toán heuristic giải bài toán trên và cho biết độ phức tạp của thuật toán đó.
- Minh họa các thuật toán trên qua ví dụ sau:  $n=10$ ;  $(t_i, f_i)$  lần lượt là  $(3; 10), (2; 8), (3; 4), (1; 1), (5; 9), (2; 3), (6; 24), (10; 25), (7; 15), (9; 23)$ ; hãy cho biết số công việc hoàn thành trễ hạn đối với thuật toán trên.
- Thuật toán heuristic ở câu a) có tìm được lời giải tối ưu cho ví dụ ở câu b) hay không? Nếu không hãy chỉ ra một lời giải tối ưu.

### CÂU 3: (3 điểm)

Tại vòng loại bảng  $A$  của một giải vô địch cờ vua gồm 8 kỳ thủ. Các kỳ thủ thi đấu vòng tròn để tính điểm. Biết rằng hiện tại:

Kỳ thủ 1 đã thi đấu với kỳ thủ 3 và 4;

Kỳ thủ 4 đã thi đấu với kỳ thủ 2, 3 và 8;

Kỳ thủ 5 đã thi đấu với kỳ thủ 6 và 8;

Kỳ thủ 7 đã thi đấu với kỳ thủ 1, 4 và 5.

Trong một buổi thi mỗi kỳ thủ chỉ thi đấu một trận. Hãy lập lịch thi đấu cho các trận còn lại sao cho số buổi cần thực hiện là ít nhất.

Hết

Câu 1.

a) Thời gian để các nút rễ stand mọc xong bao giờ  
để nút rễ lòi

Lưu rễ: L<sub>1</sub> L<sub>2</sub> L<sub>3</sub> L<sub>4</sub> L<sub>5</sub> L<sub>6</sub> L<sub>7</sub> L<sub>8</sub> L<sub>9</sub> L<sub>10</sub> L<sub>11</sub> L<sub>12</sub>

Thời gian: 22 16 8 12 8 14 4 12 12 18 6 8

Step Step các lõi rễ theo thời gian stand mọc grain

Lưu rễ: L<sub>1</sub> L<sub>10</sub> L<sub>2</sub> L<sub>6</sub> L<sub>4</sub> L<sub>8</sub> L<sub>9</sub> L<sub>3</sub> L<sub>5</sub> L<sub>12</sub> L<sub>11</sub> L<sub>7</sub>

Thời gian: 22 18 16 14 12 12 12 8 8 8 6 4

Lịch stand mọc các lõi rễ theo các nút rễ

nút rễ 1: L<sub>1</sub> (22), L<sub>8</sub> (12), L<sub>5</sub> (8), L<sub>11</sub> (6)

nút rễ 2: L<sub>10</sub> (18), L<sub>4</sub> (12), L<sub>9</sub> (12), L<sub>7</sub> (4)

nút rễ 3: L<sub>2</sub> (16), L<sub>6</sub> (14), L<sub>3</sub> (8), L<sub>12</sub> (8).

Theo lịch này, thời gian để các nút rễ stand 1, 2, 3 hoàn thành

Công việc cần mẫn là lòi lõi là 48, 46, 48.

Vì thời gian lòi lõi của lõi rễ stand mọc grain là 48

b) Khi nút rễ lòi lõi, grain, và có sự phát triển rễ  
bằng 2/3 nút rễ rễ ta có lòi lõi sau:

nút rễ 1: L<sub>1</sub> (22), L<sub>3</sub> (8), L<sub>12</sub> (8).

nút rễ 2: L<sub>6</sub> (18), L<sub>9</sub> (12), L<sub>10</sub> (6), L<sub>7</sub> (4)

nút rễ 3: L<sub>2</sub> (16), L<sub>4</sub> (12), L<sub>5</sub> (8).

Quá trình: L<sub>6</sub> (21), L<sub>9</sub> (18), ~~L<sub>10</sub> (6)~~.

Thời gian để các nút rễ rễ ra khỏi quá trình lòi lõi là

lòi lõi 38, 40, 36, 39.

Vì thời gian lòi lõi của lõi rễ stand mọc grain là 40.

Câu 2

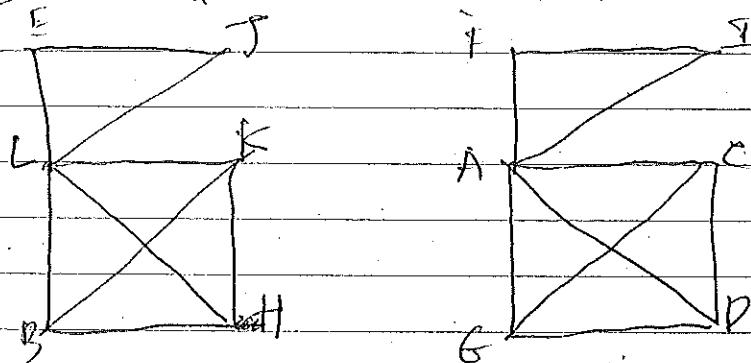
Bước 1: Xây dựng đồ thị.

Gồm một q' sao là một đường cao ở đồ thị.

Có 2 đường cao và 2 đường thẳng - tia - vay với 2 q' sao

Không đặt điểm trong cùng một đường thẳng.

Theo bài ra ta có đồ thị sau:



Bước 2: TS mèo đố-elli

Piee cao cát-ting:

Ding	A	C	D	G	B	H	K	E	J	L	F	I
------	---	---	---	---	---	---	---	---	---	---	---	---

Piee	5	3	3	3	3	3	3	3	2	2	5	2	2
------	---	---	---	---	---	---	---	---	---	---	---	---	---

Sắp xếp cao cát-ting theo bài giải

Ding	A	L	C	D	G	P	H	K	E	J	F	I
------	---	---	---	---	---	---	---	---	---	---	---	---

Piee	5	5	3	3	3	3	3	3	3	2	2	2
------	---	---	---	---	---	---	---	---	---	---	---	---

TS mèo cao cát-ting - Theo cách loại - Chia, bao

mèo	A	L	C	D	G	B	H	F	E	J	F	I
-----	---	---	---	---	---	---	---	---	---	---	---	---

1	X	X										
---	---	---	--	--	--	--	--	--	--	--	--	--

2			X			X			X	X		
---	--	--	---	--	--	---	--	--	---	---	--	--

3				X			X			X	X	
---	--	--	--	---	--	--	---	--	--	---	---	--

4					X			X				
---	--	--	--	--	---	--	--	---	--	--	--	--

Bước 3: Kép lặp lối:

Theo bài giải như ta cần 4 lối để đi đến cao cát-ting  
số lối cần phải cao bài giải:

Kép 1: A, L

Kép 2: C, B, F, F

Kép 3: D, H, J, T

Kép 4: G, K.

Câu 3.

(a) Đặt các biểu minh đê:

 $p$ : 'Gỗn lõai và rau' $q$ : 'Gỗn phan và rau đều là lõai lõai' $r$ : 'Gỗn (cây - hoa)'

Theo bài ra, ta có:

$$J: P \rightarrow q$$

$$K: q \rightarrow r$$

$$\delta: (\underline{(P \rightarrow q)} \wedge \underline{(q \rightarrow r)}) \rightarrow (\underline{r} \rightarrow \bar{P})$$

a) Khi dùng bài toán này biến đổi logic sẽ có

$$((P \rightarrow q) \wedge (q \rightarrow r)) \rightarrow (\underline{r} \rightarrow \bar{P})$$

b) Khi dùng minh bæt træe bæt phiaong phap Vua vay tha

$$((P \rightarrow q) \wedge (q \rightarrow r)) \rightarrow (\underline{r} \rightarrow \bar{P})$$

$$((\bar{p} \vee q) \wedge (\bar{q} \vee r)) \rightarrow (r \vee \bar{P})$$

$$\Leftrightarrow \bar{p} \vee q, \bar{q} \vee r \rightarrow r, \bar{P}$$

$$\Leftrightarrow \left\{ \begin{array}{l} \bar{p}, q \vee r \rightarrow r, \bar{P} \text{ (dоказано)} \\ q, \bar{q} \vee r \rightarrow r, \bar{P} \text{ (?)} \end{array} \right.$$

(1)  $\hookrightarrow$ 

$$\left\{ \begin{array}{l} q, r \rightarrow r, \bar{P} \text{ (dоказано)} \\ q, \bar{q} \rightarrow r, \bar{P} \text{ (?)} \end{array} \right.$$

(2)  $\hookrightarrow$ 

$$q \rightarrow r, \bar{P}, q \text{ (dоказано)}$$

Vay bài toán  $\hookrightarrow$  (dоказано).

Câu 4.

- + Giải quyết bài toán như sau:  $S_0$ ,  $S_1$ ,  $S_2$ ,  $S_3$ ,  $S_4$ ,  $S_5$ ,  $S_6$ ,  $S_7$ ,  $S_8$ ,  $S_9$ .
- +  $g(S)$  là số cách để xếp các ô trống để có  $g = 0$ .
- +  $h(S)$ : Giá trị của  $h(S)$  là số lượng ô trống còn lại.
- +  $f(S) = g(S) + h(S)$

Để giải bài toán, ta sẽ chia thành bốn bước:  $S_0 \rightarrow S_1 \rightarrow S_2 \rightarrow S_3$ .

$S_0$	2	8	3	$g=0$
	1	6	4	$h=4$
	7	5		$f=4$

$S_1$        $S_2$        $S_3$

$S_1$	2 8 3 $g=1$	2 8 3 $g=1$	2 8 3 $g=1$
	1 6 4 $h=5$	1 6 4 $h=3$	1 6 4 $h=5$
	7 5 $f=6$	7 6 5 $f=4$	7 5 $f=6$

$S_4$	2 8 3 $g=2$	2 8 3 $g=2$	2 8 3 $g=2$
	1 4 $h=4$	1 8 4 $h=3$	1 4 $h=3$
	7 5 $f=6$	7 6 5 $f=5$	7 6 5 $f=5$

$S_5$	2 3 $g=3$	2 3 $g=3$
	1 8 4 $h=2$	1 8 4 $h=4$
	7 6 5 $f=5$	7 6 5 $f=3$

$S_6$	1 2 3 $g=4$
	8 4 $h=1$
	7 6 5 $f=5$

$S_7$	1 2 3 $g=5$	1 2 3 $g=5$
	7 8 4 $h=2$	8 7 4 $h=0$
	6 5 $f=7$	7 6 5 $f=5$

Vậy có  $g=5$  bài toán để giải để đáp ứng yêu cầu bài toán.

Câu 5

Bước 0: Ghi nhận độ đo  $V = (\text{Pi}, \text{theory})$ .

Bước 1: Tính xác suất của các sự kiện bằng cách  
ESP (xem dưới).

+ Thúc hiện quay cao (QC)

$$V(QC = \text{mùa}) = (2/5, 3/5)$$

$$V(QC = \text{vịnh murray}) = (3/5, 0/5) = (1, 0)$$

$$V(QC = \text{murray}) = (3/4, 1/4)$$

+ Thúc hiện không đổi (ND)

$$V(ND = \text{mùa}) = (2/5, 1/5)$$

$$V(ND = \text{murray}) = (1/3, 2/3)$$

$$V(ND = \text{lau}) = (3/4, 1/4)$$

+ Thúc hiện đổi đổi (DA)

$$V(DA = \text{mùa}) = (3/6, 3/6)$$

$$V(DA = \text{murray}) = (1/6, 1/6)$$

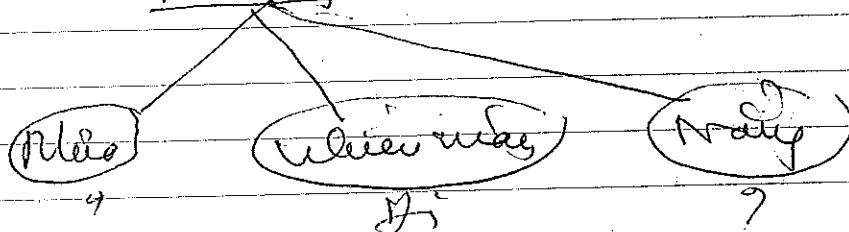
+ Thúc hiện đổi (G)

$$V(G = \text{mùa}) = (5/7, 2/7)$$

$$V(G = \text{murray}) = (3/5, 2/5)$$

Chọn thực hiện quay cao là: Thực hiện không đổi.

[Độ tin cậy]



Bước 2: Ghi nhận độ đo của các sự kiện bằng cách

ESP (tính xác suất quay cao là 1/2).

Màu Nhạt: ĐS Tín, Cao: Projekt deer

1. mùa lao theory

2. murray lao murray theory

3. lau lao nhạt theory

4. murray lao nhạt theory

5. murray lao nhạt murray. ĐT

$V(ND = \text{mùa}) = (1/2, 1/2)$

$V(ND = \text{murray}) = (0/2, 2/2) = (0, 1)$

(câu 5) (tiếp theo)  
+ Phân tích  $\Delta ABC$

$$V(\overrightarrow{PA} = \overrightarrow{BC}) = \left( \frac{0}{2}, \frac{3}{2} \right) = (0, 1)$$

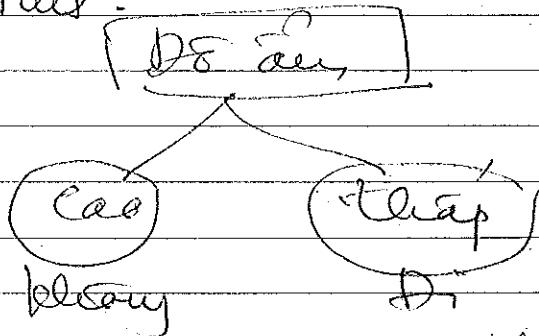
$$V(\overrightarrow{PA} = \overrightarrow{AC}) = \left( \frac{2}{2}, \frac{0}{2} \right) = (1, 0)$$

Thứ tự từ trái sang

$$V(\overrightarrow{B} = \overrightarrow{AC}) = \left( \frac{2}{2}, \frac{1}{2} \right)$$

$$V(\overrightarrow{B} = \overrightarrow{AB}) = \left( \frac{1}{2}, \frac{1}{2} \right)$$

Các vector này  $\Delta ABC$  là: Các vector  $\overrightarrow{PA}$ ,  $\overrightarrow{PB}$ ,  $\overrightarrow{PC}$  là các vector  
tùy ý của  $\Delta ABC$ :



Hỗn hợp:

Tất cả các vector đều là vector của  $\Delta ABC$  và  
tất cả chúng đều là vector của  $\Delta ABC$ .

màu | Mô hình  $\Delta ABC$  có thể là

4 | Cao Thấp Hàng D

5 | Hàng Thấp Hàng Hàng

2 | Cao Hàng Cao Hàng D

8 | Hàng Thấp Cao Hàng D

Phân tích  $\Delta ABC$

$$V(\overrightarrow{PA} = \overrightarrow{BC}) = \left( \frac{2}{2}, \frac{0}{2} \right) = (1, 0)$$

$$V(\overrightarrow{PA} = \overrightarrow{AC}) = \left( \frac{1}{2}, \frac{1}{2} \right)$$

e) Phân tích  $\Delta ABC$ :

$$V(\overrightarrow{PA} = \overrightarrow{AC}) = \left( \frac{2}{2}, \frac{1}{2} \right)$$

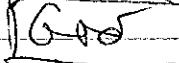
$$V(\overrightarrow{PA} = \overrightarrow{CA}) = (1, 0)$$

+ Phân tích  $\Delta ABC$

$$V(\overrightarrow{B} = \overrightarrow{AC}) = \left( \frac{2}{2}, \frac{0}{2} \right) = (1, 0)$$

$$V(\overrightarrow{B} = \overrightarrow{CA}) = (0, 1)$$

Các vector này  $\Delta ABC$  là: Các vector  $\overrightarrow{PA}$ ,  $\overrightarrow{PB}$ ,  $\overrightarrow{PC}$  là các vector  
tùy ý của  $\Delta ABC$  và là vector của  $\Delta ABC$ .



Cao 5 (lên) (theo)

Chia cao bùn - kẹo, ta có ca gác đòn, sai.

[Penny cao]

(Mèo) (Nhíe, mèo) (Mèo)

Đi

[Đèi cao]

(Cao) (Thiệp)

Đi

(Nhíe) (mèo)

Đi

Đi

bước 4: Tạo lát.

Step

Tùy ý gác đòn - kẹo, ta có gấp (kèo) Cao!

Lát 1: Nhíe Penny Lát Nhíe, mèo - kẹo Đi

Lát 2: Nhíe Penny Cao Mèo và Đèi cao (kèo - kẹo Nhíe)

Lát 3:

- Lát 4: Nhíe Penny Cao Mèo và Gatk (kèo Nhíe) Đi

Lát 5:

Câu 1:

Đặt các biến mảng để

 $p$ : biến số có giá trị xác định $q$ : biến số có giá trị xác định không xác định $r$ : biến số có giá trị xác định

Theo bài toán ta có:

$$\text{L}: p \rightarrow q$$

$$\text{B}: q \rightarrow r$$

$$\text{S}: r \rightarrow p$$

a/ Xác định bài toán bằng biểu thức logic mảng để  
 $((p \rightarrow q) \wedge (q \rightarrow r)) \rightarrow (r \rightarrow p)$ 

b/ Chứng minh bài toán bằng phương pháp phản chứng.

$$((p \rightarrow q) \wedge (q \rightarrow r)) \rightarrow (r \rightarrow p)$$

$$\Leftrightarrow ((p \vee q) \wedge (\neg q \vee r)) \rightarrow \neg r \vee p$$

$$\Leftrightarrow (p \vee q), (\neg q \vee r) \rightarrow \neg r, p$$

$$\Leftrightarrow p \vee q, \neg q \vee r, \neg r, p$$

Chọn các biến mảng để  $\neg r \vee q \vee \neg q \vee r$ 

$$(1) \Leftrightarrow \neg r \vee q, \neg q, p$$

Chọn các biến mảng để  $\neg r \vee q, \neg q$ 

$$(2) \Leftrightarrow \neg r, p$$

- Xác định bài toán bằng cách sử dụng lệnh `if`.

Theo cách xác định bài toán như sau, bài toán được giải.

Câu 2:

a/ Cài đặt giải bài toán để mảng

lệnh 1: Nếu biến a là số chẵn thì đổi biến a thành biến b

lệnh 2: Nếu biến b là số chẵn thì đổi biến b thành biến a

lệnh 3: Nếu biến a là số không chẵn và biến b là số không chẵn

thì đổi biến a thành biến b và biến b thành biến a; else đổi biến

biến a thành biến b; b = b + 1; kết thúc

b/ Điều kiện a, b, c là bài toán để mảng có thể giải

c là số số chẵn và số a và b.

VD: a = 4, b = 5, c = 3. bài toán sẽ là

grau

a = 6, b = 8, c = 5 bài toán kết quả

grau

lớp học đầu: Môn Toán

$$GL = HEP \quad 1$$

$$GL = DT \quad 2$$

$$GL = LA \quad 3$$

$$DN = HEP \quad 2$$

$$DN = BD \quad 1$$

$$HLP = BD \quad 3$$

$$BD = DT \quad 4$$

$$BD = LA \quad 2$$

$$DT = LA \quad 1$$

||. Sắp xếp các giá trị theo TT để mảng 3 đc Sắp xếp qua.

Câu 4 (\*)

(+) dùng giá trị của GTS1 để xác định  $V_i$  ( $i=1, 4$ )

$$GTS_1(V_1): 1 \rightarrow 5 \rightarrow 2 \rightarrow 4 \rightarrow 6 \rightarrow 3 \rightarrow 1$$

$$\text{cost}(V_1) = 1 + 5 + 2 + 10 + 11 + 8 + 3 =$$

$$GTS_1(V_2): 3 \rightarrow 6 \rightarrow 4 \rightarrow 2 \rightarrow 1 \rightarrow 5 \rightarrow 3$$

$$\text{cost}(V_2) = 4 + 2 + 4 + 2 + 1 + 16 = 34$$

$$GTS_1(V_3): 5 \rightarrow 2 \rightarrow 4 \rightarrow 1 \rightarrow 6 \rightarrow 3 \rightarrow 5$$

$$\text{cost}(V_3) = 5 + 2 + 7 + 20 + 11 + 9 = 54$$

$$GTS_1(V_4): 6 \rightarrow 4 \rightarrow 2 \rightarrow 1 \rightarrow 5 \rightarrow 3 \rightarrow 6$$

$$\text{cost}(V_4) = 2 + 4 + 2 + 1 + 16 + 4 = 34$$

Vậy hàng thứ nhất là hàng tốt nhất với

độ dài là 34 và có thể viết là:

$$6 \rightarrow 4 \rightarrow 2 \rightarrow 1 \rightarrow 5 \rightarrow 3 \rightarrow 6$$

(+) GTS1(V): Tùy hàng hàng-tốt nhất đi từ thành

phố V (theo chiều từ trái sang phải) như sau:

chỗ thành phố có chỉ số hàng hàng-tốt nhất

qua.

GTS2: Thực hiện p laris - kinh tree GTS1 ra, elior, hàng hàng-tốt nhất trong số đó là kết quả.

## Câu 2: Kép-Cheo

Thứ tự	a	b	hết	Kết luận
0	0	0		Bạn bè e có bao lồng
1	0	8	2	
2	6	2	3	
3	0	2	1	
4	2	0	3	
5	2	8	2	
6	6	4	3	Dừng
7				

Câu 3: b) Xây dựng đồ thị:

Đát-lee, tất cả các đội bóng là GL, PN, LA, BD, HEP-DT.  
 Các đội bóng đá với nhau  $\Rightarrow$  Tuy Cửng có 15 đường đấu.  
 Số đội có 6-7 đội. Đát-lee-đát-lee, sau còn  
 lại gian nan với các Sắp-đoch-là:

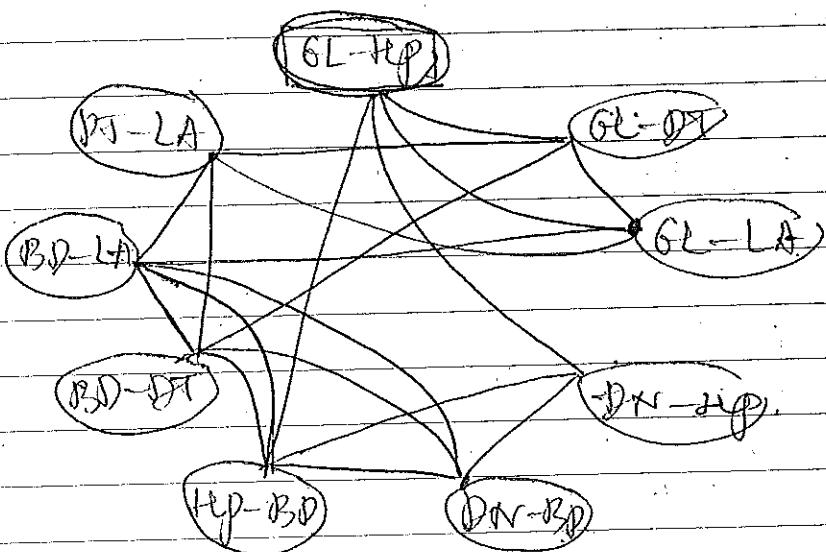
GL-HEP, GL-DT, GL-LA

DH-HEP, DH-BD, HEP-BD, BD-DT, BD-LA, DT-LA

Đát-mỗi-hai đội là một đường của đội (ch)

Cách của đội-đát là đường nối 2-hai, dây có  
cùng-ter một đội bóng.

Theo bài ra là có 7 đội-đát-lee sau:



Lưu ý: hoa-hoa có -kết-quay-kết để biểu diễn đội  
 thi-hang-mua-kết-kết-nhau

(Câu 3 (kép lèo))

Date:

	GL-Hp	GL-DT	GL-LA	DN-Hp	DN-BD	HP-BD	BD-DT	BD-LA	DT-LA
GL-Hp	0	1	1	1	0	1	0	0	0
GL-DT	1	0	1	0	0	0	1	0	1
GL-LA	1	1	0	0	0	0	0	1	1
DN-Hp	1	0	0	0	1	1	0	0	0
DN-BD	0	0	0	1	0	1	1	1	0
HP-BD	1	0	0	1	1	0	1	1	0
BD-DT	0	1	0	0	1	1	0	1	1
BD-LA	0	0	1	0	1	1	1	0	1
DT-LA	0	1	1	0	0	0	1	1	0

Bước 2: Tính mâu - dò (li - lèo) loại theo câu

Dòng	GL-Hp	GL-DT	GL-LA	DN-Hp	DN-BD	HP-BD	BD-DT	BD-LA	DT-LA
Bài	4	4	4	3	4	5	5	5	4

Sắp xếp các dòng theo bài giảm.

Dòng	HP-BD	BD-DT	BD-LA	GL-Hp	GL-DT	GL-LA	DN-BD	DT-LA	DN-Hp
Bài	5	5	5	4	4	4	4	4	3

Tổng mâu

mâu	HP-BD	BD-DT	BD-LA	GL-Hp	GL-DT	GL-LA	DN-BD	DT-LA	DN-Hp
1	X						X		
2		X				X			
3			X						X
4							X	X	
5								X	

Vay vốn - trả lãi

Bước 3: Lập, trả lời - dàn

Trắc 1: Hải Phòng - Bình Dương; Gia Lai - Đồng Tháp

Trắc 2:

Trắc 3:

Trắc 4:

Trắc 5: Đồng Tháp - Long An.

Đề bài: Dùng thiết kế 10 mâu (tổng mâu)  
chứa 10 câu, không xếp câu, dàn - lèo bài

Câu 5.

bước 0: Chọn  $v_0$  là độ dài  $v = (có, không)$

bước 1: Quy  $v_0$  là độ dài của cát (màu sáu) bằng với  $v_1$   
bae dài

+ Theo lính trắc, có:

$$\checkmark (kích cỡ = Không bẩn) = (1, 0)$$

$$\checkmark (kích cỡ = Nhicz) = (1/2, 1/2)$$

$$\checkmark (kích cỡ = Lón) = (2/4, 2/4)$$

+ Theo lính màu sáu

$$\checkmark (màu sáu = Xanh) = (3/3, 0/3) = (1, 0)$$

$$\checkmark (màu sáu = Đỏ) = (1/4, 3/4)$$

+ Theo lính hìnhdây

$$\checkmark (hình dây = Bóng gác) = (1, 0)$$

$$\checkmark (hình dây = Hình vuông) = (0/2, 2/2) = (0, 1)$$

$$\checkmark (hình dây = Hình tròn) = (2/2, 0/2) = (1, 0)$$

$$\checkmark (hình dây = Hình tam giác) = (1/2, 1/2)$$

Chọn  $v_1$  là độ dài của hìnhdây là  $v_1$  phia bên

Hình dây

(Bóng gác)

(Hình vuông)

(Hình tròn)

(Hình tam giác)

có

không

có

?

bước 2:

Xem đây, bao giờ mà có  $v_1$  với  $v_0$  là hìnhdây là hìnhdây

$$\begin{array}{|c|c|c|c|} \hline & kích cỡ & màu sáu & Chọn \\ \hline 1 & Lớn & Xanh & Chọn \\ \hline \end{array}$$

$$\begin{array}{|c|c|c|c|} \hline & kích cỡ & màu sáu & Chọn \\ \hline 2 & Lớn & Xanh & Chọn \\ \hline \end{array}$$

$$\begin{array}{|c|c|c|c|} \hline & kích cỡ & màu sáu & Chọn \\ \hline 3 & Lớn & Đỏ & Chọn \\ \hline \end{array}$$

(vì  $v_1$  là độ dài của hìnhdây là hìnhdây)

$$\checkmark (kích cỡ = Lớn) = (1/2, 1/2)$$

Theo lính kích cỡ,

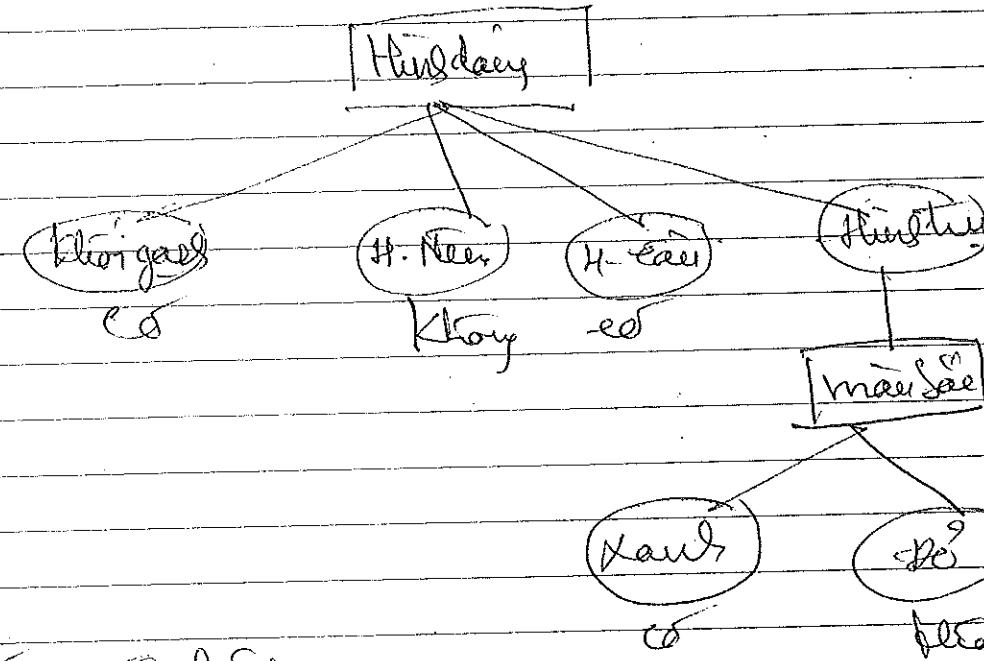
+ Theo lính màu sáu

$$\checkmark (màu sáu = Xanh) = (1, 0)$$

$$\checkmark (màu sáu = Đỏ) = (0, 1)$$

Chọn  $v_2$  là độ dài của  $v_1$  theo lính phia  
lateral

màu sáu

co 15 (trip Voca)Ta ed. Eay queed din Seeb> ~~ta ed.~~ - Replaced.

Tui eay q/din hie, ta ed - Cap leat Saw

Head 1:

2:

3:

4:

5:

Vay hiot not go

Neer hindday kien gaek hoee hing tui <sup>Co</sup> the Clear

2

3:

4:

C&gt;

SRI			
8			
9			

Clear

Co

Khoy

Câu 1.

a) Thời gian

Là phương án mà cao công việc cần để lập thời gian

Lập lại các công việc sau theo thứ tự nào là cao công việc đầu tiên song:

Nếu có máy mài nhôm (là nạp công việc kẽm) may dán keo và (Nếu có chìa khóa sẽ làm máy cưa rãnh) và một thời điểm thì máy và chìa khóa sẽ dùng để làm công việc.

$$b) L^k = (J_3, J_2, J_5, J_1, J_4, J_6)$$

Ta có phép tính sau

$$p_1 : J_3 (13) \quad J_6 (6)$$

$$p_2 : J_2 (10) \quad J_4 (6)$$

$$p_3 : J_5 (9) \quad J_1 (7)$$

Thời gian cao nhất là  $p_1, p_2, p_3$  là 13 - 10 - 9 = 16

Vậy thời gian hoàn thành là cao công việc là 19. (các máy)

Câu 2

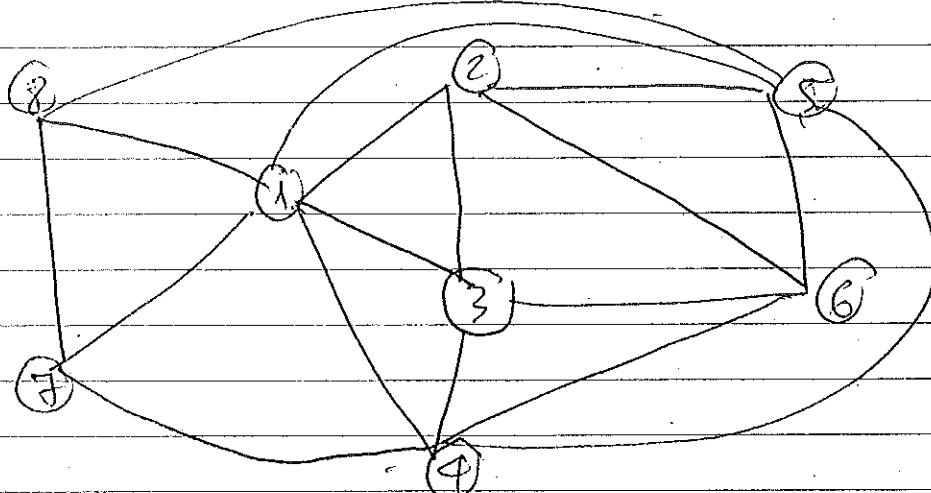
bài 1: Xây dựng đồ thị

Gọi mỗi câu đố là 1 đỉnh của đồ thị

Sau đó hãy diễns ra đường đi từ câu

về câu

Theo bài toán có đồ thị sau



bài 2: Giải ma trận - Viết lời - Lập bảng

Dính	1	2	3	4	5	6	7	8
------	---	---	---	---	---	---	---	---

Phát	6	4	4	5	5	4	3	3
------	---	---	---	---	---	---	---	---

Sắp xếp các đỉnh - Theo bắc giảm

Dính	1	4	5	2	3	6	7	8
------	---	---	---	---	---	---	---	---

Phát	6	5	5	2	4	4	3	3
------	---	---	---	---	---	---	---	---

Dính	1	2	3	4	5	6	7	8
------	---	---	---	---	---	---	---	---

màu	1	x						
-----	---	---	--	--	--	--	--	--

2		x		x				
---	--	---	--	---	--	--	--	--

3			x		x		x	
---	--	--	---	--	---	--	---	--

Bài 3: Lập bảng:

Số bước cần để trả câu là 3:

bước 1: trả câu đố 1,6

2: . . . . . 2,4,8

3: . . . . . 3,5,7

(Câu 3)

$$\text{a/ } \{(P \wedge q) \rightarrow s, (q \wedge s) \rightarrow r, q, p\} \rightarrow s$$

$$\Leftrightarrow P \vee \bar{q} \vee s, \bar{q} \vee \bar{s} \vee r, q, p \rightarrow s$$

$$\Leftrightarrow P \vee \bar{q} \vee s, \bar{q} \vee \bar{s} \vee r \rightarrow s, \bar{q}, \bar{p}$$

$$\Leftrightarrow \left\{ \begin{array}{l} P \\ \bar{q} \\ s \end{array} \right. , \bar{q} \vee \bar{s} \vee r \rightarrow s, \bar{q}, \bar{p} \quad (\text{dùi c/m})$$

$$\Leftrightarrow \left\{ \begin{array}{l} P \\ \bar{q} \\ s \end{array} \right. , \bar{q} \vee \bar{s} \vee r \rightarrow s, \bar{q}, \bar{p} \quad (\text{dùi e/m})$$

Vay KB1 suy ra dùi s.

$$\text{b/ } \{(q \rightarrow s) \wedge (p \rightarrow s)\} \rightarrow (q \vee p) \rightarrow s$$

$$\Leftrightarrow \{(\bar{q} \vee s) \wedge (\bar{p} \vee s)\} \rightarrow \bar{q} \vee \bar{p} \vee s$$

$$\Leftrightarrow \bar{q} \vee s, \bar{p} \vee s, \bar{q} \vee \bar{p} \rightarrow \bar{q} \vee \bar{p}, s$$

$$\Leftrightarrow \bar{q} \vee s, \bar{p} \vee s, q \vee p, \bar{s} \quad (1)$$

Chọn cặp mảng để  $\bar{q} \vee s$  và  $q \vee p$ .

$$(1) \Leftrightarrow p \vee s, \bar{p} \vee s, \bar{s} \quad (2)$$

Chọn cặp mảng để  $p \vee s \wedge \bar{p} \vee s$

$$(2) \Leftrightarrow s \vee s, \bar{s}$$

$\Leftrightarrow s, \bar{s}$  là cặp bù nhau

Vay KB2 suy ra dùi  $(q \vee p) \rightarrow s$ .

Câu 4.

Gồm hàng thứ nhất là các bài toán là định số

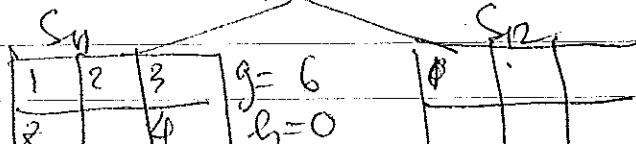
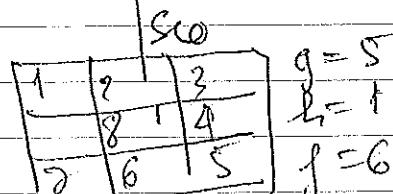
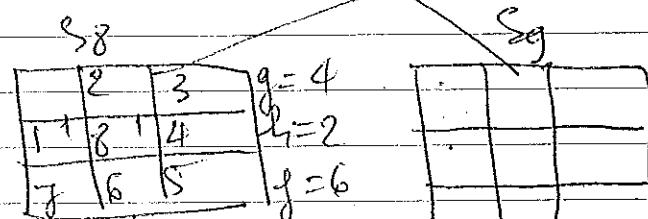
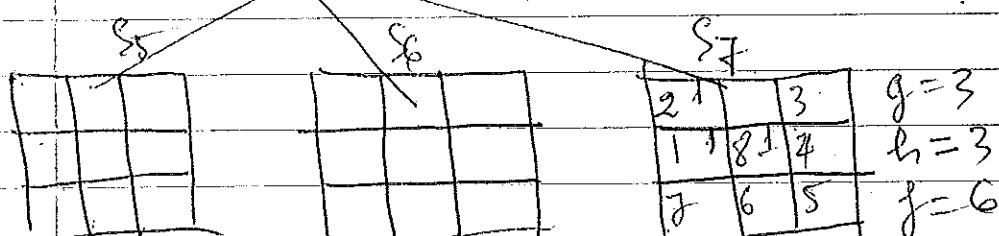
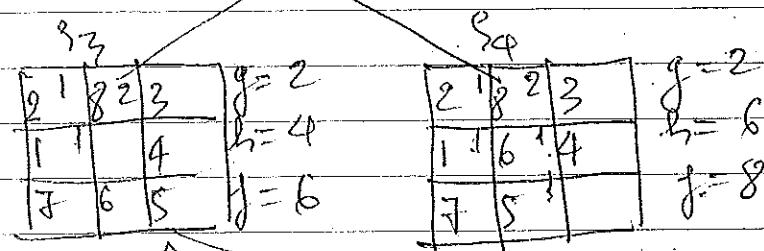
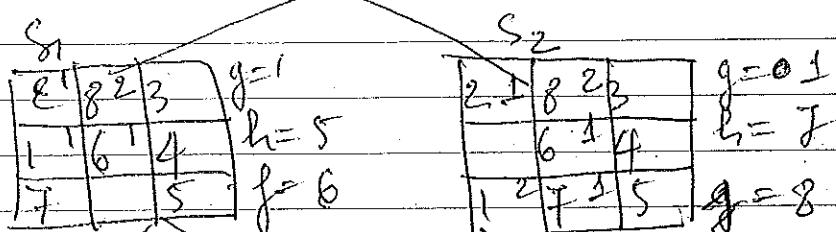
+  $g(S)$  là số cần (số bài toán) để đường đi  
tỷ lệ nhất - kí sốn S,

+  $h(S)$ : Giá trị của đường đi từ kí sốn S  
đến kí sốn S' sau

+  $f(S) = g(S) + h(S)$

Theo bài toán, ta có cách biểu diễn là giải như sau

$2'$	$8^2$	$3$	$g=0$
$1'$	$6^1$	$4$	$h=6$
$7$	$1^7$	$5$	$f=6$



## Câu 4 (kiếp véo)

Vay có  $g = 6$  bước chuyển đổi từ  $S_0$  đến  $S_{11}$  theo thứ tự sau:  
 Bước 1:  $S_0 \rightarrow S_1 \rightarrow S_2 \rightarrow S_3 \rightarrow S_4 \rightarrow S_5 \rightarrow S_6 \rightarrow S_7 \rightarrow S_8 \rightarrow S_9 \rightarrow S_{10} \rightarrow S_{11}$ .

## Câu 5

a) Bước 0: Gọi vec tơ  $\vec{d}_0$  do  $v = (\text{Thịt lợn mía}, mía)$

Bước 1:  $(\vec{d}_0 + \vec{d}_1) \rightarrow \vec{d}_1$  (do  $\vec{d}_0$  là số lượng và  $\vec{d}_1$  là số lượng vật chất ban đầu).

+ Theo tính chất:

$$\vec{v}(\vec{d}_0 + \vec{d}_1) = \left( \frac{3}{3}, 0 \right) = (1, 0)$$

$$\vec{v}(\vec{d}_1) = \text{Mía} = \left( \frac{4}{5}, \frac{4}{5} \right)$$

+ Theo tính áp suất

$$\vec{v}(\vec{d}_0) = \text{Cao} = \left( \frac{2}{4}, \frac{2}{4} \right)$$

$$\vec{v}(\vec{d}_1) = (\text{Trứng, bì}) = \left( 0, 1 \right)$$

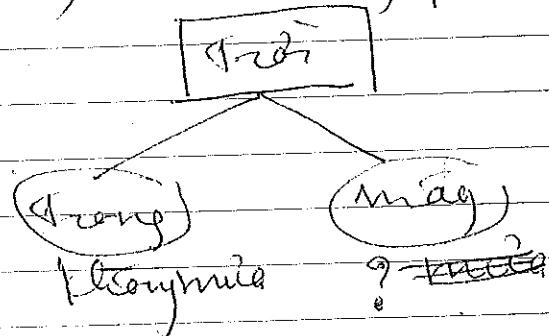
$$\vec{v}(\vec{d}_0 + \vec{d}_1) = \text{Thấp} = \left( \frac{2}{3}, \frac{1}{3} \right)$$

+ Thêm bước (giả)

$$\vec{v}(\vec{d}_0 + \vec{d}_1) = \text{Bé} = \left( \frac{2}{5}, \frac{3}{5} \right)$$

$$\vec{v}(\vec{d}_0 + \vec{d}_1) = \text{Mùn} = \left( \frac{2}{3}, \frac{1}{3} \right)$$

Còn thêm bước là  $\vec{d}_2$  làm thêm hai phần bì  
 (do  $\vec{d}_2$  có vec tơ đơn vị ~~vec tơ~~  $\vec{d}_1$  (đã định rõ là 3 mía)). Ta có kết quả



## Bước 2:

Lập bảng dữ liệu ứng với bước bàng May.

#	Áp suất	Giả	Giá trị
1	Cao	Giả	Mùn
2	Mùn	Giả	Mùn
3	Trứng	Giả	Mùn
4	Thấp	Giả	Mùn
5	Cao	Bé	Mùn
6	Bé	Bé	Mùn
7	Trứng	Bé	Mùn
8	Thấp	Bé	Mùn

(câu 4)

Trang 3

No:

Date:

Tính độ lệch chuẩn (lý thuyết)

+ Lý thuyết về biến

$$\checkmark (\text{ap suất} = \text{Cee}) = (0, 2/2) = (0, 1)$$

$$\checkmark (\text{ap suất} = \text{Army birds}) = (0, 1)$$

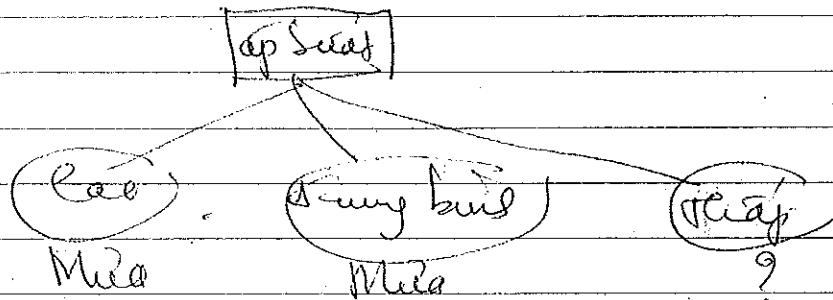
$$\checkmark (\text{ap suất} = \text{Thiệp}) = (1/2, 1/2)$$

+ Lý thuyết Grö

$$\checkmark (\text{Grö} = \text{Bắc}) = (0, 3/2) = (0, 1)$$

$$\checkmark (\text{Grö} = \text{Nam}) = (1/2, 1/2)$$

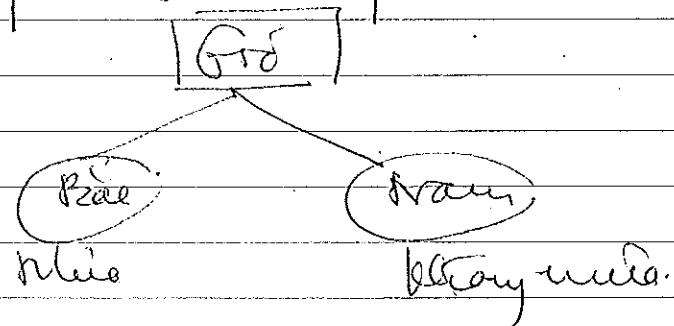
Chọn lý thuyết về biến (lý thuyết giá trị).



bài 3.

Kết quả tổng hợp về ap suất bay rải.

II	Grö	Quyết định
5	Bắc	Mèo
7	Nam	Army birds

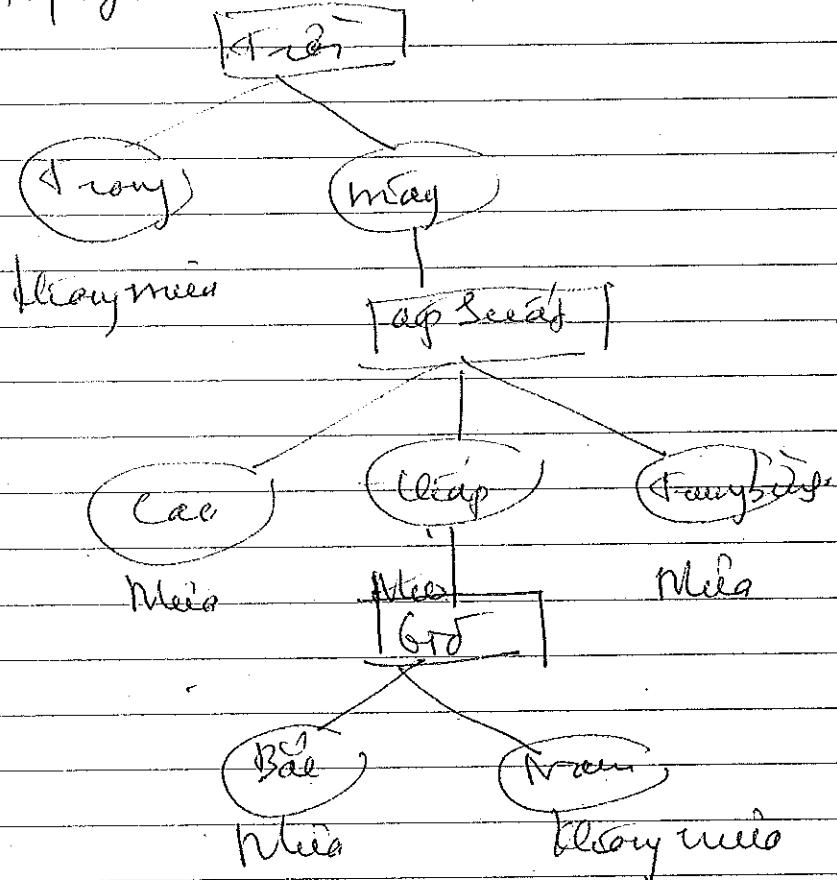


Câu 4 (H)

No: Đề 3

Date:

Ta sẽ giải quyết như sau:



Bước 4: ~~1~~ Rập hét kinh với cay gai.

Lời 1: Nếu ~~đó~~ Trong Khoa mua

2: Nếu ~~đó~~ may ~~đó~~ Pop seeds Fairytail không  
điều kiện Cao (lùi) Mia

3: Nếu ~~đó~~ may ~~đó~~ điều kiện Gia Bao Mia

4: ~~Nếu~~ Khoa mua

b) ~~đó~~ gai quyết định Khoa mua

1) ~~đó~~ gai quyết định Mia.

Câu 1.

$$a) (\overline{p \rightarrow q}, \overline{s \rightarrow r}, (\overline{q \vee s}) \rightarrow t, \overline{t}) \rightarrow (\overline{r} \wedge \overline{p})$$

$$\Leftrightarrow \overline{p \vee q}, \overline{r \vee s}, \overline{q \vee s} \vee \overline{t}, \overline{t} \rightarrow \overline{r \vee p}$$

$$\Leftrightarrow \overline{p \vee q}, \overline{r \vee s}, \overline{q \vee s} \vee \overline{t}, \overline{t}, \overline{r \wedge p} \quad (1)$$

$$\Leftrightarrow \text{Chia } \overline{p \vee q} \text{ thành } \overline{p} \vee \overline{q} \text{ và } \overline{r \wedge p}$$

$$(1) \Rightarrow \overline{q \vee r}, \overline{r \vee s}, \overline{q \vee s} \vee \overline{t}, \overline{t} \quad (2)$$

$$\text{Chia } \overline{q \vee r} \text{ thành } \overline{q} \vee \overline{r} \text{ và } \overline{r} \vee \overline{s}$$

$$(2) \Leftrightarrow \overline{q \vee s} \vee \overline{s} \vee \overline{t}, \overline{t} \quad (3)$$

$$\text{Chia } \overline{q \vee s} \text{ thành } \overline{q} \vee \overline{s} \text{ và } \overline{q} \vee \overline{s} \vee \overline{t}$$

$$(3) \Leftrightarrow \overline{t}; \overline{t} \text{ là lập biếu đối tự eş.}$$

Vậy biểu thức đã chia là đúng.

$$b) \vdash b = \{ (a \wedge b) \rightarrow c, (b \wedge c) \rightarrow d, \overline{d} \} \rightarrow (a \rightarrow b)$$

$$\Leftrightarrow \overline{a \wedge b} \vee c, \overline{b \wedge c} \vee d, \overline{d} \rightarrow \overline{a}, \overline{b}$$

$$\Leftrightarrow \overline{a \wedge b} \vee c, \overline{b \wedge c} \vee d \rightarrow \overline{a}, \overline{b}, \overline{d}$$

$$\Leftrightarrow \begin{cases} \overline{a \wedge b} \vee c, \overline{b \wedge c} \\ \overline{a \wedge b} \vee c, \overline{d} \end{cases} \rightarrow \overline{a}, \overline{b}, \overline{d} \quad (1)$$

$$\Leftrightarrow \begin{cases} \overline{a \wedge b} \vee c, \overline{b \wedge c} \\ \overline{a \wedge b} \vee c, \overline{d} \end{cases} \rightarrow \overline{a}, \overline{b}, \overline{d} \quad (\text{điều kiện})$$

$$(1) \Leftrightarrow \overline{a \vee b \vee c}, \overline{b \vee c} \rightarrow \overline{a}, \overline{b}, \overline{d}$$

$$\left\{ \begin{array}{l} \overline{a}, \overline{b} \vee \overline{c} \rightarrow \overline{a}, \overline{b}, \overline{d} \\ \overline{b} \vee \overline{c}, \overline{b} \vee \overline{c} \rightarrow \overline{a}, \overline{b}, \overline{d} \end{array} \right. \quad (\text{điều kiện})$$

$$\left\{ \begin{array}{l} \overline{b} \vee \overline{c}, \overline{b} \vee \overline{c} \rightarrow \overline{a}, \overline{b}, \overline{d} \\ \overline{b} \vee \overline{c}, \overline{d} \rightarrow \overline{a}, \overline{b}, \overline{d} \end{array} \right. \quad (\text{bằng điều kiện})$$

Vậy  $\vdash b$  bằng suy ra điều  $(a \rightarrow b)$

Câu 2

Bài 1: Xây dựng đồ thị

Gồm các chẽ là 1 định của đồ thị

Các chẽ là không chẽ ra đồng thời được nối thành  
cách của đồ thịTheo đề bài ta có đồ thị sau (biểu diễn bằng  
ma trận Kép)

A B C D E F G H I

A 0 1 0 1 1 0 0 1 1

B 1 0 1 1 0 0 0 1 1

C 0 1 0 1 0 0 0 0 0

D 1 1 1 0 1 1 0 1 1

E 1 0 0 1 0 0 0 0 0

F 0 0 0 1 0 0 1 1 1

G 0 0 0 0 0 1 0 1 0

H 1 1 0 1 0 1 1 0 1

I 1 1 0 1 0 1 0 1 0

Bài 2: Số véo đồ thị theo thứ tự tăng bao

Định A B C D E F G H I

Bé 5 5 2 7 2 4 2 6 5

Sắp xếp các đỉnh theo béo giảm

Định D H A B I F C E G

Bé 7 6 5 5 5 4 2 2 2

Định D H A B I F C E G

mà

1 X X

2 X X X X

3 X X

4 X

5 X

Bài 3: Kết luận phương án lắp ráp

Bài 1: Tô chẽ các chẽ đồ thị D, G

2

3

4

H, S, E

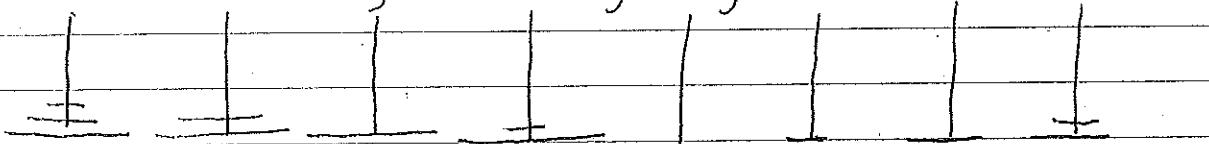
A, F

B

Câu 2.

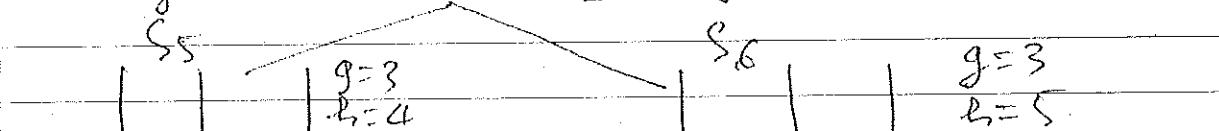
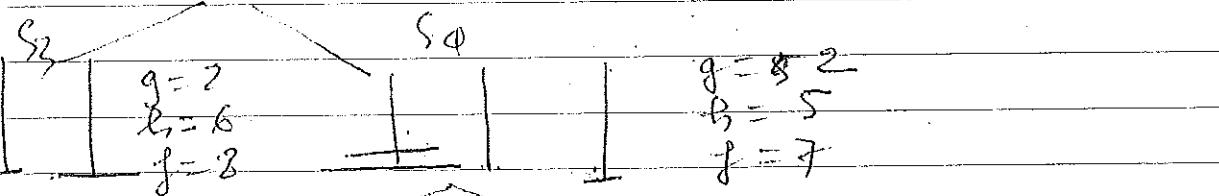
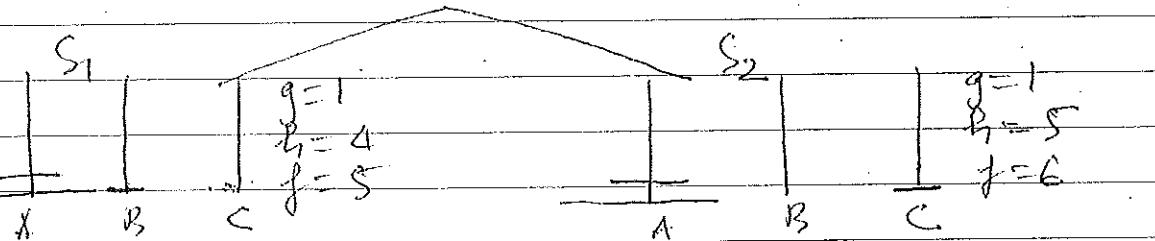
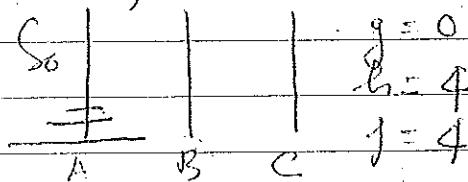
- + Sử dụng biến characteristic function ( $f(s)$ ) là để làm ít nhất phần tử chuyển đổi để đạt được trạng thái đích.
- + Trong trạng thái đầu của bài toán là  $S_0$ .
- +  $g(s)$  là số bước di chuyển của đường đi ngắn nhất để đến đích  $S$ .
- +  $f(s) = g(s) + h(s)$
- +  $h(s)$ : Giá trị của nó là tổng số bước cần đi từ trạng thái hiện tại đến đích. Trong bài này  $h(s)$  là số bước ít nhất phải di chuyển để đến trạng thái đích  $S$  về trạng thái đích.
- +  $f(s) = g(s) + h(s)$ .

Các trường hợp sau có thể có ở состояния  $C$  và giá trị của chúng, xác định bằng  $h$  là như sau:



$h=0$     $h=1$     $h=2$     $h=3$     $h=4$     $h=5$     $h=6$     $h=7$   
 (3điều)   (điều 1)   (điều 2)   (điều 3)   (không điều điều)   (điều 2)   (2điều  
 riêng rẽ)   (điều 3)   (không điều điều)   (không điều điều)   và điều riêng

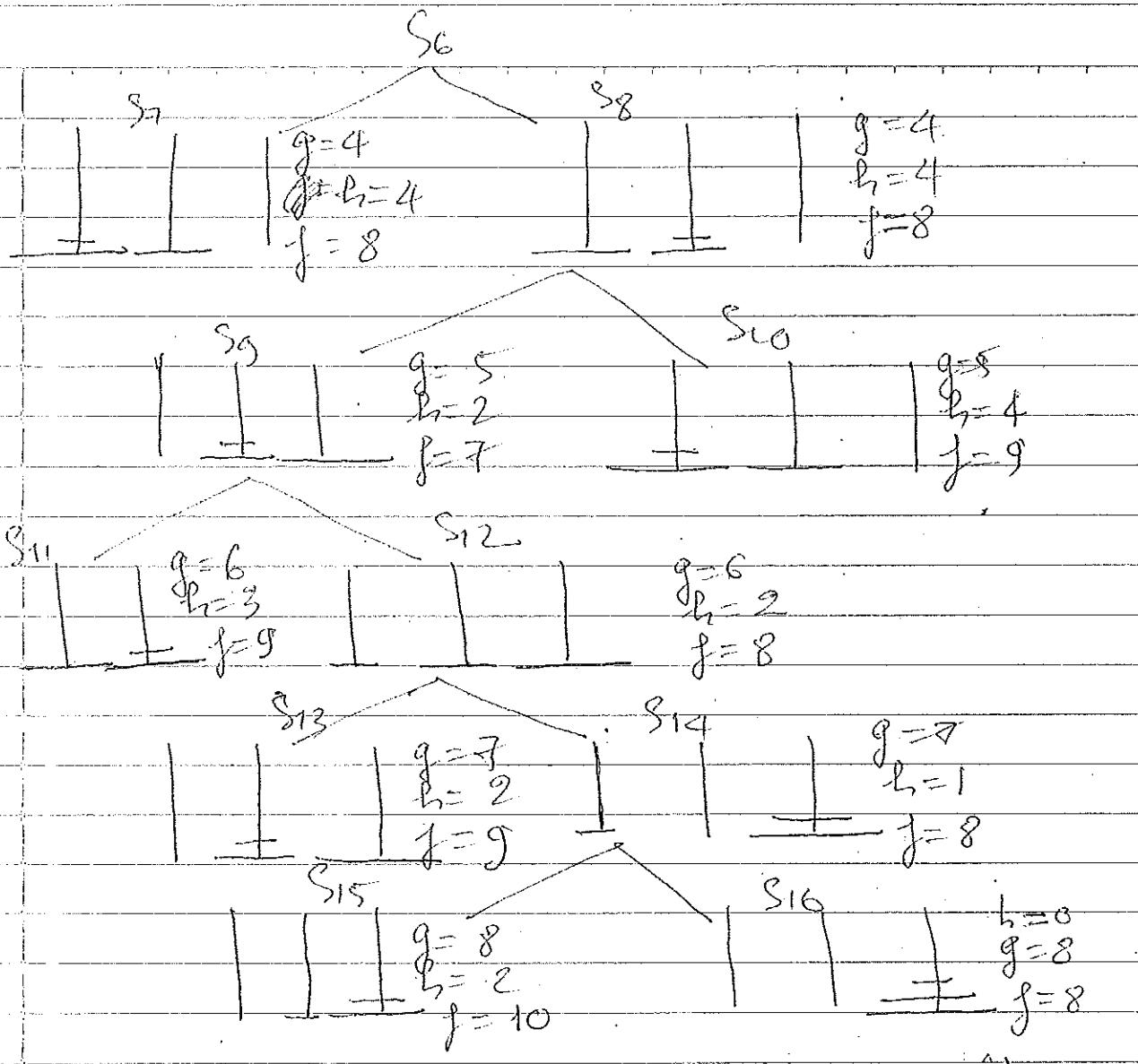
Theo bài ra, ta có cay biển dưới đây giải như sau:



Câu 3 (H-H)

Đề 4

Date: \_\_\_\_\_



Vay cõ  $g=8$  bude clugen de' di-lù-hang-hai-bæ  
tâu vê hang-hai di-ch:

$S_0 \rightarrow S_1 \rightarrow S_2 \rightarrow S_4 \rightarrow S_6 \rightarrow S_8 \rightarrow S_9 \rightarrow S_{12} \rightarrow S_{14} \rightarrow S_{16}$

Câu 4.

Chia các câu liệt kín thành 2 nhóm.

N<sub>1</sub> gồm các câu liệt D<sub>i</sub> thỏa mãn ai ≤ bi

$$N_1 = \{ D_1, D_3, D_5, D_6, D_{10}, D_{11}, D_{12} \}$$

N<sub>2</sub> gồm các câu liệt D<sub>i</sub> thỏa mãn ai > bi

$$N_2 = \{ D_2, D_4, D_7, D_8, D_9 \}$$

Sắp xếp các câu liệt trong N<sub>1</sub> theo chiều dài, theo ai

$$N_1 = \{ D_6(2), D_{11}(3), D_1(4), D_3(7), D_{12}(8), D_5(9), D_{10}(11) \}$$

Sắp xếp các câu liệt trong N<sub>2</sub> theo bi.

$$N_2 = \{ D_8(12), D_4(9), D_7(9), D_2(7), D_9(4) \}$$

Vay kết quả là (nối N<sub>2</sub> vào trước N<sub>1</sub>) :

$$\{ D_6, D_{11}, D_1, D_3, D_{12}, D_5, D_{10}, D_8, D_4, D_7, D_2, D_9 \}$$

Câu 5

c) Bài 0: gán xác suất cho  $X = (\text{Xuân}, \text{Hè}, \text{Thứ}, \text{Đông})$

Bài 1: Tính xác suất cho概率 của các chuỗi liên kết

+ Chuỗi liên kết Thứ nhất

$$V(\text{Thứ nhất} = \text{Nhịp}) = (1/4, 1/4, 1/4, 1/4)$$

$$V(\text{Thứ nhất} = \text{Nắng}) = (0, 0, 0, 1)$$

$$V(\text{Thứ nhất} = \text{Mưa}) = (1/5, 1/5, 2/5, 1/5)$$

+ Chuỗi liên kết lá cây:

$$V(\text{Lá cây} = \text{Vàng}) = (0, 2/3, 1/3)$$

$$V(\text{Lá cây} = \text{Xanh}) = (2/5, 2/5, 0, 1/5)$$

$$V(\text{Lá cây} = \text{Rừng}) = (0, 0, 1/4, 3/4)$$

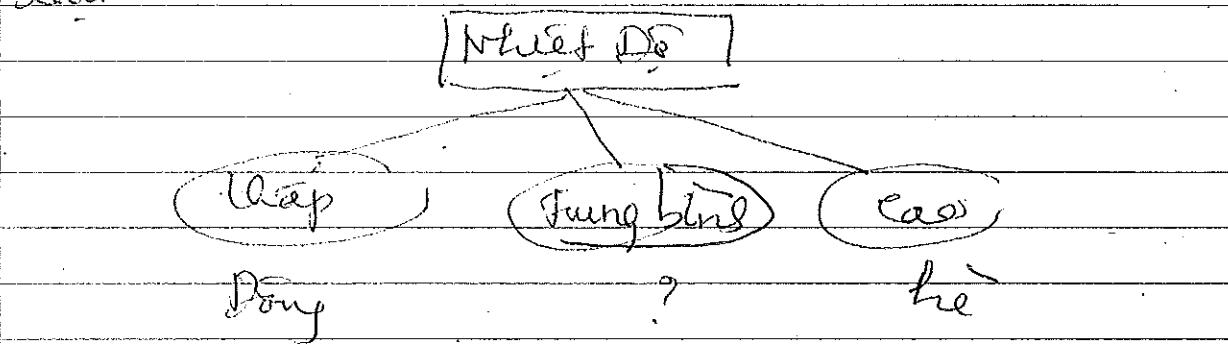
+ Chuỗi liên kết Nhịp Độ

$$V(\text{Nhịp độ} = \text{Trung bình}) = (2/5, 0, 3/5, 0)$$

$$V(\text{Nhịp độ} = \text{Thấp}) = (0, 0, 0, 1)$$

$$V(\text{Nhịp độ} = \text{Cao}) = (0, 1, 0, 0)$$

Chọn chuỗi liên kết nhịp độ làm chuỗi liên kết phái loài



Bài 2: Tính xác suất

CSDL ứng với nhịp độ bằng Trung bình

#1 Thứ nhất là cây Budgeting (Mùa)

1 Nhịp Vàng Thủ

4 Nhịp Vàng Thủ

8 Nắng Xuân Xuân

16 Mưa Rừng Thủ

12 Mưa Xuân Xuân

Lưu ý: Các số ở đây là xác suất của các chuỗi liên kết.

+ Chuỗi liên kết Thứ nhất:

$$V(\text{Thứ nhất} = \text{Nắng}) = (1/2, 0, 1/2, 0)$$

$$V(\text{Thứ nhất} = \text{Mưa}) = (1/6, 0, 3/6, 0)$$

(câu 5 (trên bìa))

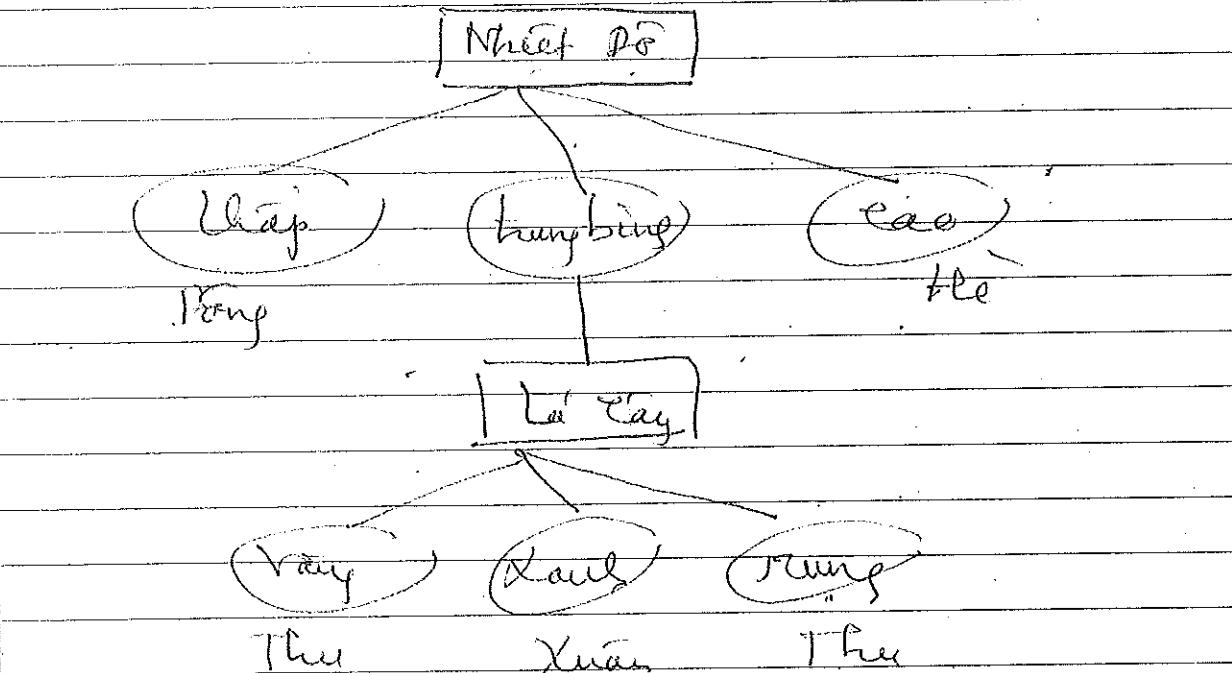
+ Thời tiết là cây

$$Y(\text{Lá cây} = \text{Vàng}) = (0, 0, 1, 0)$$

$$Y(\text{Lá cây} = \text{Xanh}) = (1, 0, 0, 0)$$

$$Y(\text{Lá cây} = \text{Rừng}) = (0, 0, 1, 0)$$

Chọn: Thời tiết là cây sau: Thời tiết phân loại  
Tổng Cây quyết định sau:



Bước 3: Tạo luận.

Tổ cây quyết định thời tiết, ta có lập luận sau:

1: Nếu nhiệt độ thấp là mùa Đông

2: Nếu nhiệt độ cao là mùa hè

3: Nếu nhiệt độ trung bình và lá cây xanh  
thì là mùa Xuân,4: Nếu nhiệt độ trung bình và lá cây vàng  
thì là cây vàng thì là mùa Thu.

b/Mùa	Thời tiết	Lá cây	Nhiệt Đè	Quyết định (Mùa)
13	Mùa	Vàng	Thấp	Đông
14	Tuyết	Rừng	Trung bình	Thu

## GỢI Ý ĐỀ 5

### CÂU 1 (1.5 điểm):

a. Thời gian (tính bằng giờ) để các nhân viên đánh máy xong luận văn tốt nghiệp là:

Luận văn: L1 L2 L3 L4 L5 L6 L7 L8 L9 L10 L11 L12

Thời gian: 20 14 7 10 6 12 5 8 10 15 4 6

Sắp xếp các luận văn theo thời gian đánh máy giảm: (0.5)

Luận văn: L1 L10 L2 L6 L4 L9 L8 L3 L5 L12 L7 L11

Thời gian: 20 15 14 12 10 10 8 7 6 6 5 4

Lịch đánh máy các luận văn của nhân viên:

Nhân viên 1: L1(20), L9(10), L5(6), L11(4)

Nhân viên 2: L10(15), L4(10), L8(8), L12(6)

Nhân viên 3: L2(14), L6(12), L3(7), L7(5)

Theo lịch trên, thời gian để các nhân viên 1,2,3 hoàn thành công việc của mình lần lượt là: 40,39,38.

Vậy thời gian hoàn thành việc đánh máy cho các luận văn là 40 (0.5)

b. Khi người quản lý tham gia, và có công suất đánh máy bằng 1/2 nhân viên. (0.5 điểm)

Ta có lịch sau đánh máy các luận văn của nhân viên và quản lý:

Nhân viên 1: L1(20), L8(8), L7(5)

Nhân viên 2: L10(15), L9(10), L12(6), L11(4)

Nhân viên 3: L2(14), L4(10), L3(7)

Quản lý: L6(24), L5(12)

Thời gian để các nhân viên và người quản lý hoàn thành công việc của mình lần lượt là 33,35,31,36

Vậy thời gian hoàn thành việc đánh máy cho các luận văn khi có người quản lý tham gia là 36 (0.5)

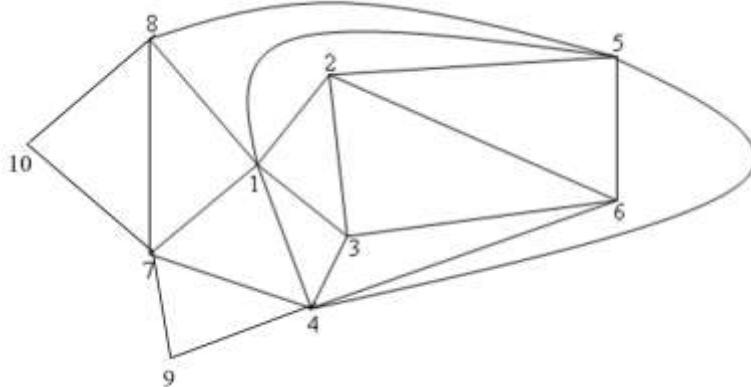
### CÂU 2 (2.0 điểm):

Bước 1: Xây dựng đồ thị (0.5 điểm)

Gọi mỗi chủ đề là một đỉnh của đồ thị

Nối hai chủ đề không diễn ra đồng thời là một cạnh của đồ thị

Theo bài ra ta có đồ thị sau:



Bước 2: Tô màu theo thuật toán tham lam

(1.0 điểm)

Đỉnh	1	2	3	4	5	6	7	8	9	10
Bậc	6	4	4	6	5	4	5	4	2	2

Sắp xếp các đỉnh theo bậc giảm

Đỉnh	1	4	5	7	2	3	6	8	9	10
Bậc	6	6	5	5	4	4	4	4	2	2

Màu	1	2	3	4	5	6	7	8	9	10
1	x					x			x	x
2		x		x				x		
3			x		x		x			

Bước 3: Lập lịch

(0.5 điểm)

Số buổi cần để tổ chức tất cả các buổi hội thảo là: 3

Buổi 1: Các chủ đề 1,6,9,10

Buổi 2: Các chủ đề 2,4,8

Buổi 3: Các chủ đề 3,5,7

### CÂU 3 (2.0 điểm):

a.Xây dựng bài toán bằng biểu thức logic mệnh đề

(0.50 điểm)

b.Chứng minh bài toán bằng phương pháp Vương Hạo

(0.75 điểm)

Chứng minh bài toán bằng phương pháp Robinson

(0.75 điểm)

### CÂU 4 (1.5 điểm):

Đưa về bài toán đóng 2 bình (có nhiều cách để làm điều này)

(0.5 điểm)

Thực hiện theo thuật toán đóng nước như đã học

(1.0 điểm)

### CÂU 5 (3.0 điểm):

Bước 0: Gọi vectơ độ đo  $v = (\text{Xuân}, \text{Hè}, \text{Thu}, \text{Đông})$

(0.25đ)

Bước 1:Tính vecto độ đo của các thuộc tính ở CSDL ban đầu

(0.75đ)

+Thuộc tính **thời tiết**

$$V_{(\text{thời tiết=nắng})} = (1/4, 1/4, 1/4, 1/4);$$

$$V_{(\text{thời tiết=tuyết})} = (0, 0, 0, 1);$$

$$V_{(\text{thời tiết=mưa})} = (1/5, 1/5, 2/5, 1/5);$$

+Thuộc tính lá cây

$$V_{(\text{lá cây=Vàng})} = (0, 0, 2/3, 1/3);$$

$$V_{(\text{lá cây=Xanh})} = (2/5, 2/5, 0, 1/5);$$

$$V_{(\text{lá cây=Rụng})} = (0, 0, 1/4, 3/4);$$

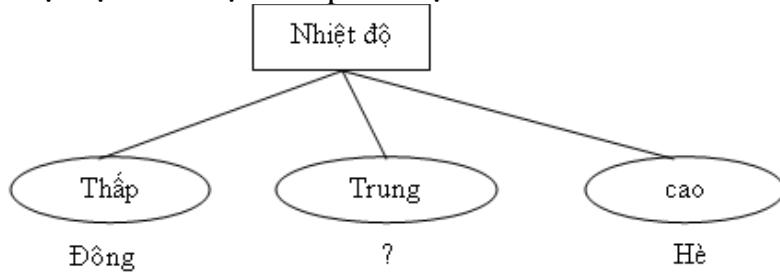
+Thuộc tính Nhiệt độ

$$V_{(\text{nhiệt độ=Trung bình})} = (2/5, 0, 3/5, 0);$$

$$V_{(\text{nhiệt độ=thấp})} = (0, 0, 0, 1);$$

$$V_{(\text{nhiệt độ=cao})} = (0, 1, 0, 0);$$

Chọn thuộc tính Nhiệt độ làm thuộc tính phân loại:



Bước 2: CSDL ứng với Nhiệt độ bằng Trung bình (0.75đ)

#	Thời tiết	Lá cây	Quyết định (Mùa)
2	Nắng	Vàng	Thu
3	Mưa	Vàng	Thu
7	Nắng	Xanh	Xuân
11	Mưa	Rụng	Thu
12	Mưa	Xanh	Xuân

Tính vector độ đo của các thuộc tính

+Thuộc tính **thời tiết**

$$V_{(\text{thời tiết}=\text{nắng})} = (1/2, 0, 1/2, 0);$$

$$V_{(\text{thời tiết}=\text{mưa})} = (1/3, 0, 2/3, 0);$$

+Thuộc tính lá cây

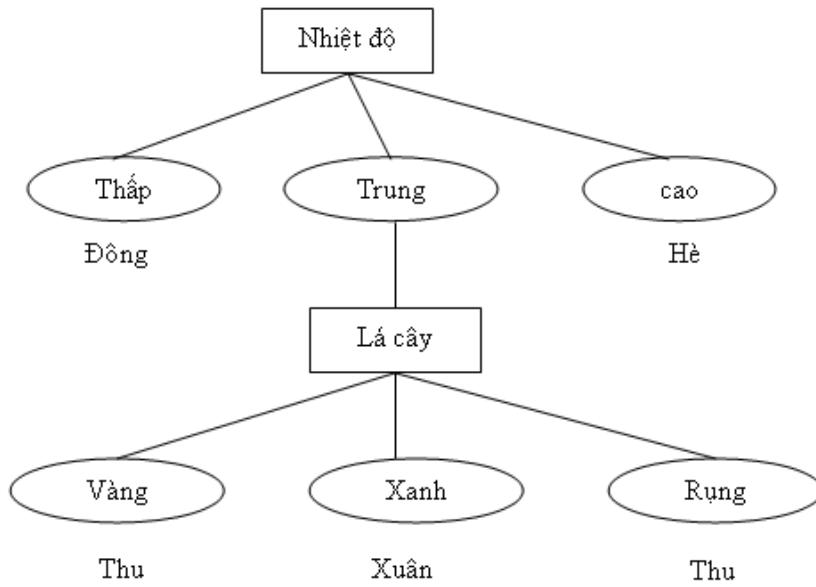
$$V_{(\text{lá cây}=\text{Vàng})} = (0, 0, 1, 0);$$

$$V_{(\text{lá cây}=\text{Xanh})} = (1, 0, 0, 0);$$

$$V_{(\text{lá cây}=\text{Rụng})} = (0, 0, 1, 0);$$

Chọn thuộc tính Lá cây làm thuộc tính phân loại

Từ các bước trên, ta có cây quyết định sau:



**Bước 3: Tập luật** Từ cây quyết định trên, ta có tập luật sau: (0.75đ)

- Luật 1: Nếu nhiệt độ thấp thì mùa Đông
- Luật 2: Nếu nhiệt độ cao thì mùa Hè
- Luật 3: Nếu nhiệt độ trung bình và lá cây Xanh thì mùa Xuân
- Luật 4: Nếu nhiệt độ trung bình và lá cây màu Vàng hoặc lá rụng thì mùa Thu

(0.50)

#	Thời tiết	Lá cây	Nhiệt độ	Quyết định (Mùa)	Luật
x	Mưa	Vàng	Cao	Hè	Luật 2
y	Tuyết	Rụng	Trung bình	Thu	Luật 4
z	Tuyết	Vàng	Trung bình	Thu	Luật 4

Hết