

# Bài tập Seminar chuyên đề

## Chương 1.

### Thuật toán chia để trị

#### 1.1.

Cho dãy  $n$  số nguyên và một số nguyên  $k$  với  $1 \leq k < n$  và  $(n \leq 10^6)$ . Hãy chuyển  $k$  phần tử đầu dãy về cuối dãy (yêu cầu không dùng mảng trung gian)

Ví dụ:

$n=9; k=3$

3 6 5 1 2 7 6 9 8

Kết quả:

1 2 7 6 9 8 3 6 5

#### 1.2.

- a. Thuật toán Mergesort
- b. Thuật toán Heap sort
- c. Thuật toán Quick sort

#### 1.3.

Bài toán tìm cặp điểm gần nhau nhất

Trong mặt phẳng tọa độ  $OXY$  cho  $n$  điểm; điểm thứ  $i$  có tọa độ  $(x_i, y_i)$  với  $i=1..n$ .

Hãy trình bày thuật toán tìm một cặp điểm gần nhau nhất và cho biết độ phức tạp của thuật toán đó.

Hãy minh họa thuật toán đề xuất qua ví dụ sau:

$n=10$

(3;5), (2;2), (0;-5), (1; -1), (1;1), (-1;1), (-2;3), (-1; -2), (3,-4), (3;0)

### Thuật toán quy hoạch động

#### 1.4.

(Dãy con tăng dài nhất)

Cho dãy số nguyên  $A = a_1, a_2, \dots, a_n$  ( $n \leq 1000, -10000 \leq a_i \leq 10000$ ).

Một dãy con của  $A$  là một cách chọn ra trong  $A$  một số phần tử giữ nguyên thứ tự.

Yêu cầu: Tìm dãy con không giảm của  $A$  có độ dài dài nhất (lưu ý: Có thể có nhiều dãy con kết quả; và ta chỉ yêu cầu tìm 01 dãy con thỏa mãn).

\*Hãy đánh giá độ phức tạp thời gian tính của thuật toán.

#### 1.5.

(ba lô 1)

Cho  $n$  món hàng, món thứ  $i$  có khối lượng là  $a[i]$  ( $i=1..n, a[i]$  là số nguyên) và một ba lô có khối lượng  $M$ .

Yêu cầu: Chọn những món hàng nào vào một ba lô sao cho tổng khối lượng của các món hàng được chọn là lớn nhất nhưng không vượt quá khối lượng ba lô ?

(mỗi món hàng chỉ được chọn hoặc không được chọn).

\*Hãy đánh giá độ phức tạp thời gian tính của thuật toán.

INPUT:

5 12

1 6 8 7 1

OUTPUT

10

1 3 5

### 1.6.

(Ba lô 2 - bài ba lô tổng quát)

Cho  $n$  món hàng, món thứ  $i$  có khối lượng  $a[i]$  và có giá trị là  $c[i]$  (trong đó  $a_i, c_i \leq 100$  và là các số nguyên) và một ba lô khối lượng  $M$ .

Yêu cầu: Chọn những món hàng nào để bỏ vào ba lô sao cho khối lượng các món hàng được chọn không vượt quá  $M$  ( $M$  là số nguyên dương  $\leq 10000$ ) và khi đó ba lô có giá trị lớn nhất.

(Lưu ý mỗi món hàng chỉ có thể chọn tối đa một lần).

\*Hãy đánh giá độ phức tạp thời gian tính của thuật toán.

INPUT:

6 12

5 7

1 5

6 3

4 6

9 14

5 8

OUTPUT:

20

1 2 6

Áp dụng:

Áp dụng thuật toán quy hoạch động tạo bảng dữ liệu  $f[k][v]$ , với  $k=1..n$ ,  $v=1..M$ ,  $n=6$ ,  $M=12$ ; mỗi cặp số  $(a_i, c_i)$  với  $i=1..6$  lần lượt là (6;8), (10;12), (5;7), (8;4), (4;3), (9;3); hãy cho biết thứ tự của các món hàng được chọn ứng với bảng dữ liệu trên.

### 1.7.

(Di chuyển trên bảng)

Cho một bảng  $A$  kích thước  $m \times n$  ô (bắt đầu từ 1); mỗi ô chứa một số nguyên.

Từ ô  $A[i, j]$  chỉ có thể di chuyển sang một trong 3 ô  $A[i, j+1]$ ,  $A[i-1, j+1]$  và  $A[i+1, j+1]$ .

Yêu cầu: Tìm vị trí ô xuất phát từ cột 1 sang cột  $n$  sao cho tổng các số ghi trên đường đi là lớn nhất. Yêu cầu xuất giá trị tổng này.

\*Hãy đánh giá độ phức tạp thời gian tính của thuật toán.

### 1.8.

(Tam giác số)

- Cho một tam giác chứa các số nguyên không âm từ 0 đến 99. Dòng thứ  $h$  của tam giác chứa  $h$  số ( $1 < h \leq 1000$ ).

- Một đường đi là hợp lệ nếu đường đi bắt đầu từ ô đỉnh của tam giác xuống đến một ô nào đó ở đáy của tam giác; và tại mỗi ô ở dòng  $i$  ta chỉ có thể đi tiếp đến ô ở dòng  $i+1$  theo cách: Hoặc sang bên trái hoặc sang bên phải.

- Yêu cầu: Tìm một đường đi có tổng số các ô trên đường đi là lớn nhất; cho biết tổng số các ô trên đường đi tìm được.

- Hãy đánh giá độ phức tạp thời gian tính của thuật toán.

### 1.9.

(Dãy con chung dài nhất)

Cho hai số nguyên dương  $M, N$  ( $0 < M, N \leq 100$ ) và hai dãy số nguyên:  $A_1, A_2, \dots, A_M$  và  $B_1, B_2, \dots, B_N$ .

Yêu cầu: Tìm một dãy  $C$  là con chung dài nhất của hai dãy  $A$  và  $B$ , nhận được từ  $A$  bằng cách xoá đi một số số hạng và cũng nhận được từ  $B$  bằng cách xoá đi một số số hạng.

\*Hãy đánh giá độ phức tạp thời gian tính của thuật toán.

M=13;N=12

3 12 5 6 1 7 61 9 0 10 11 32 4

12 1 4 61 12 0 14 11 10 3 4 32

Kết quả:

12 1 61 0 11 4

## Thuật toán quay lui

### 1.10.

Xét bàn cờ tổng quát  $n \times n$ . Một quân Hậu trên bàn cờ có thể ăn các quân khác nằm tại các ô cùng hàng, cùng cột hoặc cùng đường chéo. Hãy tìm cách xếp  $n$  quân hậu trên bàn cờ sao cho không quân nào ăn quân nào.

### 1.11.

Bài toán Mã đi tuần

Mã đi tuần hay hành trình của quân Mã là bài toán về việc di chuyển một quân Mã trên bàn cờ Vua ( $8 \times 8$ ). Quân Mã được đặt ở một ô trên một bàn cờ trống nó phải di chuyển theo quy tắc của cờ Vua để đi qua mỗi ô trên bàn cờ đúng một lần. Hãy tìm một hành trình cho cho quân Mã (có rất nhiều lời giải cho bài toán này, trong đó quân Mã có thể kết thúc tại chính ô mà nó khởi đầu; một hành trình như vậy được gọi là hành trình đóng. Có những hành trình, trong đó quân Mã sau khi đi hết tất cả 64 ô của bàn cờ (kể cả ô xuất phát), thì từ ô cuối của hành trình không thể đi về ô xuất phát chỉ bằng một nước đi; những hành trình như vậy được gọi là hành trình mở).

### 1.12.

Bài toán người đi du lịch

Cho  $n$  thành phố (được đánh số từ 1 đến  $n$ ), biết ma trận chi phí đi lại giữa các thành phố là  $C_{ij}$ . Người du lịch xuất phát từ một thành phố bất kỳ, muốn đi qua tất cả các thành phố, mỗi thành phố đúng một lần rồi quay về thành phố xuất phát. Cách đi như vậy gọi là một hành trình. Chi phí của hành trình được tính như là tổng các chi phí của các đoạn đường của nó. Cần tìm hành trình với chi phí nhỏ nhất

### 1.13.

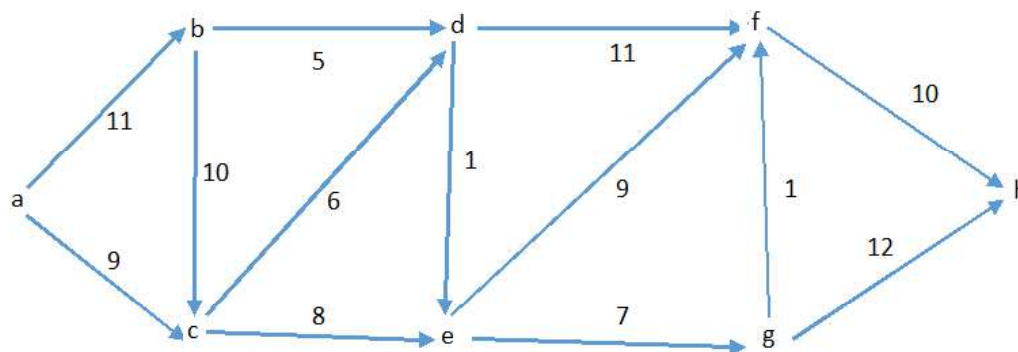
Bài toán tìm đường đi trong mê cung

Giả sử mê cung là một ma trận 0 1 trong đó 0 là đường đi, 1 là tường. Hãy tìm một đường đi từ ô vị trí  $(u,v)$  đến ô  $(z,t)$  hoặc chỉ ra không tồn tại đường đi.

## Chương 2.

### 2.1.

Cho đồ thị  $G=(V,E)$  như hình vẽ sau; trong đó các đỉnh được duyệt theo thứ tự  $a,b,c,d,e,f,g,h$



a. Trình bày thuật toán Dijkstra.

Hãy minh họa quá trình tìm đường đi ngắn nhất từ đỉnh  $a$  đến các đỉnh còn lại của đồ thị  $G$  theo thuật toán Dijkstra; từ đó hãy chỉ ra một đường đi ngắn nhất từ đỉnh  $a$  đến đỉnh  $h$  và cho biết khoảng cách ngắn nhất từ đỉnh  $a$  đến tất cả các đỉnh của đồ thị.

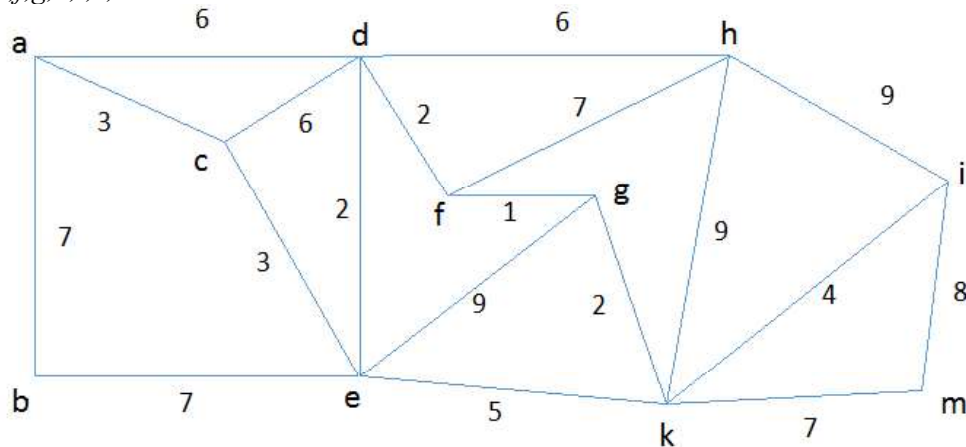
- b. Cho mạng  $G$  trong đó  $a$  là đỉnh phát,  $h$  là đỉnh thu, trọng số trên mỗi cung là khả năng thông qua của cung đó, giá trị luồng trên mỗi cung ban đầu được cho bằng 0. Hãy minh họa quá trình thực hiện thuật toán *Ford-Fulkerson* để tìm luồng cực đại trong mạng  $G$ .

- c. Trình bày thuật toán Kruskal.

Gọi  $G_1$  là đồ thị thu được từ đồ thị  $G$  bằng cách bỏ hướng trên các cung. Áp dụng thuật toán Kruskal mô tả quá trình tìm cây khung nhỏ nhất của đồ thị  $G_1$ . Vẽ cây khung nhỏ nhất tìm được; chiều dài cây khung nhỏ nhất này là bao nhiêu ?

## 2.2.

Cho đồ thị vô hướng có trọng số  $G=(V,E)$  như sau; trong đó các đỉnh được duyệt theo thứ tự  $a,b,c,d,e,f,g,h,i,k,m$ .



- a. Hãy đưa ra thứ tự duyệt các đỉnh của đồ thị theo thuật toán duyệt theo chiều rộng (BFS) bắt đầu từ đỉnh  $a$ .

Hãy đưa ra thứ tự duyệt các đỉnh của đồ thị theo thuật toán duyệt theo chiều sâu (DFS) bắt đầu từ đỉnh  $a$ .

- b. Hãy trình bày thuật toán Prim tìm cây khung nhỏ nhất của đồ thị.

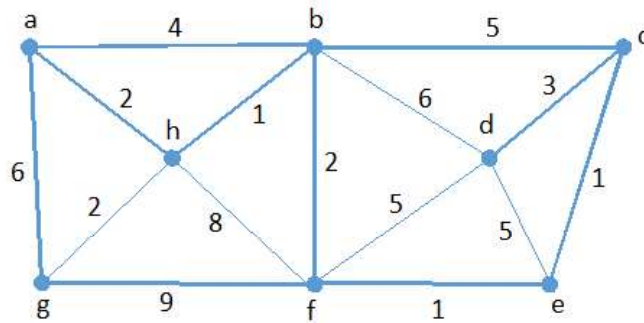
Áp dụng thuật toán Prim tìm cây khung nhỏ nhất của đồ thị  $G$  với đỉnh lựa chọn đầu tiên là  $c$ . Giá trị cây khung nhỏ nhất này là bao nhiêu ?

- c. Hãy trình bày thuật toán Dijkstra.

Hãy minh họa quá trình tìm đường đi ngắn nhất từ đỉnh  $a$  đến các đỉnh còn lại của đồ thị  $G$  theo thuật toán Dijkstra; từ đó hãy chỉ ra một đường đi ngắn nhất từ đỉnh  $a$  đến đỉnh  $m$  và cho biết khoảng cách ngắn nhất từ đỉnh  $a$  đến tất cả các đỉnh của đồ thị.

## 2.3.

Cho đồ thị vô hướng có trọng số  $G=(V,E)$  như hình vẽ sau; trong đó các đỉnh được duyệt theo thứ tự  $a,b,c,d,e,f,g,h$ .



- Hãy duyệt các đỉnh của đồ thị  $G$  theo các cách BFS, DFS bắt đầu từ đỉnh  $a$ .
- Trình bày thuật toán Kruskal.  
Minh họa quá trình tìm cây khung nhỏ nhất của đồ thị  $G$  theo thuật toán Kruskal. Vẽ cây khung nhỏ nhất tìm được, giá trị cây khung nhỏ nhất này là bao nhiêu?
- Trình bày thuật toán Dijkstra.  
Minh họa quá trình (lập bảng chi tiết) tìm đường đi ngắn nhất từ đỉnh  $a$  đến các đỉnh còn lại của đồ thị  $G$  (hình trên) theo thuật toán Dijkstra; từ đó chỉ ra một đường đi ngắn nhất từ đỉnh  $a$  đến đỉnh  $c$  và cho biết độ dài của đường đi này; đồng thời hãy cho biết khoảng cách ngắn nhất từ đỉnh  $a$  đến tất cả các đỉnh khác.

#### 2.4.

Cho đồ thị hai phía  $G = (X \cup Y, E)$ ; trong đó  $X = \{1, 2, 3, 4, 5, 6, 7, 8\}$ ,  $Y = \{a, b, c, d, e, f, g, h\}$ , danh sách 18 cạnh của tập  $E$  được cho như sau:

1 $d$	4 $a$	6 $h$
1 $e$	4 $c$	7 $b$
2 $a$	4 $d$	7 $c$
2 $f$	5 $b$	8 $e$
3 $f$	5 $h$	8 $f$
3 $g$	6 $f$	8 $g$

Hãy minh họa quá trình thực hiện thuật toán trên để tìm cặp ghép có nhiều cạnh được ghép nhất của đồ thị  $G$

#### 2.5.

Áp dụng thuật toán Hungari tìm một cặp ghép có trọng số lớn nhất của đồ thị hai phía được biểu diễn bằng ma trận trọng số sau:

6
6 9 6 1 4 9
3 2 1 7 8 6
5 7 3 4 1 9
8 8 3 7 6 6
1 4 1 8 9 7
7 2 1 4 3 5

#### 2.6.

Nghiên cứu thuật toán cặp ghép lớn nhất trên đồ thị tổng quát

#### 2.7.

Nghiên cứu thuật toán cặp ghép có trọng số lớn nhất trên đồ thị tổng quát

## Chương 3.

### 3.1.

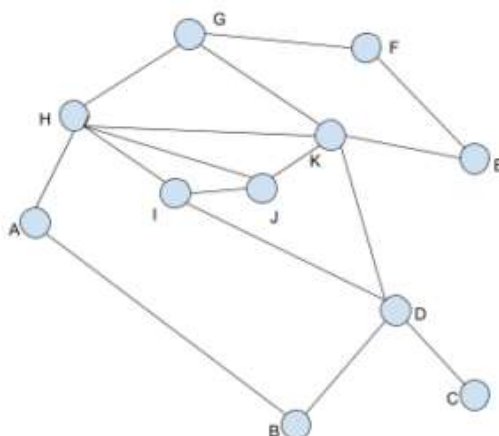
- Hãy trình bày sơ đồ thuật toán tham lam (greedy algorithm).
- Trình bày thuật toán tham lam giải bài toán TSP
- Áp dụng: Cho bài toán người bán hàng (Travelling Salesman Problem) có ma trận chi phí như sau:

$\infty$	18	21	14	10	19
16	$\infty$	20	10	27	23
17	9	$\infty$	22	18	13
16	13	18	$\infty$	11	19
18	14	15	19	$\infty$	8
14	19	20	11	21	$\infty$

Hãy sử dụng thuật toán tham lam GTS2 (Greedy Travelling Salesman) để tìm hành trình tốt nhất với  $p=4$ :  $v_1=1$ ;  $v_2=3$ ;  $v_3=4$ ;  $v_4=6$ .

### 3.2.

- Trình bày thuật toán tham lam Welsh - Powell giải bài toán tô màu đồ thị
- Áp dụng: Tô màu đồ thị sau



### 3.3.

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 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.

### 3.4.

Một hội thảo khoa học được tổ chức với  $n$  chủ đề khác nhau được ký hiệu là:  $c_1, c_2, \dots, c_n$ . Mỗi chủ đề hội thảo cần được tổ chức trong một buổi; trong đó có  $m$  cặp chủ đề/hoặc nhóm chủ đề không được diễn ra đồng thời trong cùng một buổi.

- Hãy trình bày một thuật toán tham lam để lập lịch các chủ đề trên vào các buổi sao cho số buổi hội thảo diễn ra là ít nhất có thể.
- Minh họa thuật toán trên qua ví dụ sau:  $n=10$ ; các chủ đề của hội thảo ký hiệu là 1,2,3,4,5,6,7,8,9,10; có  $m=21$  cặp chủ đề sau đây không được diễn ra đồng thời (1,2); (1,3); (1,4); (1,5); (1,7); (1,8); (2,3); (2,5); (2,6); (2,10); (3,4); (3,6); (3,10); (4,5); (4,6); (4,7); (5,6); (5,8); (5,9); (7,8); (8,9).

### 3.5.

Một giải vô địch cờ vua gồm 7 kỳ thủ. Các kỳ thủ thi đấu vòng tròn để tính điểm. Biết rằng Kỳ thủ 1 đã thi đấu với kỳ thủ 3, 4, 7; Kỳ thủ 2 đã thi đấu với kỳ thủ 4; Kỳ thủ 3 đã thi đấu với kỳ thủ 4; Kỳ thủ 4 đã thi đấu với kỳ thủ 7; Kỳ thủ 5 đã thi đấu với kỳ thủ 6,7. Trong một buổi thì 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.

### 3.6.

Một dịch vụ đánh máy luận văn tốt nghiệp có 3 nhân viên và 1 quản lý làm việc. Dịch vụ nhận được yêu cầu đánh máy một số luận văn với thông tin về số trang của mỗi luận văn như bảng sau:

Luận văn	$L_1$	$L_2$	$L_3$	$L_4$	$L_5$	$L_6$	$L_7$	$L_8$	$L_9$	$L_{10}$	$L_{11}$	$L_{12}$
Số trang	150	270	120	210	90	240	270	210	90	120	180	60

Giả sử trong mỗi giờ, mỗi nhân viên đánh máy được 10 trang, giả thiết các nhân viên có công suất đánh máy như nhau.

- Hãy phân công các luận văn trên cho ba nhân viên đánh máy sao cho thời gian hoàn thành việc đánh máy các luận văn là sớm nhất.
- Trong trường hợp người quản lý cũng tham gia đánh máy, nhưng công suất của người quản lý chỉ bằng  $\frac{3}{4}$  công suất của nhân viên. Hãy tìm cách phân công các luận văn trên cho 3 nhân viên và người quản lý đánh máy, sao cho thời gian hoàn thành việc đánh máy các luận văn là sớm nhất.

### 3.7.

Một trung tâm máy tính hiệu năng cao cần lập lịch cho  $n$  khách hàng thuê sử dụng các dịch vụ. Các khách hàng được đánh số từ 1 đến  $n$ , với mỗi khách hàng thứ  $i$ , trung tâm biết được thông tin về thời điểm ngày bắt đầu và thời điểm ngày kết thúc  $(s_i, f_i)$  cần đăng ký thuê. Tại mỗi thời điểm (ngày) bất kỳ, trung tâm chỉ có thể cho duy nhất một khách hàng thuê. Hỏi trung tâm cần chọn những khách hàng nào để cho thuê thì số lượng khách hàng được chọn là nhiều nhất?

Giả sử thời điểm bắt đầu, kết thúc là thông tin về ngày được quy ước theo số thứ tự ngày trong năm; ngày 1/1 là ngày đầu tiên của năm (Ví dụ các khách hàng có nhu cầu đăng ký thuê là (1,4) và (4,13) thì không thể chọn đồng thời do thời điểm đăng ký của các khách hàng này là giao nhau).

- Hãy trình bày một thuật toán tham lam giải bài toán trên.
- Minh họa thuật toán trên qua ví dụ sau:  $n=14$ ; thông tin thời điểm đăng ký thuê của các khách hàng lần lượt là (27,30); (26,28); (11,18); (11,15); (28,30); (29,31); (22,26); (20,25); (6,9); (6,12); (4,8); (3,5), (17,30); (24,31).

### 3.8.

Cho  $n$  chi tiết  $D_1, D_2, \dots, D_n$  cần được lần lượt gia công trên hai máy  $A, B$  (một chi tiết sau khi gia công trên máy  $A$  thì được tiếp tục gia công trên máy  $B$ . Mỗi máy khi được cho thực hiện một chi tiết nào đó thì nó sẽ làm cho tới khi hoàn chỉnh phần việc của nó). Thời gian gia công chi tiết  $D_i$  trên máy  $A$  là  $a_i$ ; thời gian gia công chi tiết  $D_i$  trên máy  $B$  là  $b_i$  với  $i=1, 2, \dots, n$ .

- Hãy trình bày một thuật toán tìm lịch gia công các chi tiết trên hai máy  $A, B$  sao cho tất cả các chi tiết được hoàn thành là sớm nhất có thể.

- b. Áp dụng cho bài toán cụ thể sau:  $n=10$ ; các chi tiết cần gia công là  $D_1, D_2, \dots, D_{10}$   
Thời gian gia công các chi tiết  $D_i$  trên các máy  $A, B$  được cho như ở bảng sau:

Công việc Máy	$D_1$	$D_2$	$D_3$	$D_4$	$D_5$	$D_6$	$D_7$	$D_8$	$D_9$	$D_{10}$
$A$	5	5	5	12	5	6	9	11	8	12
$B$	7	8	4	10	3	9	8	14	10	7

Hãy tìm lịch gia công các chi tiết trên hai máy  $A, B$ .

### 3.9.

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 ngắn gọn một thuật toán tham lam giải bài toán trên.
- Mình họa 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; 24), (7; 15), (9; 23); khi đó hãy cho biết số công việc hoàn thành trễ hạn ứng với thuật toán trên.

### 3.10.

Hãy sử dụng thuật toán tham lam giải bài toán sau đây:

Cho  $n = 12$  đoạn thẳng sau; các đoạn thẳng có các điểm đầu mút nằm trên trục hoành.

$\{(18;21), (6;9), (2;4), (3;6), (13; 16), (4;5), (10;17), (16;20), (8;12), (7,15), (1;3), (6;8)\}$ .

Ví dụ đoạn (18;21): có tọa độ điểm đầu là (18;0), tọa độ điểm cuối là (21;0).

Hãy tìm tập chứa nhiều đoạn thẳng không giao nhau và có số đoạn được chọn là nhiều nhất có thể.

### 3.11.

Giải bài toán phân việc với  $n=4$

11 9 1 4  
8 13 7 6  
12 2 14 11  
10 13 15 2

Dòng  $i$  cột  $j$  cho biết hiệu quả khi phân công nhân viên  $i$  thực hiện công việc  $j$ . Hãy tìm một cách phân công sao cho tổng hiệu quả phân công là lớn nhất.

### 3.12.

Tìm hiểu thuật toán tham lam giải một trong các bài toán sau:

- Max clique
- Cây steiner nhỏ nhất (Steiner minimal trees problem)
- Bài toán định tuyến xe (vehicle routing problem - VRP)



## Chương 4.

### 4.1.

- Hãy trình bày các khái niệm thuật toán heuristic, thuật toán metaheuristic.
- Hãy trình bày sơ đồ thuật toán local search algorithm.
- Hãy trình bày sơ đồ thuật toán genetic algorithm.

### 4.2.

So sánh ưu điểm, khuyết điểm của thuật toán heuristic với các thuật toán tìm lời giải chính xác (chẳng hạn như thuật toán quy hoạch động, thuật toán quay lui, thuật toán nhánh cận) trong việc giải bài toán tối ưu.

### 4.3.

- Hãy trình bày lược đồ thuật toán tìm kiếm địa phương (local search algorithm).
- Cho bài toán người du lịch (travelling salesman problem - *TSP*) với  $n=9$  có ma trận chi phí như sau; trong đó mỗi ứng viên lời giải của bài toán là một dãy hoán vị gồm  $n$  phần tử của tập hợp  $\{1,2,3,4,5,6,7,8,9\}$  (chi phí của mỗi lời giải phải bao gồm chi phí đi từ thành phố cuối cùng trong hành trình quay trở về thành phố xuất phát).

$\infty$	42	38	10	42	20	22	26	27
46	$\infty$	25	10	40	38	31	17	35
44	44	$\infty$	14	13	13	36	8	13
20	23	44	$\infty$	35	16	31	40	14
27	47	33	28	$\infty$	30	30	33	19
36	8	16	27	16	$\infty$	47	31	41
38	49	12	45	14	32	$\infty$	26	10
27	33	35	29	42	19	33	$\infty$	24
15	33	31	41	4	15	15	13	$\infty$

Hãy tính chi phí của lời giải  $\{5\ 1\ 7\ 8\ 9\ 4\ 6\ 2\ 3\}$ .

- Hãy trình bày ý tưởng các chiến lược tìm kiếm một lời giải lân cận đối với *TSP*: chiến lược “đảo”, chiến lược “chèn”, chiến lược “dời chỗ”, chiến lược “trao đổi qua lại”.
- Áp dụng: (*TSP* có ma trận chi phí như trên)
  - Cho lời giải: 5 1 7 8 9 4 6 2 3; hỏi chiến lược “đảo” với 2 điểm (vị trí) được chọn là 3 và 7 thì lời giải sinh ra có chi phí là bao nhiêu ?
  - Cho lời giải: 5 1 7 8 9 4 6 2 3; hỏi chiến lược “chèn” với thành phố được chọn có vị trí là 5 và vị trí chèn vào là vị trí 2 (chèn thành phố ở vị trí 5 vào tại vị trí 2) thì lời giải sinh ra có chi phí là bao nhiêu ?
  - Cho lời giải: 5 1 7 8 9 4 6 2 3; hỏi chiến lược “dời chỗ” với hành trình con được chọn là từ vị trí 6 đến vị trí 9 và vị trí cần chèn hành trình vào là vị trí 3 thì lời giải sinh ra có chi phí là bao nhiêu ?
  - Cho lời giải: 5 1 7 8 9 4 6 2 3; hỏi chiến lược “trao đổi qua lại” với hai thành phố được chọn tại các vị trí 3 và vị trí 9 thì lời giải sinh ra có chi phí là bao nhiêu ?

### 4.4.

- Hãy trình bày lược đồ thuật toán di truyền (genetic algorithm).
- Hãy chọn trình bày 2 trong số các vấn đề sau đây của thuật toán di truyền: tạo lời giải ban đầu, điều kiện dừng, phép lai (crossover), phép đột biến (mutation), phép chọn lọc (selection).

- c. Hãy trình bày ý tưởng các phép lai  $PMX$ ,  $OX$ ,  $CX$  của thuật toán di truyền đối với  $TSP$ .
- d. Áp dụng: (xét trong ngữ cảnh sử dụng thuật toán di truyền giải  $TSP$ )
- d<sub>1</sub>. Cho hai cá thể  $p_1 = \{1\ 2\ 3\ / 4\ 5\ 6\ 7\ / 8\ 9\}$  và  $p_2 = \{4\ 5\ 1\ / 3\ 8\ 2\ 9\ / 7\ 6\}$ ; hãy tìm các cá thể con  $p_1', p_2'$  sinh ra từ phép lai  $PMX$ .
- d<sub>2</sub>. Cho hai cá thể  $p_1 = \{1\ 2\ 3\ / 4\ 5\ 8\ 7\ / 9\ 6\}$  và  $p_2 = \{4\ 5\ 1\ / 3\ 8\ 7\ 6\ / 2\ 9\}$ ; hãy tìm các cá thể con  $p_1', p_2'$  sinh ra từ phép lai  $OX$ .
- d<sub>3</sub>. Cho hai cá thể  $p_1 = \{1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\}$  và  $p_2 = \{4\ 1\ 2\ 8\ 7\ 6\ 9\ 3\ 5\}$ ; hãy tìm các cá thể con  $p_1', p_2'$  sinh ra từ phép lai  $CX$ .

#### 4.5.

Hãy trình bày lược đồ thuật toán tìm kiếm lân cận biến đổi

#### 4.6.

Hãy trình bày lược đồ thuật toán bees.

#### 4.7.

Xây dựng một thuật toán metaheuristic giải một trong các bài toán sau:

- Max clique
- Cây steiner nhỏ nhất (Steiner minimal trees problem)
- Bài toán định tuyến xe (vehicle routing problem - VRP)

### Chương 5.

#### 5.1.

- a. Cho các luật đúng sau đây:

$$\begin{aligned} a \wedge d &\rightarrow e; \\ b \wedge c &\rightarrow f; \\ e \wedge f &\rightarrow g; \\ a; \\ b; \\ a &\rightarrow c; \\ b &\rightarrow d; \\ c &\rightarrow e; \\ g &\rightarrow h; \end{aligned}$$

Hãy sử dụng cơ chế suy diễn tiến kiểm tra xem  $h$  có được suy ra từ các luật trên hay không ?

- b. Cho các luật sau đây là đúng:

$$\begin{aligned} t &\rightarrow u; \\ p \wedge h &\rightarrow v; \\ p &\rightarrow q; \\ q &\rightarrow r; \\ p \wedge q &\rightarrow s; \\ p \wedge r &\rightarrow h; \\ q \wedge s &\rightarrow t; \\ p; \end{aligned}$$

Hãy sử dụng cơ chế suy diễn tiến kiểm tra xem  $u \wedge v$  có được suy ra từ các luật trên hay không ?

- c. Cho 8 luật sau đây là đúng:

$$\begin{aligned} a; \\ b; \end{aligned}$$

$c$ ;  
 $a \wedge b \rightarrow d$ ;  
 $a \wedge c \rightarrow e$ ;  
 $b \wedge d \rightarrow f$ ;  
 $f \rightarrow g$ ;  
 $a \wedge e \rightarrow h$ ;

Hãy sử dụng cơ chế suy diễn tiến hoặc cơ chế suy diễn lùi kiểm tra xem  $h \wedge g$  có được suy ra từ các luật trên hay không ?

## 5.2.

- Hãy nêu các phương pháp biểu diễn tri thức.
- Hãy trình bày định nghĩa suy diễn tiến (*forward chaining*), suy diễn lùi (*backward chaining*).
- Giả sử tri thức của bài toán được phát biểu như sau:  
 $R_1$ : IF *tỉ lệ lãi suất giảm* THEN *giá trái phiếu sẽ tăng*  
 $R_2$ : IF *tỉ lệ lãi suất tăng* THEN *giá trái phiếu sẽ giảm*  
 $R_3$ : IF *tỉ lệ lãi suất không đổi* THEN *giá trái phiếu không đổi*  
 $R_4$ : IF *giá đô la lên cao* (so với các tiền tệ khác) THEN *tỉ lệ lãi suất giảm*  
 $R_5$ : IF *giá đô la giảm* THEN *tỉ lệ lãi suất tăng*  
 $R_6$ : IF *giá trái phiếu giảm* THEN *mua trái phiếu*  
 -Một khách hàng thấy rằng *tỉ lệ lãi suất không đổi*, cô ấy muốn biết có nên mua trái phiếu hay không ? Hãy thực hiện suy diễn tiến hoặc suy diễn lùi để cho cô ấy lời khuyên.  
 -Một khách hàng thấy rằng *giá đô la giảm*, anh ta muốn biết có nên mua trái phiếu hay không ? Hãy thực hiện suy diễn tiến hoặc suy diễn lùi để cho anh ấy lời khuyên.

## 5.3.

Giả sử tri thức của bài toán được phát biểu như sau:

$R_1$ : IF *có 10000 đô la và có bằng cao đẳng* THEN *nên là nhà đầu tư trong lĩnh vực bảo hiểm*.  
 $R_2$ : IF *có thu nhập hàng năm ít nhất là 40000 đô la và có bằng cao đẳng* THEN *nên đầu tư vào lĩnh vực cổ phiếu gia tăng*.  
 $R_3$ : IF *chưa đến 30 tuổi và đang đầu tư trong lĩnh vực bảo hiểm* THEN *nên đầu tư trong lĩnh vực cổ phiếu gia tăng*.  
 $R_4$ : IF *chưa đến 30 tuổi và lớn hơn 22 tuổi* THEN *đã có bằng cao đẳng*.  
 $R_5$ : IF *muốn đầu tư vào lĩnh vực cổ phiếu gia tăng* THEN *cổ phiếu ấy nên là của IBM*.  
 Có một nhà đầu tư 25 tuổi và có 10.000 đô la mong được khuyên nên hay không nên đầu tư vào cổ phiếu IBM (sự kiện được biết: Một nhà đầu tư 25 tuổi, có 10000 đô la).  
 Dựa vào cơ sở tri thức và các sự kiện được biết trên, hãy sử dụng suy diễn tiến hoặc suy diễn lùi để cho anh ấy lời khuyên.

## 5.4.

Cho trước một tập hợp các luật ( $R_1, R_2, R_3$ ), câu hỏi đặt ra là “Chúng ta nên mua nhà hay không ?”

$R_1$ : if (lạm phát thấp) then lãi suất là thấp else lãi suất là cao  
 $R_2$ : if (lãi suất cao) then giá nhà cao  
 $R_3$ : if (giá nhà cao) then không mua nhà else mua nhà.

- Thực hiện một suy diễn lùi với sự kiện được biết là “lạm phát cao”.
- Thực hiện một suy diễn tiến với sự kiện được biết là “lạm phát thấp”.

## 5.5.

- Trình bày định nghĩa bài toán phân lớp dữ liệu.
- Trình bày một thuật toán để phân lớp dữ liệu.

c. Cho cơ sở dữ liệu sau:

TT	Quang cảnh	Nhiệt độ	Độ ẩm	Gió	Quyết định
1	Mưa	Ấm	Cao	Nhẹ	Không
2	Nhiều mây	Nóng	Cao	Nhẹ	Đi
3	Nhiều mây	Ấm	Cao	Mạnh	Đi
4	Nắng	Ấm	Thấp	Nhẹ	Đi
5	Nắng	Lạnh	Thấp	Mạnh	Không
6	Mưa	Nóng	Cao	Mạnh	Không
7	Nắng	Ấm	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	Ấm	Thấp	Mạnh	Đi

Từ cơ sở dữ liệu trên hãy rút ra bộ luật cho sự quyết định đi/không đi theo thuật toán Quinlan.

### 5.6.

Cho cơ sở dữ liệu sau:

TT	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	?

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

### 5.7.

- Hãy trình bày các khái niệm: Trí tuệ nhân tạo, Học máy, Cây quyết định.
- Trình bày một thuật toán để phân lớp dữ liệu.
- Cho cơ sở dữ liệu sau:

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

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

Áp dụng bộ luật trên hãy cho biết quyết định (Mùa) của các mẫu x và y

### 5.8.

- a. Hãy áp dụng một trong ba thuật toán Quinlan, ILA (Inductive Learning Algorithm), ID3 tìm bộ luật cho sự quyết định (buys\_computer) cho bảng dữ liệu gồm 14 mẫu sau:

RID	age	income	student	credit_rating	buys_computer
1	youth	high	no	fair	no
2	youth	high	no	excellent	no
3	middle_aged	high	no	fair	yes
4	senior	medium	no	fair	yes
5	senior	low	yes	fair	yes
6	senior	low	yes	excellent	no
7	middle_aged	low	yes	excellent	yes
8	youth	medium	no	fair	no
9	youth	low	yes	fair	yes
10	senior	medium	yes	fair	yes
11	youth	medium	yes	excellent	yes
12	middle_aged	medium	no	excellent	yes
13	middle_aged	high	yes	fair	yes
14	senior	medium	no	excellent	no

- b. Áp dụng bộ luật trên cho biết quyết định của các mẫu x, y với  
 $x = \{\text{middle\_aged, medium, yes, fair}\};$   
 $y = \{\text{youth, high, yes, excellent}\}$

### 5.9.

- a. Hãy áp dụng một trong ba thuật toán Quinlan, ILA (Inductive Learning Algorithm), ID3 tìm bộ luật cho sự quyết định (Class) cho bảng dữ liệu sau:

No.	Attributes				Class
	Outlook	Temperature	Humidity	Windy	
1	sunny	hot	high	false	N
2	sunny	hot	high	true	N
3	overcast	hot	high	false	P
4	rain	mild	high	false	P
5	rain	cool	normal	false	P
6	rain	cool	normal	true	N
7	overcast	cool	normal	true	P
8	sunny	mild	high	false	N
9	sunny	cool	normal	false	P
10	rain	mild	normal	false	P
11	sunny	mild	normal	true	P
12	overcast	mild	high	true	P
13	overcast	hot	normal	false	P
14	rain	mild	high	true	N

- b. Áp dụng bộ luật trên cho biết quyết định của các mẫu  $x, y, z$  với  
 $x = \{\text{overcast, cool, high, false}\};$   
 $y = \{\text{sunny, hot, normal, true}\};$   
 $z = \{\text{rain, cool, high, true}\}$

#### 5.10.

- Trình bày định nghĩa bài toán phân cụm dữ liệu
- Trình bày thuật toán  $k$ -means giải bài toán phân cụm dữ liệu
- Áp dụng thuật toán  $k$ -means giải bài toán sau:

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

Sử dụng thuật toán  $k$ -means giải bài toán phân cụm dữ liệu trên theo độ đo khoảng cách euclid (kết quả cần các thông tin: tâm mỗi cụm và liệt kê các điểm nằm trong mỗi cụm đó).

#### 5.11.

Trong mặt phẳng tọa độ  $OXY$ , cho 9 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

+Trình bày chi tiết các bước;

+Kết quả cần các thông tin: Tâm mỗi cụm và liệt kê các điểm nằm trong mỗi cụm đó.

#### 5.12.

Áp dụng thuật toán  $k$ -means giải bài toán sau:

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

Sử dụng thuật toán  $k$ -means giải bài toán phân cụm dữ liệu trên theo độ đo khoảng cách manhattan (kết quả cần các thông tin: tâm mỗi cụm và liệt kê các điểm nằm trong mỗi cụm đó).

#### 5.13.

Trình bày khái niệm khai phá dữ liệu. Trình bày ngắn gọn các giai đoạn của quá trình khai phá dữ liệu.

#### 5.14.

Cho một cơ sở dữ liệu gồm các giao dịch sau:

$$o_1 = \{i_1, i_3, i_4\}$$

$$o_2 = \{i_1, i_3\}$$

$o_3 = \{i_3, i_5\}$   
 $o_4 = \{i_1, i_2, i_4, i_5\}$   
 $o_5 = \{i_2, i_4\}$   
 $o_6 = \{i_1, i_2, i_4, i_5\}$

- Hãy tìm độ hỗ trợ (support) và độ tin cậy (confidence) của các luật kết hợp:  $i_1 \rightarrow i_2$ ;  $i_1, i_4, i_5 \rightarrow i_2$ .
- Hãy sử dụng thuật toán Apriori tìm tất cả các tập phổ biến tối đại (frequent itemset) với  $\min\_sup = 30\%$ .
- Với  $\min\_conf = 45\%$ , hãy liệt kê tất cả các luật kết hợp mạnh (strong association rule) từ các tập phổ biến tìm được ở câu b.

### 5.15.

Cho tập mặt hàng  $I = \{i_1, i_2, i_3, i_4, i_5, i_6, i_7, i_8\}$  và tập giao tác  $O = \{o_1, o_2, o_3, o_4, o_5, o_6\}$

$o_1 = \{i_1, i_7, i_8\}$   
 $o_2 = \{i_1, i_2, i_6, i_7, i_8\}$   
 $o_3 = \{i_1, i_2, i_6, i_7\}$   
 $o_4 = \{i_1, i_7, i_8\}$   
 $o_5 = \{i_3, i_4, i_5, i_6, i_8\}$   
 $o_6 = \{i_1, i_4, i_5\}$

- Hãy tìm độ hỗ trợ (support) và độ tin cậy (confidence) của các luật kết hợp:  $i_7 \rightarrow i_8$ ;  $i_8 \rightarrow i_7$ ;  $i_1, i_2 \rightarrow i_7$ ;  $i_7 \rightarrow i_1, i_2$ .
- Tìm tất cả các tập phổ biến theo ngưỡng  $\minsup = 0.3$
- Tìm tất cả các luật kết hợp theo ngưỡng  $\minsup = 0.3$  và ngưỡng  $\min\_conf = 1$  được tạo từ các tập phổ biến của câu b.

## Chương 6.

### 6.1.

- Hãy trình bày định nghĩa hệ thống thông minh (intelligent systems), nêu tên gọi các loại hệ thống thông minh.
- Hãy trình bày sự khác biệt cơ bản giữa các phần mềm tin học thông thường và các hệ cơ sở tri thức?

### 6.2.

Hãy trình bày định nghĩa hệ thống thông minh (intelligent systems), nêu tên gọi các loại hệ thống thông minh. Hãy trình bày sự khác biệt cơ bản giữa các phần mềm tin học thông thường và các hệ cơ sở tri thức?

### 6.3.

Hãy trình bày khái niệm hệ chuyên gia (expert system), trình bày cấu trúc cơ bản của một hệ chuyên gia.

### 6.4.

Hãy trình bày hệ hỗ trợ ra quyết định (decision support system), nêu các đặc trưng và khả năng cơ bản của một hệ hỗ trợ ra quyết định.

### 6.5.

Hãy trình bày các khái niệm: kho dữ liệu (data warehouse), cơ sở dữ liệu thông minh.

### 6.6.

Hãy trình bày tác tử thông minh khái (intelligent agent), trình bày cấu trúc cơ bản của một tác tử thông minh, trình bày các đặc tính cơ bản của tác tử thông minh.

Hết