Module: **INT102**
Assignment 2

1. Assessment

The tasks contribute 10% to the overall assessment of INT102

2. Submission

Please complete the assessment tasks using Microsoft Word and submit it in PDF via Learning. Please including **your name and student ID** on the cover page and name your file as **name_student ID**.

3. Deadline

26-May- 2021, Wednesday, 24:00. Time zone in Beijing, China (GMT+8)

Question 7 is compulsory.

Name: Lincheng Shi (施林呈)
Student_ID: 1927978

Question 1

1. A: 5-1-3=1
   G: 5-1-2=2
   C: 5
   T: 5-1-1=3

A	G	C	T	-
1	2	5	3	5

2.

A	G	C	A	A	T	G	A	A
A	T	G	A	A				
	A	T	G	A	A			
				A	T	G	A	A

So the number of comparisons is 9.

Question 2

a	b	c	d	e	
0	∞	∞	∞	∞	a
0	4	∞	∞	∞	b
0	4	5	∞	∞	c
0	4	5	12	∞	d
0	4	5	12	6	
0	4	5	12	-2	e
0	0	5	12	-2	b
0	0	5	8	-2	d

So the shortest path length from a to a is 0,
the shortest path length from a to b is 0,
the shortest path length from a to c is 5,
the shortest path length from a to d is 8,
the shortest path length from a to e is -2.

Question 3

1.

	“”	G	A	G	T
“”	0	0	0	0	0
A	0	↑0	↖1	←1	←1
G	0	↖1	↑1	↖2	←2
C	0	↑1	↑1	↑2	↑2
C	0	↑1	↑1	↑2	↑2
C	0	↑1	↑1	↑2	↑2
T	0	↑1	↑1	↑2	↖3

So the length of the Longest Common Subsequence between sequences of GAGT and AGCCCT is 3.

2.

	“”	G	A	G	T

“”	0	0	0	0	0
A	0	↑0	↖1	←1	←1
G	0	↖1	↑1	↖2	←2
C	0	↑1	↑1	↑2	↑2
C	0	↑1	↑1	↑2	↑2
C	0	↑1	↑1	↑2	↑2
T	0	↑1	↑1	↑2	↖3

So one of the longest common subsquences of GAGT and AGCCCT is AGT.

Question 4

1.a.

Complete the following table using the formula

$F(i,j)=\max(F(i-1,j-1)+s(x_i,y_j), F(i-1,j)+d, F(i,j+1)+d)$

and set $F(0,0)=0, F(i,0)=i*d, F(0,j)=j*d$

		G	A	G	T
	0	←-1	←-2	←-3	←-4
A	↑-1	↖↑←-2	↖0	←-1	←-2
G	↑-2	↖0	↑←-1	↖1	←0
A	↑-3	↑-1	↖1	↑←0	↑←-1
C	↑-4	↑-2	↑0	↑←-1	↖↑←-2
C	↑-5	↑-1	↑-1	↑←-2	↖↑←-3
T	↑-6	↑-2	↑-2	↑←-3	↖-1

b.

An optimal global alignment of GAGT and AGACCT:

- GA - - GT

AGACC -T

2.a.

Complete the following table using the formula

$F(i,j)=\max(F(i-1,j-1)+s(x_i,y_j), F(i-1,j)-d, F(i,j+1)-d,0)$

and set $F(0,0)=0$

		G	A	G	T
	0	0	0	0	0
A	0	0	↖1	0	0
G	0	↖1	0	↖2	←1
A	0	0	↖2	↑1	0
C	0	0	0	0	0
C	0	0	0	0	0
T	0	0	0	0	↖1

b.

An optimal local alignment of GAGT and AGACCT are found:

A G

A G

Question 5

	a	b	c	d	e
a	0	4	5	2	1
b	4	0	4	3	1
c	5	4	0	1	8
d	2	3	1	0	6
e	1	1	8	6	0

To reduce the complexity, we can assume that, d is before e in the tour. The lower bound for each node can be computed by

Node 0: lb=$\lceil[(1+2)+(1+3)+(1+4)+(1+2)+(1+1)]/2\rceil$=9

Node 1: lb=$\lceil[(4+1)+(4+1)+(4+1)+(2+1)+(1+1)]/2\rceil$=10

Node 2: lb=$\lceil[(5+1)+(3+1)+(5+1)+(2+1)+(1+1)]/2\rceil$=11

Node 3: lb=$\lceil[(2+1)+(3+1)+(4+1)+(2+1)+(1+1)]/2\rceil$=9

Node 4: ignored as d is not before e.

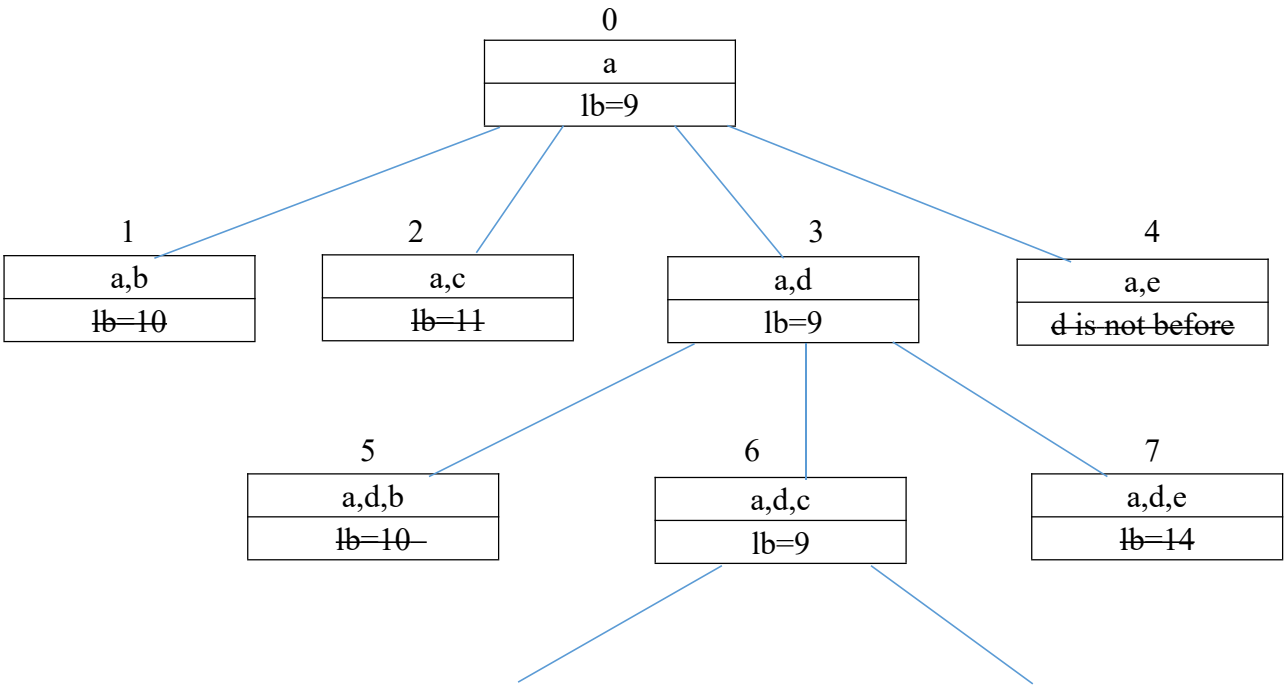
Node 5: lb=$\lceil[(2+1)+(3+1)+(4+1)+(2+3)+(1+1)]/2\rceil$=10

Node 6: lb=$\lceil[(2+1)+(3+1)+(4+1)+(2+1)+(1+1)]/2\rceil$=9

Node 7: lb=$\lceil[(2+1)+(3+1)+(4+1)+(6+2)+(6+1)]/2\rceil$=14

Node 8: lb=$\lceil[(2+1)+(4+1)+(4+1)+(2+1)+(1+1)]/2\rceil$=9

Node 9: lb=$\lceil[(4+2)+(4+1)+(1+8)+(2+1)+(1+8)]/2\rceil$=16



8	9
a,d,c,b,(e,a)	a,d,c,e,(b,a)
lb=9(optimal tour)	lb=16(inferior tour)

Solutions: a,d,c,b,e,a

Question 6

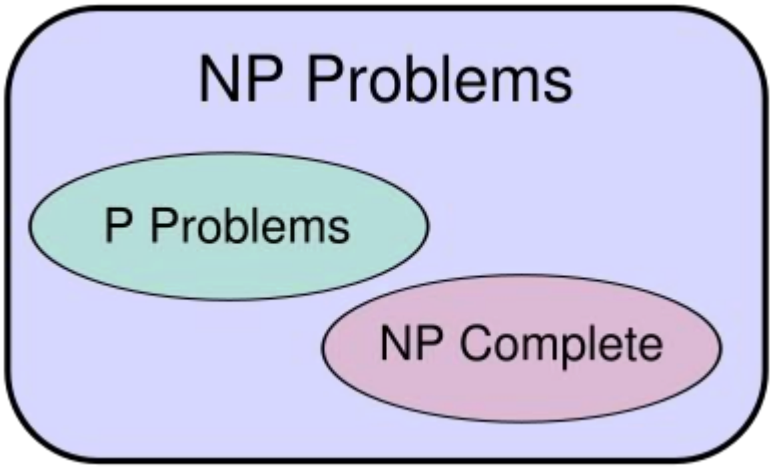
P problem is all decision problems that can be solved in worst-case polynomial time.
NP problem is all problems that can be verified in polynomial time.
All P problem can be verified in polynomial time, so P=NP or P⊂NP(P≠NP).

Problem A is NP-complete if

- Problem A is in NP
- For any Problem A' in NP, A' is reducible to A in polynomial time

So NPC⊂NP, P problem is not NPC problem.

As a result, we can get the diagram as shown below:



1. contradict

NP=NPC is wrong.

2. not contradict
3. contradict

NP=P∪NPC is wrong.

4. contradict

P∩NPC≠{} is wrong

5. not contradict

Question 7

Yes, I do.