

ĐỀ BÀI

Một mạng từ là một đồ thị có hướng với mỗi đỉnh là một từ, mỗi cạnh nối giữa hai từ w_1 và w_2 thể hiện khoảng cách từ w_1 tới w_2 (khác với khoảng cách từ w_2 đến w_1 nếu có). Mỗi từ chỉ xuất hiện duy nhất trong đồ thị một lần. Mạng từ được lưu trong hai tệp văn bản là 'vocab.txt' và 'graph.txt'. Tệp 'vocab.txt' có dòng đầu lưu tổng số từ khác nhau V (V là số nguyên dương < 1000), V dòng sau mỗi dòng có định dạng " w id" trong đó w là từ, id là chỉ số tương ứng của nó. Tệp 'graph.txt' dưới dạng danh sách cạnh trong đó dòng đầu là số cạnh m (m là số nguyên dương < 1000), m dòng tiếp theo mỗi dòng là một cạnh có định dạng " $distance$ id1 id2" với $distance$ là khoảng cách và id1 và id2 là chỉ số của hai từ tương ứng (id1 và id2 luôn khác nhau). Distance là số thực dạng double ($distance > 0$) có 5 số sau dấu phẩy.

Hãy viết một chương trình dưới dạng menu thực hiện các chức năng sau:

1. **Read:** Đọc tệp 'vocab.txt' và 'graph.txt' và lưu vào đồ thị có hướng.
2. **Adjacency:** Nhập vào hai từ w_1 và w_2 . Nếu w_2 liền kề với w_1 thì in ra màn hình khoảng cách tương ứng với 5 số sau dấu phẩy. Nếu không, in ra thông báo 'Not adjacent'.
3. **Path:** Nhập vào hai từ w_1 và w_2 . In ra màn hình độ dài đường đi ngắn nhất từ w_1 đến w_2 với 5 số sau dấu phẩy. Nếu không tồn tại đường đi thì in ra thông báo 'No path'.
4. **Summarize:** Một văn bản được lưu trong tệp văn bản 'input.txt' với dòng đầu là tổng số câu s trong văn bản (s là số nguyên dương, $20 < s < 100$). Trong s dòng tiếp theo lưu các câu trong văn bản, mỗi dòng bắt đầu bằng tổng số từ trong câu l (l là số nguyên dương, $3 < l < 50$) và tiếp theo là l từ cách nhau bởi dấu cách. Các từ là dạng chuỗi có độ dài tối đa 20 kí tự. Hãy in ra màn hình bản tóm tắt của văn bản là n câu có **điểm bé nhất** theo thứ tự tăng dần của điểm, mỗi câu trên một dòng. Nếu các câu có cùng điểm thì ưu tiên theo thứ tự xuất hiện sớm hơn trong văn bản. Trong đó n là một số nguyên dương nhập vào từ bàn phím ($n < 10$). Điểm của một câu được tính như sau: *Xét tất cả các cặp từ **liên tiếp nhau** trong câu. Với mỗi cặp từ đó, nếu tồn tại cạnh này trong mạng từ thì điểm của câu (score) được cộng thêm khoảng cách tương ứng trong đồ thị, số cặp (count) được cộng thêm 1. Nếu không tồn tại cạnh này trong mạng từ thì không cập nhật gì cả. Điểm của câu là khoảng cách trung bình của các cặp từ trong câu xuất hiện trong mạng từ (score/count).*
5. **Exit:** Thoát khỏi chương trình.

Lưu ý:

- Tất cả các từ trong mạng từ, nhập từ bàn phím, hay xuất hiện trong văn bản đều là dạng chuỗi có độ dài tối đa 20 kí tự.
- Nên dùng cấu trúc jrb để lưu đồ thị
- Các chức năng được tính điểm riêng biệt
- Phải giải phóng bộ nhớ trước khi thoát khỏi chương trình

Ví dụ một mạng từ với tệp 'vocab.txt':

61
[start] 0
aikido 1
is 2
a 3
japanese 4
martial 5
art 6
[end] 7
it 8
was 9
developed 10
by 11
morihei 12
ueshiba 13
based 14
on 15
philosophy 16
arts 17
training 18
and 19
religious 20
beliefs 21
the 22
word 23
often 24
translated 25
as 26
way 27
of 28
unifying 29
with 30
life 31
energy 32
or 33
harmonious 34
spirit 35
wanted 36
to 37
create 38
an 39
where 40
people 41
could 42
defend 43
themselves 44
without 45
harming 46
their 47
attacker 48
using 49
attackers 50
ki 51
against 52

them 53
he 54
each 55
practitioner 56
develop 57
both 58
physically 59
spiritually 60

và tệp 'graph.txt':

79
0.50000 0 1
0.33333 1 2
1.00000 2 3
1.00000 3 4
1.00000 4 5
1.00000 5 6
1.00000 6 7
1.00000 0 8
1.00000 8 9
1.00000 9 10
1.00000 10 11
1.00000 11 12
1.00000 12 13
1.00000 13 7
1.00000 2 14
1.00000 14 15
1.00000 15 13
1.00000 13 16
1.00000 16 5
1.00000 5 17
1.00000 17 18
1.00000 18 19
1.00000 19 20
1.00000 20 21
1.00000 21 7
1.00000 0 22
1.00000 22 23
1.00000 23 1
1.00000 2 24
1.00000 24 25
1.00000 25 26
0.50000 26 22
0.50000 22 27
0.50000 27 28
1.00000 28 29
1.00000 29 30
1.00000 30 31
1.00000 31 32
1.00000 32 33
1.00000 33 26
1.00000 28 34
1.00000 34 35
1.00000 35 7
1.00000 0 13

```

1.00000 13 36
1.00000 36 37
1.00000 37 38
1.00000 38 39
1.00000 39 6
1.00000 6 40
1.00000 40 41
1.00000 41 42
1.00000 42 43
1.00000 43 44
1.00000 44 45
1.00000 45 46
1.00000 46 47
1.00000 47 48
1.00000 48 11
1.00000 11 49
1.00000 49 22
1.00000 22 50
1.00000 50 51
1.00000 51 52
1.00000 52 53
1.00000 53 7
1.00000 0 54
1.00000 54 36
1.00000 36 55
1.00000 55 56
1.00000 56 28
1.00000 28 1
1.00000 1 37
1.00000 37 57
1.00000 57 58
1.00000 58 59
1.00000 59 19
1.00000 19 60
1.00000 60 7

```

Với đồ thị này ta có:

- (Adjacency) Trọng số của cạnh “was developed’ là 1.00000
- (Path) Khoảng cách ngắn nhất từ ‘aikido’ tới ‘the’ là 3.83333
- (Summarize) Với $n = 1$, câu có điểm bé nhất là [start] aikido is a japanese martial art [end]” với văn bản ‘input.txt’ dưới đây:

```

6
8 [start] aikido is a japanese martial art [end]
8 [start] it was developed by morihei ueshiba [end]
14 [start] aikido is based on ueshiba philosophy martial arts training and re-
ligious beliefs [end]
23 [start] the word aikido is often translated as the way of unifying with life
energy or as the way of harmonious spirit [end]
24 [start] ueshiba wanted to create an art where people could defend themselves
without harming their attacker by using the attackers ki against them [end]
14 [start] he wanted each practitioner of aikido to develop both physically and
spiritually [end]

```

Testcase:

Nội dung tệp 'vocab.txt':

86
[start] 0
aikido 1
is 2
a 3
japanese 4
martial 5
art 6
[end] 7
it 8
was 9
developed 10
by 11
morihei 12
ueshiba 13
based 14
on 15
philosophy 16
arts 17
training 18
and 19
religious 20
beliefs 21
the 22
word 23
often 24
translated 25
as 26
way 27
of 28
unifying 29
with 30
life 31
energy 32
or 33
harmonious 34
spirit 35
wanted 36
to 37
create 38
an 39
where 40
people 41
could 42
defend 43
themselves 44
without 45
harming 46
their 47
attacker 48
using 49

attackers 50
ki 51
against 52
them 53
he 54
each 55
practitioner 56
develop 57
both 58
physically 59
spiritually 60
performed 61
blending 62
moves 63
force 64
attack 65
rather 66
than 67
coming 68
this 69
achieved 70
techniques 71
are 72
completed 73
many 74
different 75
throws 76
joint 77
locks 78
which 79
can 80
be 81
combined 82
defenses 83
one 84
grappling 85

Nội dung tệp 'graph.txt'

126
0.25000 0 1
0.20000 1 2
1.00000 2 3
1.00000 3 4
1.00000 4 5
1.00000 5 6
1.00000 6 7
1.00000 0 8
1.00000 8 9
1.00000 9 10
1.00000 10 11
1.00000 11 12
1.00000 12 13
1.00000 13 7
1.00000 2 14
1.00000 14 15

1.00000 15 13
1.00000 13 16
1.00000 16 5
1.00000 5 17
1.00000 17 18
1.00000 18 19
1.00000 19 20
1.00000 20 21
1.00000 21 7
0.50000 0 22
1.00000 22 23
1.00000 23 1
1.00000 2 24
1.00000 24 25
1.00000 25 26
0.50000 26 22
0.33333 22 27
0.50000 27 28
1.00000 28 29
1.00000 29 30
1.00000 30 31
1.00000 31 32
1.00000 32 33
1.00000 33 26
1.00000 28 34
1.00000 34 35
1.00000 35 7
1.00000 0 13
1.00000 13 36
1.00000 36 37
1.00000 37 38
1.00000 38 39
1.00000 39 6
1.00000 6 40
1.00000 40 41
1.00000 41 42
1.00000 42 43
1.00000 43 44
1.00000 44 45
1.00000 45 46
1.00000 46 47
1.00000 47 48
1.00000 48 11
0.50000 11 49
0.33333 49 22
0.50000 22 50
0.50000 50 51
0.50000 51 52
0.50000 52 53
0.50000 53 7
1.00000 0 54
1.00000 54 36
1.00000 36 55
1.00000 55 56

1.00000 56 28
1.00000 28 1
1.00000 1 37
1.00000 37 57
1.00000 57 58
1.00000 58 59
1.00000 59 19
1.00000 19 60
1.00000 60 7
1.00000 2 61
1.00000 61 11
1.00000 11 62
1.00000 62 30
1.00000 30 22
1.00000 27 22
1.00000 22 48
1.00000 48 63
1.00000 63 49
1.00000 22 64
1.00000 64 28
1.00000 28 22
1.00000 22 65
1.00000 65 66
1.00000 66 67
1.00000 67 68
1.00000 68 52
1.00000 52 8
1.00000 8 7
1.00000 0 69
1.00000 69 2
1.00000 2 70
1.00000 70 11
1.00000 22 71
1.00000 71 72
1.00000 72 73
1.00000 73 30
1.00000 30 74
1.00000 74 75
1.00000 75 76
1.00000 76 33
1.00000 33 77
1.00000 77 78
1.00000 78 79
1.00000 79 80
1.00000 80 81
1.00000 81 82
1.00000 82 30
1.00000 30 75
1.00000 75 83
1.00000 83 7
1.00000 2 84
1.00000 84 28
1.00000 28 74
1.00000 74 85

1.00000 85 17

1.00000 17 7

Nội dung tệp ‘input.txt’:

10

8 [start] aikido is a japanese martial art [end]

8 [start] it was developed by morihei ueshiba [end]

14 [start] aikido is based on ueshiba philosophy martial arts training and religious beliefs [end]

23 [start] the word aikido is often translated as the way of unifying with life energy or as the way of harmonious spirit [end]

24 [start] ueshiba wanted to create an art where people could defend themselves without harming their attacker by using the attackers ki against them [end]

14 [start] he wanted each practitioner of aikido to develop both physically and spiritually [end]

24 [start] aikido is performed by blending with the way the attacker moves using the force of the attack rather than coming against it [end]

12 [start] this is achieved by using the attackers ki against them [end]

20 [start] the techniques are completed with many different throws or joint locks which can be combined with different defenses [end]

9 [start] aikido is one of many grappling arts [end]

No	Testcase	Input	Output	Score
1	Read			2.0
2	Adjacency	[start] aikido	0.25000	0.5
3		it on	Not adjacent	0.5
4	Path	[start] [end]	2.00000	0.5
5		developed [start]	No path	0.5
6	Summarize	1	[start] this is achieved by using the attackers ki against them [end]	1.0
7		2	[start] this is achieved by using the attackers ki against them [end] [start] aikido is a japanese martial art [end]	1.0
8		3	[start] this is achieved by using the attackers ki against them [end] [start] aikido is a japanese martial art [end] [start] the word aikido is often translated as the way of unifying with life energy or as the way of harmonious spirit [end]	1.0
9	Exit		Có giải phóng bộ nhớ	1.0
10	Menu			2.0
			Total	10