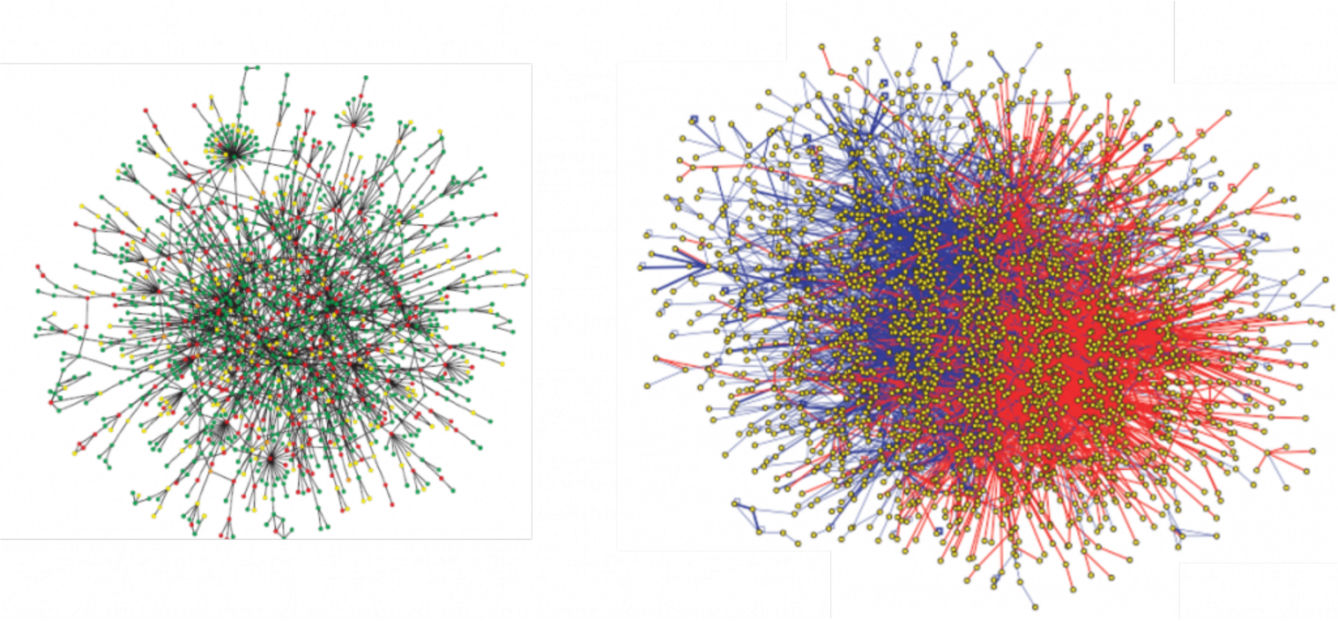
Problem 2: Not Another Pandemic [HackerRank]

By Nalin Ranjan and Sacheth Sathyanarayanan

Since the advent of the digital age, computers have been put to good use solving problems in the natural sciences — one of the greatest scientific achievements to date was the completion of the Human Genome Project in 2003, which was in a large part due to major advances in computing!

1 Background: Protein-Protein Interactomes

Protein-protein interactions are the fundamental building blocks of the more complex biological processes that life carries out. This problem is about protein-protein interactomes, which are a "mathematical" way of illustrating these interactions and their impact on resulting gene expression. The interactions are illustrated in a graph, where the vertices (nodes) represent either proteins or genes, and there are edges between proteins that interact, as well as between proteins and genes that depend on each other. Protein-protein interactomes can get really complex really fast, which only further reinforces the need for efficient computational algorithms to analyze them.



Interactomes with thousands of nodes and millions of interactions. [Image Credit: Center for Cancer Systems Biology]

2 Problem Statement

Consider a protein interactome where all nodes except one represent proteins and the remaining node represents a gene that is expressed. As it turns out, the world you are in is weathering a tough pandemic caused by the novel COVID-21 virus, whose major symptom seems to be a fetid body odor.

Our diligent scientists have been able to identify the cause of this symptom as the inhibition of the single gene GG21. They also know that the COVID-21 virus binds to the protein SRC6 and targets the chains of protein-protein interactions that start from SRC6 and lead up to the expression of GG21. For any such chain, COVID-21 starts by sabotaging the first interaction with some probability of success. It then experiences a cooldown period where it lays dormant for the duration of the next interaction. It then awakens to sabotage the third interaction along the chain, lays dormant for the fourth interaction, and so on, sabotaging every other interaction with some probability of success independent of the outcomes of every other sabotage attempt. If any of the interactions along the chain are successfully sabotaged, then the chain as a whole fails. As an example, consider the following chain from SRC6:

$$\text{SRC6} \longrightarrow \text{Protein 1} \longrightarrow \text{Protein 2} \longrightarrow \text{Protein 3} \longrightarrow \text{Protein 4} \longrightarrow \text{GG21}$$

COVID-21 will only attempt to sabotage the SRC6 $\longrightarrow$ Protein 1 interaction and the Protein 2 $\longrightarrow$ Protein 3 interaction. (It doesn't sabotage the Protein 4 $\longrightarrow$ GG21 interaction because that isn't an interaction between two proteins.) If any of these two interactions are successfully sabotaged by COVID-21, then the entire chain fails.

Your task is to analyze a given interactome that includes SRC6 and GG21 and find the chain of interactions from SRC6 to GG21 that is most resilient to COVID-21. More precisely, you want to find the chain of interactions from SRC6 to GG21 that endures all of COVID-21's sabotages with the highest probability. Remember, whether or not a given sabotage succeeds is independent of the outcomes of other sabotage attempts, so the probability that a given chain endures all of the sabotages is just the product of the probabilities of every other interaction in the chain enduring COVID-21's sabotage (starting with the first one).

Input

The first line consists of an integer N, which is how many nodes are in the interactome, followed by a whitespace, followed by another integer M, which is the number of interactions in the interactome. The nodes are numbered from 0 to $N-1$. You can assume that node 0 represents SRC6, node $N-1$ represents GG21, and every other node represents some other protein. Every subsequent line represents an interaction and contains three space-separated numbers: the first node of the interaction, the second node of the interaction, and the probability that COVID-21 successfully sabatoges that interaction (in that order).

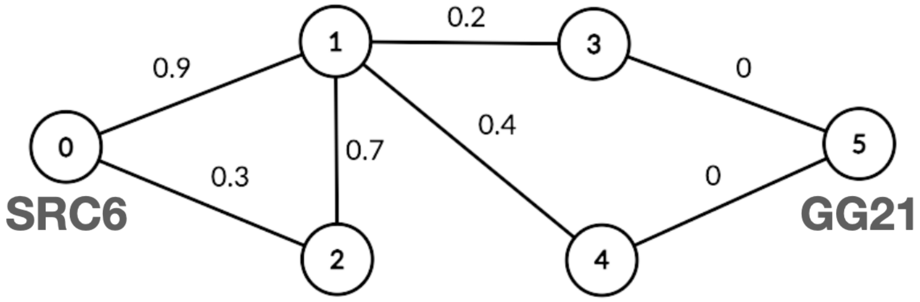
Output

Your output should consist of a single number, which is the *natural log* (i.e. logarithm with base $e \approx 2.718281828$) of the probability of the most resilient chain enduring all of COVID-21's sabotages. Since this will be a floating point number, please print your answer to seven digits after the decimal. Your answer will be accepted if it is within 10^{-5} of the correct answer.

Constraints

You can assume that $3 \leq N \leq 10^4$ and that the number of interactions is no more than $N^2/25$ if $N > 100$ (because interactomes usually aren't extremely dense). You may also assume that there will always exist at least one chain of interactions starting at SRC6 and ending at GG21, and that if an interaction is between some protein and GG21 (i.e. node $N-1$), then the probability of a successful sabotage is zero. Finally, you can assume that COVID-21's success probability in attempting to sabotage any given *protein-protein* interaction is between 0.0001 and 0.9999, and will always have no more than 7 digits of precision (in base 10).

Example 0



Input:

```
6 7
0 1 0.9
0 2 0.3
1 2 0.7
1 3 0.2
1 4 0.4
3 5 0
4 5 0
```

Output:

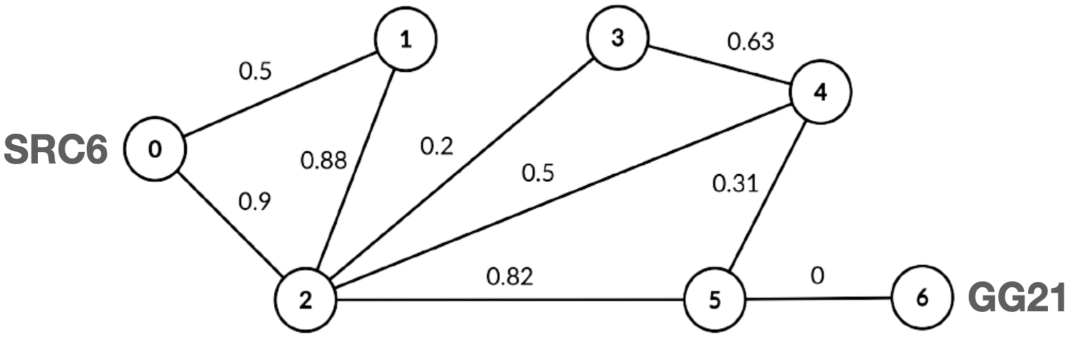
```
-0.5798185
```

Explanation: The most resilient chain from SRC6 to GG21 is

$$0 \longrightarrow 2 \longrightarrow 1 \longrightarrow 3 \longrightarrow 5$$

COVID-21 will only attempt to sabatoge the $0 \longrightarrow 2$ interaction and the $1 \longrightarrow 3$ interaction. The probability of the $0 \longrightarrow 2$ interaction enduring the sabotage is $1 - 0.3 = 0.7$, and the probability of the $1 \longrightarrow 3$ interaction enduring the sabotage is $1 - 0.2 = 0.8$. Thus the probability of this chain enduring all of COVID-21's sabotages is $0.7 \cdot 0.8 = 0.56$. The answer returned is $\ln(0.56)$ to seven decimal places.

Example 1



Input:

```
7 9
0 1 0.5
0 2 0.9
1 2 0.88
2 3 0.2
2 4 0.5
2 5 0.82
3 4 0.63
4 5 0.31
5 6 0
```

Output:

```
-1.2873544
```

Explanation: The most resilient chain from SRC6 to GG21 is

$$0 \longrightarrow 1 \longrightarrow 2 \longrightarrow 3 \longrightarrow 4 \longrightarrow 5 \longrightarrow 6$$

COVID-21 will attempt to sabatoge the $0 \longrightarrow 1$ interaction, the $2 \longrightarrow 3$ interaction, and the $4 \longrightarrow 5$ interaction. The probability of these individual interactions enduring a sabotage are $1 - 0.5 = 0.5$, $1 - 0.2 = 0.8$, and $1 - 0.31 = 0.69$, respectively. Thus the probability of this chain enduring all of COVID-21's sabotages is $0.5 \cdot 0.8 \cdot 0.69 = 0.276$. The answer returned is $\ln(0.276)$ to seven decimal places.