

Árvores Filogenéticas Herdadas

University of Ulm local Contest  Alemanha

Timelimit: 3

Entre outras coisas, Biologia Molecular Computacional lida com o processamento de seqüências genéticas. Considerando a relação evolutiva de duas seqüências, podemos dizer que eles estão intimamente relacionados, se eles não diferem muito. Podemos representar a relação por uma árvore, colocando seqüências de ancestrais acima de seqüências de seus descendentes. Tais árvores são chamadas árvores filogenéticas.

Considerando que uma tarefa da filogenia é inferir uma árvore a partir de seqüências de dados, vamos simplificar um pouco as coisas e proporcionar uma estrutura de árvore - esta será uma árvore binária completa. Você receberá as n folhas da árvore. Claro que você sabe, n é sempre uma potência de 2. Cada folha é uma seqüência de aminoácidos (designadas pelos códigos compostos de um caractere que você pode ver na figura). Todas as seqüências serão de igual comprimento l . Sua tarefa é derivar a seqüência de um ancestral comum com custos mínimos.

Aminoácido			Aminoácido		
Alanine	Ala	A	Leucine	Leu	L
Arginine	Arg	R	Lysine	Lys	K
Asparagine	Asn	N	Methionine	Met	M
Aspartic Acid	Asp	D	Phenylalanine	Phe	F
Cysteine	Cys	C	Proline	Pro	P
Glutamine	Gln	Q	Serine	Ser	S
Glutamic Acid	Glu	E	Threonine	Thr	T
Glycine	Gly	G	Tryptophan	Trp	W
Histidine	His	H	Tyrosine	Tyr	Y
Isoleucine	Ile	I	Valine	Val	V

Os custos são determinados como se segue: cada nó interior da árvore é marcado com uma seqüência de comprimento L . O custo de uma extremidade da árvore é o número de posições nas quais as duas seqüências diferem nas extremidades. O total custo é a soma dos custos em todas as bordas. A seqüência de um ancestral comum de todas as seqüências é então encontrado na raiz da árvore. Um antepassado comum ideal é um ancestral comum com os custos totais mínimos.

Entrada

O arquivo de entrada contém vários casos de teste. Cada caso de teste começa com dois números inteiros N e L , denota o número de seqüências para as folhas e do seu comprimento, respectivamente. A entrada é terminada por $N = L = 0$. Caso contrário, $1 \leq N \leq 1024$ e $1 \leq L \leq 1000$. Em seguida, siga N palavras de comprimento L sobre o alfabeto dos aminoácidos. Eles representam as folhas de uma árvore binária completa, a partir da esquerda para a direita.

Saída

Para cada caso de teste, imprima uma linha contendo algum ancestral comum ideal ótimo e os custos totais mínimos.

Exemplo de Entrada	Exemplo de Saída
4 3	AGA 3
AAG	AGA 4
AAA	AGA 4
GGA	R 2
AGA	W 0
	Y 1
	Q 0
4 3	
AAG	
AGA	
AAA	
GGA	
4 3	
AAG	
GGA	
AAA	
AGA	
4 1	
A	
R	
A	
R	
2 1	
W	
W	
2 1	
W	
Y	
1 1	
Q	
0 0	