

CPTS 350 HW 1

① Given: a number n and two primes p and q

Question: is it the case that $n = p \cdot q$?

language: $\{ \langle n \rangle : n = p \cdot q \}$ $\langle n \rangle$ is string encoding of number n .

Given: a number n

Question: is it the case that $n = p \cdot q$ for some primes p and q ?

language: $\{ \langle n \rangle : n = p \cdot q \}$ $\langle n \rangle$ is string encoding of number n .

Given: an NFA A and a word w

Question: Does A accept w ?

language: $\{ \langle A \rangle : A \text{ accepts } w \}$ $\langle A \rangle$ is string encoding of NFA, A .

Given: an NFA, A

Question: is there a word w such that A accepts w ?

language: $\{ \langle A \rangle : \text{there is a } w \text{ accepted by } A \}$ $\langle A \rangle$ is string encoding of NFA, A .

② Prove $2n^3 - 18n$ is $O(n^3)$ and $O(n^4)$ but not $O(n^2 \log n)$.

Proof: $2n^3 - 18n \leq 3 \cdot n^3$ for almost all n . Take $C=4$, the result follows.

Proof: $\limsup_{n \rightarrow \infty} \frac{2n^3 - 18n}{n^4} = \limsup_{n \rightarrow \infty} \frac{1}{n} < \infty$. Because $\limsup_{n \rightarrow \infty} \frac{f(n)}{g(n)} < \infty$,

the result follows.

Proof: Assume $2n^3 - 18n$ is indeed $O(n^2 \log n)$. Then, by definition there is a constant C and N_0 such that $2n^3 - 18n \leq C \cdot n^2 \log n$ for all $n \geq N_0$. $\frac{2n^3 - 18n}{n^2 \log n} \leq C$ for all $n \geq N_0$. Sending

$n \rightarrow \infty$, $\infty \leq C$. This is a contradiction and the result follows.

(2) Prove $3n^3 2^{2n}$ is $2^{O(n)}$

Proof: $3n^3 2^{2n} \leq 2^{O(n)}$ for almost all n . Take $C=3$,

~~the result follows~~. To show $3n^3 2^{2n} \leq 2^{O(n)}$ for almost all

n , $\limsup_{n \rightarrow \infty} \frac{3n^3 2^{2n}}{2^{O(n)}} = 0 < \infty$, therefore the result follows.

③ A. Comparison of protein structures is extremely important because the 3-D structure of a protein can most likely determine its functionality. Structurally similar molecules often have similar properties which can be highly effective in the discovery of new drugs.

A protein consists of a chain/chains of amino acid residues folded onto themselves. Ligands are molecules that connect to another molecule and perform a function.

Scientists can combat disease by finding a ligand that docks with proteins important for disease, disrupting normal function. This is important because once the structure of disease causing ligands is determined, it can be added to the database and compared to similar molecules. Thus it can be concluded that an algorithm which can accurately compare these protein molecules would be huge in determining which molecules are more likely to be diseased and can help tremendously in the development of new medical drugs.

Reference: Science direct, com/science/article/pii/S2001037014600774

B. I would ~~describe~~ represent the protein molecule with an array of jpeg pictures(2D) to fully represent the 3-D molecule. Then we could compare the indices of the arrays to check for similarity between the 2-D molecules and determine a score for the whole 3-D similarity.

C. Global Similarity - Scoring entire structures likeness by looking at the entire molecule to find molecules of a similar structure.

Pros - similar structures have similar functions.

- the folding structure plays a big part in determining function, and this method takes that into account.

Cons - tedious and hard to store the entire molecule.

- Just a similar structure doesn't always mean similar functions.

Local Similarity - Scoring similarity among structural pieces of the protein molecule in order to compare with similar motifs that have key functional properties.

Pros - micro particles of the molecule will likely have similar functions if they have a similar structure.

- very efficient.

Cons - doesn't take into account the full structure.

- can miss crucial folds when comparing similarity.

Global algorithm

- Store entire rough molecule shape in memory.
- Compare size, folds, number of atoms with online database to see similar functionality.
- Record results.

Local algorithm

- Scan molecule for key features.
- Store key features in memory.
- Compare features with online database with similar features.
- Record results.

Global algorithm is less efficient in finding small differences in the parts of the molecule that perform certain biological functions. But more efficient in comparing folds which can lead to similar functionality w/ similar fold structure.

Local algorithm is less efficient in looking at the entire structure but very good at finding small differences in certain parts of the protein molecule. This is a great way to look further into the differences in functions between parts of a protein molecule after the entire structure has been compared.