

Q1

Denote that the length of a snake's DNA is n , and a snake with venom level x must have DNA length of $5x$. So with a n length of DNA, the maximum venom level is $\lfloor n/5 \rfloor$. Which means the time complexity is $O(n/5) * O(n) = O(n^2)$ when using brute force. That is check venom level from 1 to $\lfloor n/5 \rfloor$, and each round we use greedy search runs in $O(n)$.

We can use binary search to check the venom level so that we can run in $O(n \log n)$.

First, we count the numbers n_s, n_n, n_a, n_k, n_e of occurrences of each letters, 'S', 'N', 'A', 'K', 'E' in the given DNA sequence and denote left boundary $l = 0$, right boundary $r = \min\{n_s, n_n, n_a, n_k, n_e\}$, which means we have a list contains range l to r stands for possible venom level. Apart from that, variable *max_level* to store the max venom level with initial value 0.

Then we do binary search, use while loop, let $mid = (l + r) // 2$ and use greedy algorithm to check whether the given DNA can derive level mid DNA sequence by deleting zero or more letters. If so, let $l = mid + 1$, *max_level* = mid , otherwise, let $r = mid - 1$, and continue the search in order to find the max level. Until the loop ends ($l = r$), and we can find the result.

For greedy method, since the format of venom DNA sequence is fixed, so for the given snake DNA we check from the start, is there have mid 's 'S' followed by mid 's 'N', mid 's 'A', mid 's 'K' and mid 's 'E', If it is, so this DNA can have level mid , otherwise not and do next loop.

This algorithm runs in $O(\log n)$ for binary search and $O(n)$ for greedy method each loop, so total time complexity is $O(n \log n)$.

Optimality: as I mentioned above, when use greedy method to traverse the given DNA sequence, first check whether exists mid 's 'S', if not we can break this loop and do next, that is the result must have 'S'* mid - 'N'* mid - 'A'* mid - 'K'* mid - 'E'* mid order. we cannot find any optimal solution with the random order of those letters, which means given any optimal solution we cannot find difference, so greedy solution is the optimal solution.