

Mid-term Break Problem Set (ICTPA)

[Ungraded]

1. A DNA sequence is a string consisting of A, G, T, C (capital letters). [These letters are also called *nucleotides*] Write a function `is_DNAseq` that takes an input string and returns `True` if it is a valid DNA sequence, and `False` otherwise.
2. The *complementary strand* (a.k.a. complement) of a DNA sequence is the reverse of the sequence such that every occurrence of A is replaced by T, G by C, T by A, and C by G. For example, the complement of “AGTACTGA” is “TCAGTACT”. See Figure 1. Write a function that takes a DNA sequence

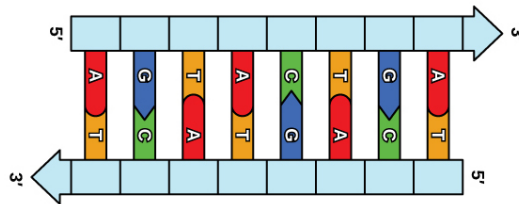


Figure 1: A depiction of complementary DNA strands

- (as a string) and returns its complement. [What do you do when the input is not a valid DNA sequence? Play around with the `assert` statement.]
3. A *period* of a string `s1` is a string `s2` (which is not necessarily different to `s1`) such that `s1 == k*s2` for some positive integer `k`, i.e., `s1` is `s2` copied `k` times. For example, the periods of “aaaa” are “aaaa”, “aa”, and “a”. In this case, the shortest period is “a”. Write a function `shortest_period` that takes an input string `s` and outputs the shortest period of `s`.
 4. ROT13 is a weak form of encryption that involves “rotating” each letter in a word by 13 places. To rotate a letter means to shift it through the alphabet, wrapping around to the beginning if necessary, so ‘A’ shifted by 3 is ‘D’ and ‘Z’ shifted by 1 is ‘A’. Write a function called `rotate_word` that takes a string and an integer as parameters, and that returns a new string that contains the letters from the original string *rotated* by the given amount. For example, “cheer” rotated by 7 is “jolly” and “melon” rotated by -10 is “cubed”. [You might want to use the built-in functions `ord`, which converts a character to a numeric code, and `chr`, which converts numeric codes to characters.]