Math 560 Extra Credit Fall 2021

Due: Friday, November 19

1. **Corrupted Text**

   You are given a string of n characters $S[1 \dots n]$, which you believe might be a corrupted text document in which all of the punctuation, spaces, and formatting (capital letters) has been removed. For example, the string could look something like:

   'duringthewholeofadulldarkandsoundlessdayintheautumnoftheyear...'

   You wish to reconstruct the original document using a dictionary or report that this is not possible. The dictionary is available in the form of a Boolean function `dict(w)` that will return `True` if `w` is a valid word and `False` otherwise, with each call to `dict(w)` taking $O(1)$ constant time. (Note: this constant time lookup can be accomplished using a hash table to store the dictionary.)

   (a) [1pt] Describe a **recursive** use-it or lose-it algorithm for determining if the input string S is a valid text document, i.e., is a sequence of valid words.

   (*Hint*: try asking whether a specific substring is in the dictionary. What two pieces of information are needed to specify a substring?)

   (b) [1pt] Describe, in clear English or pseudo-code, how you would convert your recursive algorithm into a Dynamic Programming solution. Indicate clearly how the DP table is set up and in what order the cells should be filled.

   (c) [1pt] Describe how you could construct the original document using your DP table.

   (d) [0.5pt] What is the runtime of your Dynamic Programming algorithm? Give a brief explanation.