School of Computing and Information Systems
The University of Melbourne
COMP90049 Knowledge Technologies (Semester 2, 2019)
Workshop exercises: Week 4

Suppose that we have observed the token `lended`, and we have a dictionary as follows:

```
addendum
blenders
commodity
deaden
end
leader
leant
lent
lemonade
pleading
```

1. Which, if any, of the above dictionary entries would be returned using a Neighborhood Search with a neighborhood of 2? 3?
2. With respect to the input string `lended` and the dictionary entry deaden, calculate the following:
   (a) the Global Edit Distance, using the parameter $[m, i, d, r] = [+1, -1, -1, -1]$
   (b) the Local Edit Distance, using the parameter $[m, i, d, r] = [+1, -1, -1, -1]$
   (c) the N-Gram Distance, using $n = 2$
   (d) the Jaro-Winkler Similarity, using $\ell <= 4$ and $p = 0.1$
3. Find the best approximate match (or matches, if there are ties) in the dictionary for the string `lended`, based on the following methods; consider different parameters where necessary:
   (a) the Global Edit Distance
   (b) the Local Edit Distance
   (c) the N-Gram Distance
   (d) the Jaro-Winkler Similarity
4. Assuming that the "correct" (intended) dictionary entry was `lent`, calculate the *precision* of each of the above methods of finding approximate entries from the dictionary.