School of Computing and Information Systems
The University of Melbourne
COMP90049
Knowledge Technologies (Semester 1, 2018)
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. 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 $n = 2$

2. 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) Soundex

3. Assuming that the "correct" (intended) dictionary entry was `lent`, calculate the precision of the following methods of finding approximate entries from the dictionary.

    (a) Neighbourhood search, with a neighbourhood of 1

    (b) Neighbourhood search, with a neighbourhood of 2

    (c) Neighbourhood search, with a neighbourhood of 3

    (d) Global Edit Distance, with a parameter $[m, i, d, r] = [1, -1, -1, -1]$

    (e) Local Edit Distance, with a parameter $[m, i, d, r] = [1, -1, -1, -1]$

    (f) N-gram Distance, where n is 2 (and padding with terminals)

    (g) Using the Soundex transformation, and then looking for exact matches

    (h) Using the Soundex transformation, and then permitting a 1-neighbourhood