CS221 Spring 2019 Homework 3

SUNet ID: yijiez
Name: Yijie Zhuang
Last updated: April 21, 2019

By turning in this assignment, I agree by the Stanford honor code and declare that all of this is my own work.

Problem 1: word segmentation

(a) Example: “Wewenttoamovieinstead.”
1-gram cost function: c(in)=0.01 c(instaed)=0.05 c(stead)=0.1
Sentence using greedy algorithm: “We went to a movie in stead”
Correct sentence: “We went to a movie instead”

Problem 2: Vowel Insertion

(a) Example: “why d w stck t bd hbts”
possibleFills(’t’) = [’to’, ’at’], possibleFills(’bd’) = [’bad’, ’bed’]
2-gram cost function: c(to, bed)=0.1, c(to, bad)=0.5
c(bed, habits) = 1, c(bad, habits) = 0.01
Sentence using greedy algorithm: “why do we stick to bed habits”
Correct sentence: “why do we stick to bad habits”

Problem 3: putting it together

(a) **States**: (1) Starting index of the new word (a). (2) Last chosen word by inserting space and vowels (w’).
Actions: Insert the space at index (b) and choose one word (w) in possibleFills(query[a: b]).
Cost: The bigram cost of the last chosen word (w’) and the current chosen word (w).
Initial state: 0, SENTENCE_BEGIN
End test: starting index == len(sentence)

(b) N/A

(c) To make sure $Cost_{rel}(s, a) <= Cost(s, a)$, we can have:

$$u_b(w) = \min_{w' \in allWordsInDictionary} b(w', w) \tag{1}$$

For the relaxed problem:
States: Starting index of the new word (a).

Actions: Insert the space at the index (b).
Cost: $\min_{w \in possibleFills(query[a:b])} u_b(w)$.
Initial state: 0
End test: starting index == len(sentence)

Proof of $Cost_{rel}(s, a) <= Cost(s, a)$:
$Cost_{rel}(s, a) = \min_{t \in possibleFills} u_b(t) \leq u_b(t')$
$= \min_{w' \in allWords} b(w', t) \leq b(lastChosenWord, t) = Cost(s, a)$

(d) UCS is a special case of A*. Because UCS explores states in order of PastCost(s) and A* explores states in order of PastCost(s) + h(s). We can say UCS is A* with h(s) = 0.

BFS is a special case of UCS. We can say BFS is UCS with cost of all edges equals 1.