

Universidade Federal do Ceará - Campus de Crateús
Bacharelado em Ciência da Computação
1º Trabalho Prático - Inteligência Artificial

Alunos
Fábio José dos Santos Sousa
Francisco Hartur Lopes de Alcântara

Resposta das questões 2.1, 3.1 e 4.1

2. Segmentação de Palavras

1. (2 pontos) Considere o seguinte algoritmo guloso:

- (a) inicie do início da string
- (b) encontre a posição final da próxima palavra que minimize a função de custo
- (c) repita o processo, iniciando do fim do segmento encontrado recentemente

Mostre que esse algoritmo guloso é subótimo. Em particular, forneça um exemplo de uma string em que essa abordagem falha em obter a segmentação da entrada com o menor custo.

O algoritmo ganancioso é problemático, pois ele não minimiza o custo futuro, dada uma segmentação, por exemplo: **'yourflightisdelayed'** e, em seguida, diga que o custo é o seguinte **{'you':5, 'your':6, 'rf': 10, 'flight': 4, 'light': 4, 'is': 4, 'delay': 3}**, Então o algoritmo seleciona **'you'** ao invés de **'your'**, a próxima palavra seria **'rf'**, pois qualquer outra palavra não começa com **'rf'**. Quando você termina a frase, torna-se **'you rf light is delayed'**, com o custo total de $5 + 10 + 4 + 4 + 3 = 25$ que é maior que o ideal $5 + 4 + 4 + 3 = 15$ (**'you flight is delay'**).

3. Inserção de Vogais

1. (2 pontos) Considere o seguinte algoritmo guloso: da esquerda para direita, escolha repetidamente a melhor inserção imediata da vogal para a palavra sem vogal atual, dada a inserção que foi escolhida para a palavra sem vogal anterior. Esse algoritmo não leva em consideração inserções futuras além da palavra atual. Encontre um contra-exemplo realista para mostrar que esse algoritmo guloso é subótimo.

Se selecionarmos a sentença **'th rt f fg'** e o seguinte bigram custos **{'the art': 5, 'the route': 4, 'art of': 2, 'route of': 10, 'of fugue': 2, 'of fog': 10}**. Quando o algoritmo seleciona **'the'** como primeira palavra e escolhe a próxima entre **'the art'** e **'the route'**, o menor custo imediato é **'the route'**. A única inserção subsequente possível é **'route of'**. Continuando indo para baixo, nós escolhemos **'of fugue'**. Então o custo total do bigram dessa sensação, **'the route of fugue'**, é

de $4 + 10 + 2 = 16$, mas se 'the art' for escolhida primeiro. 'the art of fugue' tem apenas o custo de $5 + 2 + 2 = 9$.

4. Problema Integrado

1. (2 pontos) Considere o problema de busca de encontrar inserções ótimas de espaços e vogais. Formalize-o como problema de busca. Quais estados, custos, estado inicial e teste de objetivo? Busque encontrar a representação mínima de estados.

Esse problema pode ser formalizado dessa maneira

Estado: (índice atual de caracteres digitalizados, palavra anterior)

Ação: (a possível palavra que será usada)

Custos: custo de bigram da palavra anterior com a palavra atual (ação)

Estado inicial: (o índice inicial de 0 e o token inicial)

Teste de objetivo: o índice de caracteres digitalizados é igual ao comprimento total da consulta.