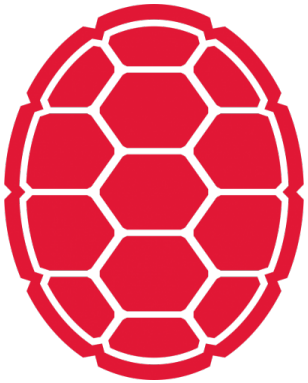

Language Models

Computational Linguistics: Jordan Boyd-Graber
University of Maryland
EXERCISE

Exercise

- Start with restaurant we had before
- Assume you see `<s> b b a c </s>`; add those counts to tables
- Compute probability of `b` following `a` ($\theta = 1.0, \delta = 0.5$)
- Compute the probability of `a` following `b`
- Compute probability of `</s>` following `<s>`

A busy night at the restaurant

Unigram Restaurant

a 3 b 1 c 1 </s> 1

<s> Restaurant

a 1

a Restaurant

a 2 b 1 c 1

b Restaurant

a 1

c Restaurant

</s> 1

A busy night at the restaurant

Unigram Restaurant

a 3 b 1 c 1 </s> 1

<s> Restaurant

a 1 b 1

a Restaurant

a 2 b 1 c 1

b Restaurant

a 1

c Restaurant

</s> 1

A busy night at the restaurant

Unigram Restaurant

a 3 b 2 c 1 </s> 1

<s> Restaurant

a 1 b 1

a Restaurant

a 2 b 1 c 1

b Restaurant

a 1

c Restaurant

</s> 1

A busy night at the restaurant

Unigram Restaurant

a 3 b 2 c 1 </s> 1

<s> Restaurant

a 1 b 1

a Restaurant

a 2 b 1 c 1

b Restaurant

a 1 b 1

c Restaurant

</s> 1

A busy night at the restaurant

Unigram Restaurant

a 3 b 3 c 1 </s> 1

<s> Restaurant

a 1 b 1

a Restaurant

a 2 b 1 c 1

b Restaurant

a 1 b 1

c Restaurant

</s> 1

A busy night at the restaurant

Unigram Restaurant

a 3 b 3 c 1 </s> 1

<s> Restaurant

a 1 b 1

a Restaurant

a 2 b 1 c 1

b Restaurant

a 2 b 1

c Restaurant

</s> 1

A busy night at the restaurant

Unigram Restaurant

a 3 b 3 c 1 </s> 1

<s> Restaurant

a 1 b 1

a Restaurant

a 2 b 1 c 2

b Restaurant

a 2 b 1

c Restaurant

</s> 1

A busy night at the restaurant

Unigram Restaurant

a 3 b 3 c 1 </s> 1

<s> Restaurant

a 1 b 1

a Restaurant

a 2 b 1 c 2

b Restaurant

a 2 b 1

c Restaurant

</s> 2

A busy night at the restaurant

Unigram Restaurant

a 3 b 3 c 1 </s> 1

<s> Restaurant

a 1 b 1

a Restaurant

a 2 b 1 c 2

b Restaurant

a 2 b 1

c Restaurant

</s> 2

As you see more data, bottom restaurants do more work.

b following a

$$= \frac{1-\delta}{\theta+5} + \frac{\theta+3\delta}{\theta+5} p(\mathrm{b}) \tag{1}$$

$$= \frac{1-\delta}{\theta+5} + \frac{\theta+3\delta}{\theta+5} \left(\frac{3-\delta}{\theta+8} + \frac{\theta+4\delta}{\theta+8} \frac{1}{V} \right) \tag{2}$$

(3)

b following a

$$= \frac{1-\delta}{\theta+5} + \frac{\theta+3\delta}{\theta+5} p(\text{b}) \tag{1}$$

$$= \frac{1-\delta}{\theta+5} + \frac{\theta+3\delta}{\theta+5} \left(\frac{3-\delta}{\theta+8} + \frac{\theta+4\delta}{\theta+8} \frac{1}{V} \right) \tag{2}$$

$$\tag{3}$$

b following a

$$= \frac{1-\delta}{\theta+5} + \frac{\theta+3\delta}{\theta+5} p(\text{b}) \tag{1}$$

$$= \frac{1-\delta}{\theta+5} + \frac{\theta+3\delta}{\theta+5} \left(\frac{3-\delta}{\theta+8} + \frac{\theta+4\delta}{\theta+8} \frac{1}{V} \right) \tag{2}$$

$$\tag{3}$$

0.23

a following b

$$= \frac{2-\delta}{\theta+3} + \frac{\theta+2\delta}{\theta+3} p(\mathrm{a}) \tag{4}$$

$$= \frac{2-\delta}{\theta+3} + \frac{\theta+2\delta}{\theta+3} \left(\frac{3-\delta}{\theta+8} + \frac{\theta+4\delta}{\theta+8} \frac{1}{V} \right) \tag{5}$$

(6)

a following b

$$= \frac{2-\delta}{\theta+3} + \frac{\theta+2\delta}{\theta+3} p(\text{a}) \tag{4}$$

$$= \frac{2-\delta}{\theta+3} + \frac{\theta+2\delta}{\theta+3} \left(\frac{3-\delta}{\theta+8} + \frac{\theta+4\delta}{\theta+8} \frac{1}{V} \right) \tag{5}$$

$$\tag{6}$$

a following b

$$= \frac{2-\delta}{\theta+3} + \frac{\theta+2\delta}{\theta+3} p(\text{a}) \tag{4}$$

$$= \frac{2-\delta}{\theta+3} + \frac{\theta+2\delta}{\theta+3} \left(\frac{3-\delta}{\theta+8} + \frac{\theta+4\delta}{\theta+8} \frac{1}{V} \right) \tag{5}$$

$$\tag{6}$$

0.55

</s> following <s>

$$= \frac{\theta + 2\delta}{\theta + 2} p(</\mathrm{s}>) \tag{7}$$

$$= \frac{\theta + 2\delta}{\theta + 2} \left(\frac{1 - \delta}{\theta + 8} + \frac{\theta + 4\delta}{\theta + 8} \frac{1}{V} \right) \tag{8}$$

(9)

</s> following <s>

$$= \frac{\theta + 2\delta}{\theta + 2} p(\text{</s>}) \tag{7}$$

$$= \frac{\theta + 2\delta}{\theta + 2} \left(\frac{1 - \delta}{\theta + 8} + \frac{\theta + 4\delta}{\theta + 8} \frac{1}{V} \right) \tag{8}$$

$$\tag{9}$$

</s> following <s>

$$= \frac{\theta + 2\delta}{\theta + 2} p(</\text{s}>) \tag{7}$$

$$= \frac{\theta + 2\delta}{\theta + 2} \left(\frac{1 - \delta}{\theta + 8} + \frac{\theta + 4\delta}{\theta + 8} \frac{1}{V} \right) \tag{8}$$

$$\tag{9}$$

0.08