$A \cap B$ : A かつ B.

$$P(B \mid A) = \frac{P(A, B)}{P(A)}$$

$$\Longleftrightarrow P(A, B) = P(A) \cdot P(B \mid A)$$

$$P(\text{"its water is so transparent"})$$

$$= P(\text{its})$$
$$\times P(\text{water} \mid \text{its})$$
$$\times P(\text{is} \mid \text{its water})$$
$$\times P(\text{so} \mid \text{its water is})$$
$$\times P(\text{transparent} \mid \text{its water is so})$$

$$P(w_1, w_2, \dots w_k) \approx \prod_i P(w_i \mid w_{i-k} \dots w_{i-1})$$

$$P(\text{the} \mid \text{its water is so transparent})$$

$$\approx P(\text{the} \mid \text{that})$$ Unigram

$$\approx P(\text{the} \mid \text{transparent that})$$ Bigram.

$$P(w_1, w_2, \ldots, w_k)$$

$$= P(w_1) \times P(w_2 \mid w_1)$$

$$\times P(w_3 \mid w_1, w_2) \ldots$$

$$\times P(w_k \mid w_{k-2}, w_{k-1})$$

$$\prod_i P(w_i \mid w_{i-2}, w_{i-1})$$

Trigram:

$$P(w_1, w_2, \cdots w_k)$$

$$= P(w_1) \times P(w_2 | w_1)$$

$$\times \prod_{i=3}^{k} P(w_k | w_{k-2}, w_{k-1})$$

You can continue to longer N-grams, but the combinations of words become more specific, so might be possible that N-grams not being present in the training data.

$$P(I \mid \langle s \rangle) = \frac{2}{3}$$

$$P(am \mid I) = \frac{C(I\ am)}{C(I)} = \frac{2}{3}$$

$$P(w_i \mid w_{i-1}) = \frac{C(w_{i-1}, w_i)}{C(w_{i-1})}$$

In practice,

these probabilities should be calculated under log to avoid. underflow

$$\log(p_1 \times p_2 \times p_3 \times p_4)$$

$$= \log p_1 + \log p_2 + \log p_3 + \log p_4$$

<s> I am Sam </s>

<s> Sam I am </s>

<s> I do not like green eggs and ham

Applying Trigram

$$P = P(w_1) \times P(w_2 | w_1) \times \prod_{i=2}^{k} P(w_k | w_{k-2,\, k-1})$$

$$P(<s>) = \frac{3}{3} = 1$$

$$P(I | <s>) = \frac{2}{3} = 0.67$$

$$P(am | <s>\ I) = \frac{1}{2} = 0.5$$

<s> I am Sam </s>

<s> Sam I am </s>

<s> I do not like green eggs and ham

$$P(\text{Sam} \mid \text{I am}) = \frac{1}{2} = 0.5$$

$$P(\langle /s \rangle \mid \text{am Sam}) = \frac{1}{1} = 1$$

$$P(\langle s \rangle \mid \text{Sam } \langle /s \rangle) = \frac{1}{1} = 1$$

$$P(w_3 \mid w_1, w_2) = \frac{C(w_1, w_2, w_3)}{C(w_1, w_2)}$$

$P$

$$\sqrt[120000]{\frac{1}{\left(\frac{1}{4}\right)^{90000} \times \left(\frac{1}{120000}\right)^{3000}}}$$

$$\sqrt[4]{\frac{1}{\frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} \times \frac{1}{120000}}}$$

$$\underbrace{0000\cdots0}_{91}\ \underbrace{123456789}_{9}$$

$$PP(W) = \sqrt[10]{\frac{1}{\left(\frac{1}{100}\right)^{1} \times \left(\frac{91}{100}\right)^{9}}}$$

$$\fallingdotseq 1.725$$ #

$$P(\text{i want chinese food})$$

$$= P(\text{i} \mid \text{<s>}) \times P(\text{want} \mid \text{i})$$

$$\times P(\text{chinese} \mid \text{want}) \times P(\text{food} \mid \text{chinese})$$

$$\times P(\text{</s>} \mid \text{food})$$

$$= 0.19 \times 0.33 \times 0.0065 \times 0.52 \times 0.40$$

$$= 0.19 \times 0.21 \times 0.0029 \times 0.052 \times 0.40$$

(2) The un-smoothed one is higher because the smoothed probability will provide some prob to the situation that will never happen, (probability = 0), so the probability will be more even in the smoothed probability.

$\boxed{3}$

$$P(\text{Sam} \mid \text{am})$$

$$= \frac{C(\text{am Sam}) + 1}{C(\text{am}) + V}$$

the number of vocab

$$= \frac{2+1}{3+V}$$

$$= \frac{2+1}{3+11} = \frac{3}{14} = 0.214$$

$$P(w_3 \mid w_1, w_2)$$

$$= \frac{C(w_1, w_2, w_3) + 1}{C(w_1, w_2) + V}$$

$$P(w_i \mid w_{i-1}) = \frac{c(w_{i-1}, w_i)}{c(w_{i-1})}$$

<s> I am Sam </s>
<s> Sam I am </s>
<s> I do not like green eggs and ham </s>

$$P(I \mid <s>) = \frac{c(I, <s>)}{c(<s>)} = \frac{2}{3}$$

$$= \frac{2}{3}$$

$$P(am \mid I) = \frac{c(am, I)}{c(I)} = \frac{2}{3} = 0.67$$

$$P(a \mid b) = P(b) \cdot P(a, b)$$

$$=$$

$P(A) = 0.4$

$P(B) = 0.3$

$P(C) = 0.2$

$P(D) = 0.1$

$$\sqrt[N]{\prod_{i=1}^{N} \frac{1}{P(W_i)}}$$

$$= \sqrt[4]{\frac{1}{0.4 \times 0.3 \times 0.4 \times 0.2}}$$

$$= \frac{10000}{4 \times 4 \times 3 \times 2} = \sqrt[4]{\frac{10000}{96}}$$

"A C"

$P(A \mid \langle s \rangle) \times P(C \mid A) \times P(\langle /s \rangle \mid C)$

$= 0.6 \times 0.7 \times 1.0$

$= \frac{6}{10} \times \frac{7}{10}$

$\sqrt[3]{\frac{100}{42}} =$

$$\frac{1}{0.5 \times 0.3 \times 0.1 \times 0.1 \times 0.1}$$

$$= \frac{100000}{15}$$

$$\sqrt[5]{\frac{10000}{15}} =$$