Perplexity Example

corpus:

dogs bark often
cats meow sometimes
dogs don't meow
cats don't bark

unigram:

dogs	bark	often	cats	meow	sometimes	don't
$\frac{2}{12}$	$\frac{2}{12}$	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{2}{12}$	$\frac{1}{12}$	$\frac{2}{12}$

test sentences:

dogs don't bark
cats meow

PP(dogs don't bark cats meow)

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

$$= \sqrt[5]{\frac{1}{P(\text{dogs}) \cdot P(\text{don't}) \cdot P(\text{bark}) \cdot P(\text{cats}) \cdot P(\text{meow})}}$$

$$= \sqrt[5]{\frac{1}{\frac{2}{12} \cdot \frac{2}{12} \cdot \frac{2}{12} \cdot \frac{2}{12} \cdot \frac{2}{12}}}$$

$$= \left(\frac{1}{\frac{2}{12} \cdot \frac{2}{12} \cdot \ldots \cdot \frac{2}{12}} \right)^{\frac{1}{5}}$$

$$= \left(\frac{2}{12} \cdot \frac{2}{12} \cdot \ldots \cdot \frac{2}{12} \right)^{-\frac{1}{5}}$$

$$= 6$$

Unigram PP calculation using $\log_{10}$ probs

dogs bark often cats meow sutms don't

$\log(\frac{1}{6})$ $\log(\frac{1}{6})$ $\log(\frac{1}{12})$ $\log(\frac{1}{6})$ $\log(\frac{1}{6})$ $\log(\frac{1}{12})$ $\log(\frac{1}{6})$

$\approx$ $-.78$ $-.78$ -1.08 $-.78$ $-.78$ -1.08 $-.78$

PP (dogs don't bark cats meow)

$$= \sqrt[N]{\prod_i \frac{1}{P(\omega_i)}}$$

$$= \sqrt[N]{\frac{1}{10^{\sum \log(P(\omega_i)}}}$$

$$= \sqrt[5]{\frac{1}{10^{-.78 - .78 - .78 - .78 - .78}}}$$

$$= \left(\frac{1}{10^{-3.9}}\right)^{\frac{1}{5}}$$

$$= \left(10^{-3.9}\right)^{-\frac{1}{5}}$$

$$= 10^{+\frac{3.9}{5}}$$

$$= 6$$

$$PP = 10^{-\frac{\sum \log(P(\omega_i))}{N}}$$

<s> dogs bark often </s>
<s> cats meow sometimes </s>
<s> dogs don't meow </s>
<s> cats don't bark </s>

bigram:

unigram counts ↓

	dogs	bark	often	cats	meow	sometimes	don't	</s>	
<s>	2/4			2/4					4
dogs		1/2					1/2		2
bark			1/2					1/2	2
often								1	1
cats					1/2		1/2		2
meow						1/2		1/2	2
smtms								1	1
don't		1/2			1/2				2

= # n-grams

$N = 4 + 3 = 7$

test sentences: <s> dogs don't bark </s>
<s> cats meow </s>

$$PP(\text{<s> dogs don't bark </s> <s> cats meow </s>})$$

$$= \sqrt[7]{\frac{1}{P(\text{dogs}|\text{<s>}) \cdot P(\text{don't}|\text{dogs}) \cdot P(\text{bark}|\text{don't}) \cdot P(\text{</s>}|\text{bark}) \cdot P(\text{cats}|\text{<s>}) \cdot P(\text{meow}|\text{cats}) \cdot P(\text{</s>}|\text{meow})}}$$

$$= \sqrt[7]{\frac{1}{\frac{2}{4} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{2}{4} \cdot \frac{1}{2} \cdot \frac{1}{2}}}$$

$$= \left(\frac{1}{.0078}\right)^{\frac{1}{7}}$$

$$= (.0078)^{-\frac{1}{7}}$$

$$= 2$$

(as expected, less than perplexity of the unigram model)

	dogs	bark	often	cats	meow	sometimes	don't	</s>
<s>	-.3			-.3				
dogs		-.3					-.3	
bark			-.3					-.3
often								0
cats					-.3		-.3	
meow						-.3		-.3
smtms								0
don't		-.3			-.3			

using base 10:

$10^x = a$ ← $P(w_2|w_1)$

$\log_{10}(a) = x$ ← $\log_{10}(P(w_2|w_1))$

in the probability matrix, replace each $P(w_2|w_1)$ with its log: $\log_{10}(P(w_2|w_1))$

$PP(\text{<s> dogs don't bark </s> <s> cats meow </s>})$

$$= \sqrt[N]{\prod_{i}^{N} \frac{1}{P(w_2|w_1)}}$$

$$= \sqrt[N]{\frac{1}{10^{\sum \log(P(w_2|w_1))}}}$$

$$= \left(\frac{1}{10^{\sum \log(P(w_2|w_1))}}\right)^{\frac{1}{N}}$$

$$= \left(10^{\sum \log(P(w_2|w_1))}\right)^{-\frac{1}{N}}$$

$$= 10^{-\frac{\sum \log(P(w_2|w_1))}{N}}$$

$$= 10^{-\frac{\left[\log(P(\text{dogs}|\text{<s>})) + \log(P(\text{don't}|\text{dogs})) + \log(P(\text{bark}|\text{don't})) + \log(P(\text{</s>}|\text{bark})) + \log(P(\text{cats}|\text{<s>})) + \log(P(\text{meow}|\text{cats})) + \log(P(\text{</s>}|\text{meow}))\right]}{N}}$$

$$= 10^{-\left(\frac{-.3 - .3 - .3 - .3 - .3 - .3 - .3}{7}\right)} = 10^{-\left(\frac{-2.1}{7}\right)} = 2$$