CS224N Assignment #4

2. Analyzing NMT Systems

(c) NMT Translation C_1

unigrams = {the, love, can, always, do}

$$p_1 = \frac{0+1+1+1+0}{5} = \frac{3}{5}$$

bigrams = {the love, love can, can always, always do}

$$p_2 = \frac{0+1+1+0}{4} = \frac{1}{2}$$

$c = 5$, $r^* = 4$

$$BLEU = BP \times e^{\sum_{n=1}^{4} \lambda_n \log p_n}$$

$$= 1 \cdot e^{0.5 \cdot \log \frac{3}{5} + 0.5 \cdot \log \frac{1}{2}}$$

$$= 0.770$$

NMT Translation C_2

unigrams = {love, can, make, anything, possible}

$$p_1 = \frac{1+1+1+1}{5} = \frac{4}{5}$$

bigrams = {love can, can make, make anything, anything possible}

$$p_2 = \frac{1+0+0+1}{4} = \frac{1}{2}$$

$c = 5, \quad r^* = 4$

$$\text{BLEU} = \text{BP} \times e^{\sum_{n=1}^{4} \lambda_n \log p_n}$$

$$= 1 \times e^{0.5 \cdot \log \frac{4}{5} + 0.5 \cdot \log \frac{1}{2}}$$

$$= 0.820$$

c_2 is a better translation and I agree.