

1. Given the following bigram probabilities:

	i	want	to	eat	chinese	food	lunch	spend
i	0.002	0.33	0	0.0036	0	0	0	0.00079
want	0.0022	0	0.66	0.0011	0.0065	0.0065	0.0054	0.0011
to	0.00083	0	0.0017	0.28	0.00083	0	0.0025	0.087
eat	0	0	0.0027	0	0.021	0.0027	0.056	0
chinese	0.0063	0	0	0	0	0.52	0.0063	0
food	0.014	0	0.014	0	0.00092	0.0037	0	0
lunch	0.0059	0	0	0	0	0.0029	0	0
spend	0.0036	0	0.0036	0	0	0	0	0

$$P(i | <s>) = 0.25$$

$$P(\text{english} | \text{want}) = 0.0011$$

$$P(\text{food} | \text{english}) = 0.5$$

$$P(</s> | \text{food}) = 0.68$$

Compute the probability of “I want English food” and “I want Chinese food” using bigram model

2. Define perplexity and explain its relationship with the branching factor of a language.
3. What is the purpose of smoothing? Explain add-1 smoothing.
4. What is the shortcoming of add-1 smoothing that add-k smoothing overcomes and how?
5. Explain (i) backoff and interpolation and (ii) Kneser-Ney smoothing techniques