Prelim Notes

Dylan Eric Robert Shubham

1. Natural Language Processing and Speech

1.1. Chapter 1-4 (July 18)

Discussions

- Are n-grams defined over word forms or word roots ? Ans - application specific.
- Good-Turing smoothing : we derived Eq 4.27 using Eq 4.26 and briefly discussed the approximation used for larger N (Simple-Good turing)
- Brief discussion of interpolation and katz-backoff.
- Discussed back-off in Kneser Ney. An unanswered question was regarding implementation of back-off from n-grams to (n-1)-grams for $n > 1$ (do we use context or back-off to Kneser-Ney probabilities ?)

Topics for Review

- Kneser-Ney
- Perplexity
- Good Turing (formula)

2. Artificial Intelligence