Language Modeling Part I

What is language modeling?

- Assigns probabilities to sequences of tokens
  - Tokens could be words, characters; in the general case character strings

P(`pocket`|`whats in your`)

- Why should we approach language understanding this way?
  - There are many applications for which this is useful, for example:
    - Machine translation

    or   P(`pocket`|`from capital one. whats in your`)

        P(`wallet`|`from capital one. whats in your`)

    - Speech recognition

    or   P(`yeast`|`i live in ____ village`)

        P(`east `|`i live in ____ village`)

  - Most importantly, we can train deep neural networks using language modeling procedures to learn features from large amounts of text.