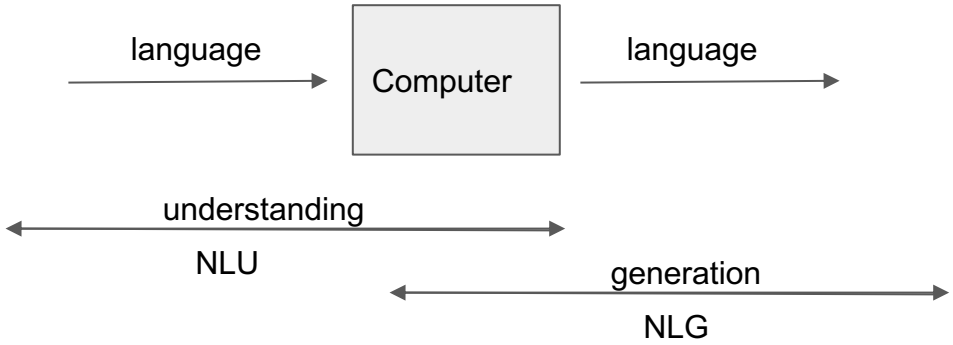
ELL 881: Fundamentals of Deep Learning
Lec 01: Deep Learning & NLP

Vineet Kumar
July 27, 2018

What is NLP? (1/2)

What is NLP? (2/2)

Using computers to process human or natural languages



Why is NLP hard? (1/3)

“At last, a computer that understands you like your mother”

Why is NLP hard? (2/3)

"At last, a computer that understands you like your mother"

1. It understands you as well as your mother understands you
2. It understands (that) you like your mother
3. It understands you as well as it understands your mother

Why is NLP hard? (3/3)

"At last, a computer that understands you like your mother"

1. It understands you as well as your mother understands you
2. It understands (that) you like your mother
3. It understands you as well as it understands your mother

Ambiguity at many levels (1 / 2)

At the acoustic level (speech recognition)

“..., a computer that understands you like your mother”

“..., a computer that understands you lie cured mother”

Ambiguity at many levels (2 / 2)

At the acoustic level (speech recognition)

"..., a computer that understands you like your mother"

Much more likely!

"..., a computer that understands you lie cured mother"

Word Representations (1 / 5)

- Vast majority of NLP work regards words as atomic symbols

Word Representations (2 / 5)

- Vast majority of NLP work regards words as atomic symbols
- In terms of vectors, a word is represented by a vector with a **single** 1 and rest zeroes. Also known as one-hot encoding

  [0 0 0 0 0 1 0 0 …….]

Word Representations (3 / 5)

- Vast majority of NLP work regards words as atomic symbols
- In terms of vectors, a word is represented by a vector with a **single** 1 and rest zeroes. Also known as one-hot encoding

- Problem with one-hot representation
  - Hotel = [0 0 0 0 0 1 0 0 …….] AND
  - Motel = [0 0 0 0 0 0 1 0 …….] = 0

Adapted from Deep Learning for NLP Tutorial Socher et al. NAACL 2013

Word Representations (4 / 5)

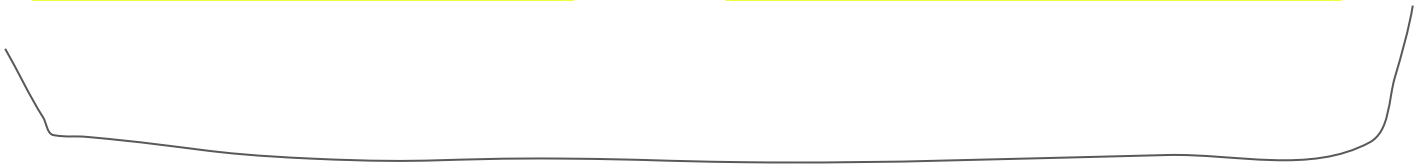
"You shall know a word by the company it keeps" (JR Firth 1957)

- You can get a lot of value by representing a word by means of its neighbors
- One of the most successful ideas of modern statistical NLP
  - government debt problems turning into banking crises as has happened in
  - saying that Europe needs unified banking regulation to replace the hodgepodge

Word Representations (5 / 5)

"You shall know a word by the company it keeps" (JR Firth 1957)

- You can get a lot of value by representing a word by means of its neighbors
- One of the most successful ideas of modern statistical NLP
  - government debt problems turning into banking crises as has happened in
  - saying that Europe needs unified banking regulation to replace the hodgepodge



Context words can be used to represent *banking*

Adapted from Deep Learning for NLP Tutorial Socher et al. NAACL 2013

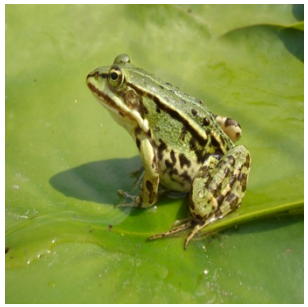
Word similarities

Nearest words to frog:

1. frogs
2. toad
3. litoria
4. leptodactylidae
5. rana
6. lizard
7. eleutherodactylus

litoria

leptodactylidae



rana

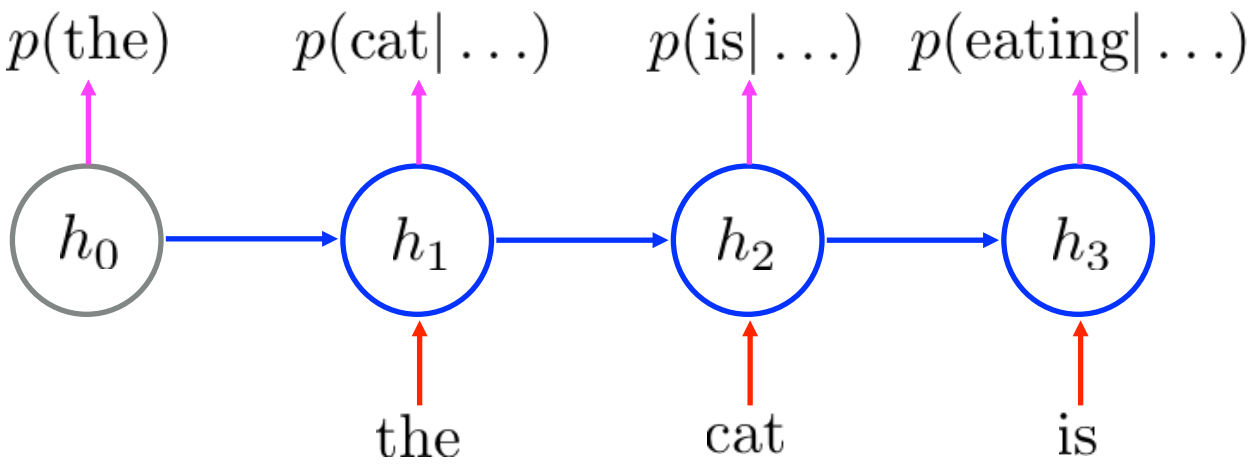
eleutherodactylus

http://nlp.stanford.edu/projects/glove/

Adapted from http://cs224d.stanford.edu/lectures/CS224d-Lecture1.pdf

Dialogue agents / Response Generation

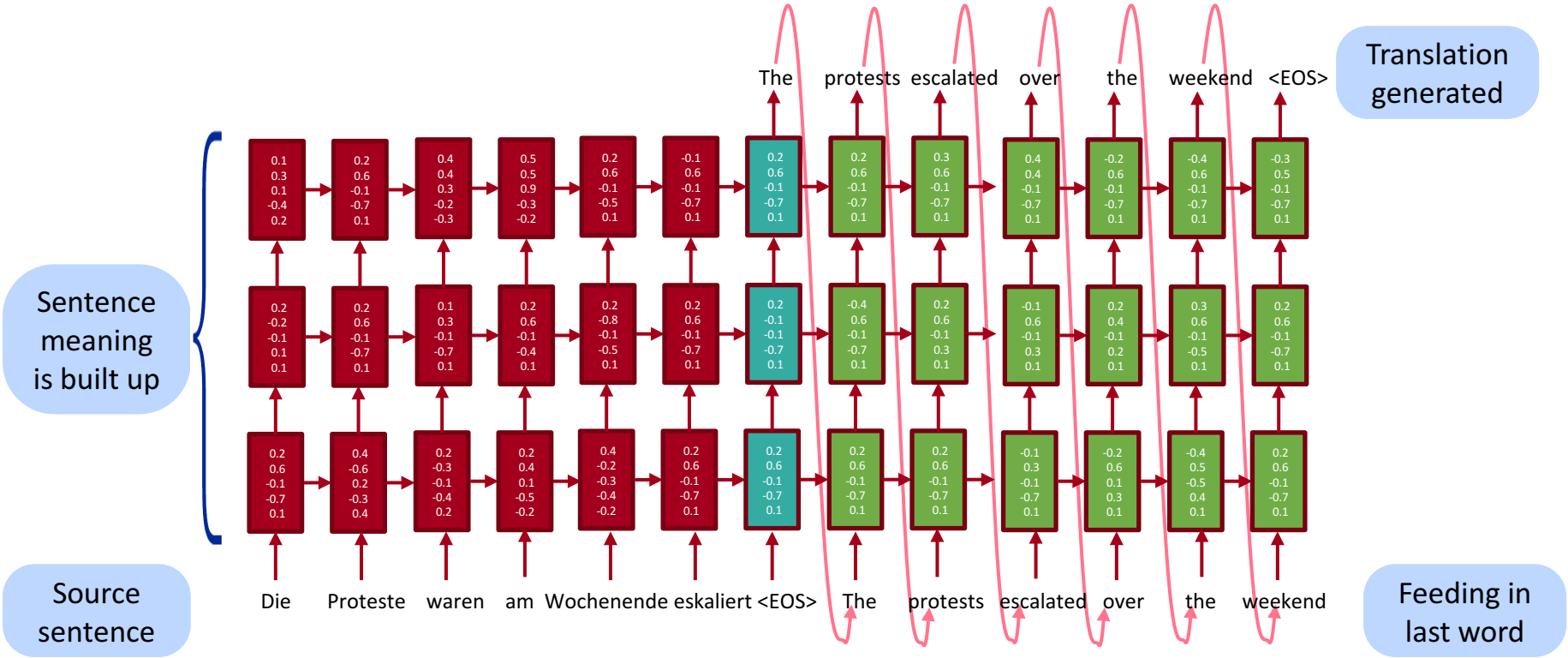
- A simple, successful example is the auto-replies available in the Google Inbox app
- An application of the powerful, general technique of **Neural Language Models**, which are an instance of Recurrent Neural Networks



Adapted from http://cs224d.stanford.edu/lectures/CS224d-Lecture1.pdf

Neural Machine Translation

Source sentence is mapped to **vector**, then output sentence generated
[Sutskever et al. 2014, Bahdanau et al. 2014, Luong and Manning 2016]



Now live for some languages in Google Translate (etc.), with big error reductions!

Adapted from http://cs224d.stanford.edu/lectures/CS224d-Lecture1.pdf