

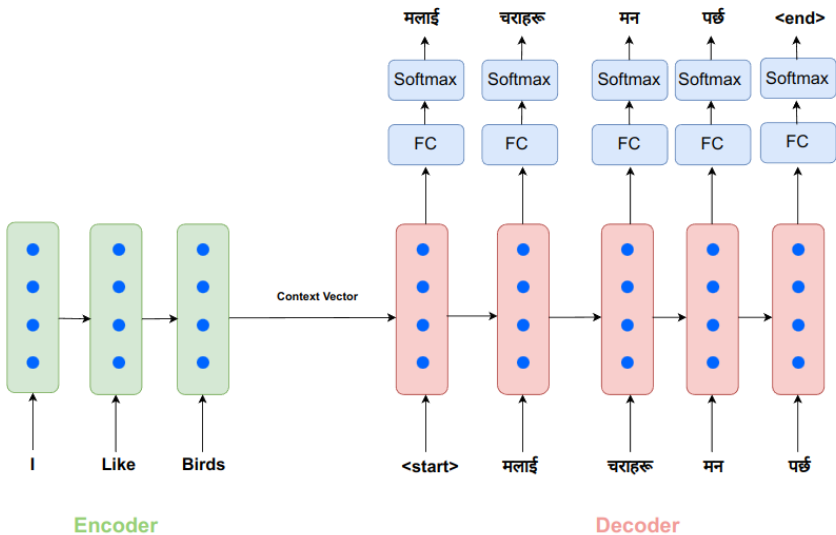
Machine Translation

Jordan Boyd-Graber

University of Maryland

Neural Models

Adapted from material by Philipp Koehn, Mohit Iyyer, Luke Zettlemoyer, Kalpesh Krishna, Karthik Narasimhan, Greg Durrett, Chris Manning, Dan Jurafsky



From Aayush Neupane

Neural MT for the Win

- Fringe in 2014: First seq2seq paper published
- 2016: Google Translate switches from SMT to NMT

Neural MT for the Win

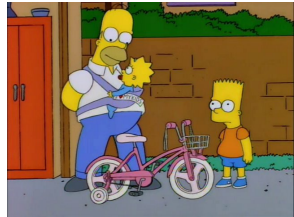
- Fringe in 2014: First seq2seq paper published
- 2016: Google Translate switches from SMT to NMT
- Why?
 - ▶ Representations allow generalization
 - ▶ Not tied to “words”
 - ▶ Attention allows long-range generalization

Distributed Representations

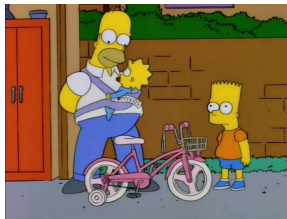
Distributed Representations



Distributed Representations



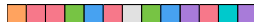
Distributed Representations



Distributed Representations



Distributed Representations



Weasel Trainer



Weasel Trainer



BIRDS AREN'T REAL



THE
NEWS WITH
SHEPARD
SMITH

**"BIRDS AREN'T REAL" IS A PARODY
CONSPIRACY THEORY "WITH A PURPOSE"**





Birds Aren't Real

A large white billboard with a black border. On the top edge of the billboard, there is a row of approximately 15 small black birds perched. The billboard is set against a blue sky with some clouds. The billboard is supported by a metal structure, and the word 'LAMAR' is visible on the structure. The number '054830' is also visible on the right side of the structure.

LAMAR

054830

A billboard is shown against a sky with a sunset or sunrise. The billboard has a white face with a black border. The text 'Words Aren't Real' is written in a large, bold, black sans-serif font. The word 'Aren't' is slightly more prominent than 'Words' and 'Real'. Above the billboard, a row of small birds is silhouetted against the sky. The billboard is supported by a dark metal frame. At the bottom of the billboard, there is a brown section with the 'LAMAR' logo in the center and the number '054830' on the right. The sky is a mix of blue and orange, with a few clouds. The overall mood is contemplative and artistic.

Words Aren't Real

LAMAR

054830

Word / Phrase Translation in Close Languages

Language	Sentence
English	How much does the weasel trainer charge?
Portuguese	Quanto cobra o treinador de doninhas?
Spanish	¿Cuánto cobra un entrenador de comadreja?

What finer-grained outputs makes possible

- Capturing agreement markers (case, gender, number)
- Honorific / tense particles
- Topic markers
- Otherwise needs to be memorized (limits generalization)

Ist dir kalt?

Are you cold?

Ist dir kalt?

Are you cold?

Ist dir kalt?

Are you cold?

Ist dir ohne deine blaue Jacke kalt?

Are you cold without your blue jacket?

Ist dir kalt?

Are you cold?

Ist dir ohne deine blaue Jacke kalt?

Are you cold without your blue jacket?

Ist dir kalt?

Are you cold?

Ist dir ohne deine Jacke kalt?

Are you cold with your jacket?



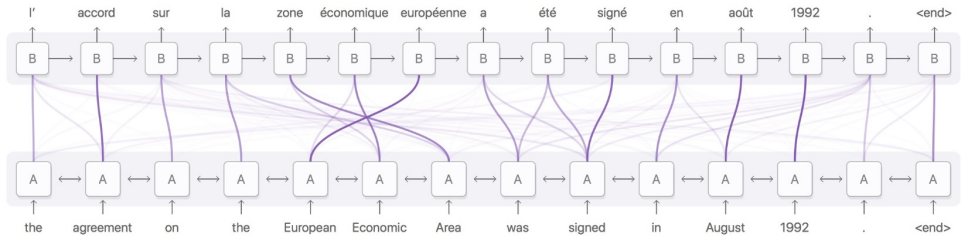
Ist dir ohne deine neue Kettensäge und Hockeymaske kalt?

Are you cold without your new chainsaw and hockey mask?

Ist dir ohne deine **neue** Kettensäge und Hockeymaske kalt?



Are you cold without your new chainsaw and hockey mask?



From Chris Olah <https://distill.pub/2016/augmented-rnns/>

