The role of Artifical Intelligence within Natural Language

Cristina España-Bonet

Universität des Saarlandes (UdS)

Deutsche Forschungszentrum für Künstliche Intelligenz (DFKI)

Toward Multilingual Public Services in Europe

Brussels, Belgium
17th October 2018

Artificial Intelligence

What is AI?

Taxonomy of AI definitions (2)

To reason human

- *"The exciting new effort to make computers think . . . machines with minds, in the full and literal sense"* (Haugeland, 1985)
- *"The automation of activities that we associate with human thinking, activities such as decision-making, problem solving, learning. . . "* (Bellman, 1978)

To reason rationally

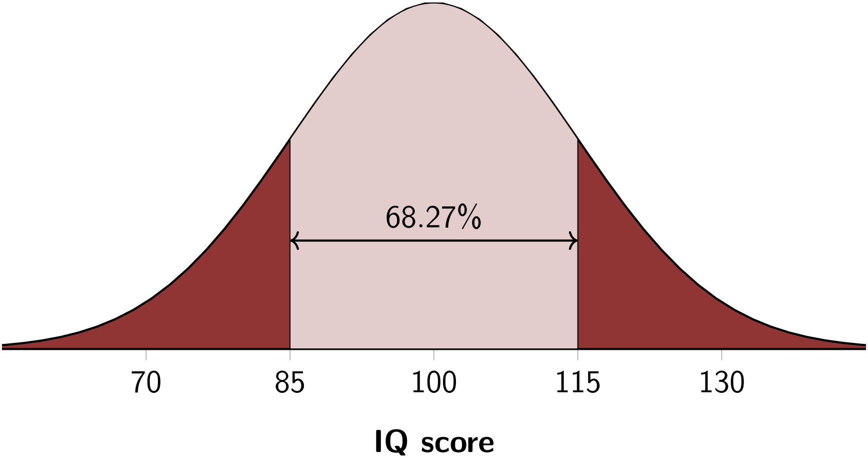
- *"The study of mental faculties through the use of computational models."* (Charniak i McDermott, 1985)
- *"The study of the computations that make it possible to perceive, reason, act."* (Winston, 1992)

To act human

- *"The art of creating machines that perform functions that require intelligence when performed by people."* (Kurzwil, 1990)
- *"The study of how to make computers do things at which, at the moment, people do better."* (Rich i Knight, 1991)

To act rationally

- *"The field of study that seeks to explain and emulate intelligent behavior in terms of computational processes."* (Schalkoff, 1990)
- *"The branch of computer science concerned with automation of intelligent behavior"* (Luger i Stubblefield, 1993)


68.27%
70
85
100
115
130
IQ score

Artificial Intelligence

What is (A)I?

What is (A)I?

- No universally accepted definition
- Is the complement easier to define?

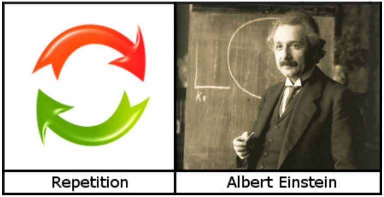
What is (A)I?

- No universally accepted definition
- Is the complement easier to define?
- **Stupidity?**

  Stupidity is doing the same thing over and over again and expecting different results!

Insanity Is Doing the Same Thing Over and Over Again and Expecting Different Results

Albert Einstein? Narcotics Anonymous? Max Nordau? George Bernard Shaw? George A. Kelly? Rita Mae Brown? John Larroquette? Jessie Potter? Werner Erhard?



https://quoteinvestigator.com/2017/03/23/same/

- Anyway, how do we avoid stupidity?

- Anyway, how do we avoid stupidity?
- By **learning** from examples, mistakes, the environment, etc.
- Key part of intelligence

- Anyway, how do we avoid stupidity?
- By **learning** from examples, mistakes, the environment, etc.
- Key part of intelligence
- Machine learning (deep learning!)

Artificial Intelligence

Achievements: Deep Blue, Chess

Learning vs. Brute Force

- 1997: Deep Blue versus Garry Kasparov
- Deep Blue won by brute force
- Chess: moves/step ~ 35
- *Thinking* 8 steps ahead means $\sim 2 \cdot 10^{12}$ states

Artificial Intelligence

Learning vs. Brute Force

- 1997: Deep Blue versus Garry Kasparov
- Deep Blue won by brute force
- Chess: moves/step ~ 35
- *Thinking* 8 steps ahead means $\sim 2 \cdot 10^{12}$ states

That's brute force!

Understanding of **Natural Languages**
cannot be approached by brute force

Natural Language understanding is an
AI-complete problem

(see later)

Artificial Intelligence

Achievements: AlphaGo, Go

Learning vs. Brute Force

- 2016: AlphaGo versus Lee Sedol
- Go: moves/step ~ 250
- Search space $> 10^{100}$!
- AlphaGo won using a Monte-Carlo tree search guided by deep neural networks with reinforcement learning

Learning vs. Brute Force

- 2016: AlphaGo versus Lee Sedol
- Go: moves/step ~ 250
- Search space $> 10^{100}$!
- AlphaGo won using a Monte-Carlo tree search guided by deep neural networks with reinforcement learning

That's (deep) learning!

Achievements: IBM Watson @Jeopardy

Learning vs. Brute Force

- 2011: IBM Watson @Jeopardy
- Question Answering System:
- Natural language processing, knowledge representation, reasoning, information retrieval...

Natural Language Understanding

- 2011: IBM Watson @Jeopardy
- Question Answering System:
- Natural language processing, knowledge representation, reasoning, information retrieval...

Natural Language Understanding

- 2018: health, weather, chatbots...

Artificial Intelligence

Achievements: (Neural) Machine Translation

Microsoft

Translator Home Solutions Products Languages Blog Support More

Neural Machine Translation reaches historic milestone: human parity for Chinese to English translations

Rate this article ★★★★★

Microsoft Translator March 14, 2018

Share 12 7 in 0 0

Natural Languages

1. Artificial Intelligence
2. Natural Languages
3. Natural Language Processing
4. Neural Networks and Deep Learning
5. Tomorrow (today!) in Europe

Natural Languages

Description

Human languages are a tool to
communicate thoughts and they are

elegant, **efficient**, **flexible**, **complex**

Natural Languages

Description

Human languages are a tool to communicate thoughts and they are

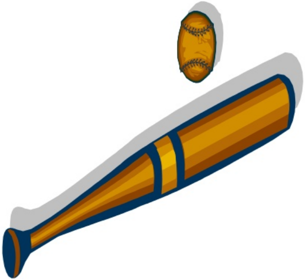
elegant, **efficient**, **flexible**, **complex**

Cool for a human, but for a machine...

Natural Languages

One word/sentence may mean many things

Homonymy and Polysemy

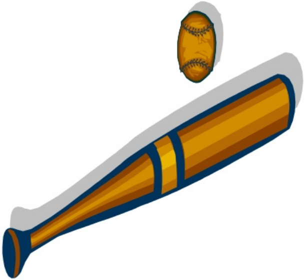


Natural Languages

One word/sentence may mean many things

Homonymy and Polysemy

bat



Natural Languages

Meaning depends on context

Bats!

Natural Languages

Meaning depends on context

Bats! *(I'm inside a cave)*

Natural Languages

Meaning depends on context

Bats! *(I'm inside a cave)*

This **bat** is brown

Natural Languages

Meaning depends on context

Bats! *(I'm inside a cave)*

This **bat** is brown (*I'm watching a baseball match*)

Natural Languages

Meaning depends on context

Bats! *(I'm inside a cave)*

This **bat** is brown (*I'm watching a baseball match)*

Streaks and slumps are as common to baseball as **bats**

He **bats** the flies away

We **bat** around a wide variety of issues

When he told me what he'd done, I didn't **bat** an eye

Natural Languages

Many ways of saying the same thing

Synonymy

gift **present**

Natural Languages

Literal and figurative language (metaphors)

This coffee shop **is an ice box**

vs.

It is very cold in this coffee shop

Natural Languages

Language is culture

The early bird catches the worm

vs.

Morgenstund hat Gold im Mund

vs.

Qui matina fa farina

vs.

A quien madruga, Dios le ayuda

Natural Languages

Language is culture

Jurafsky & Martin, 2007 (Ch.25)

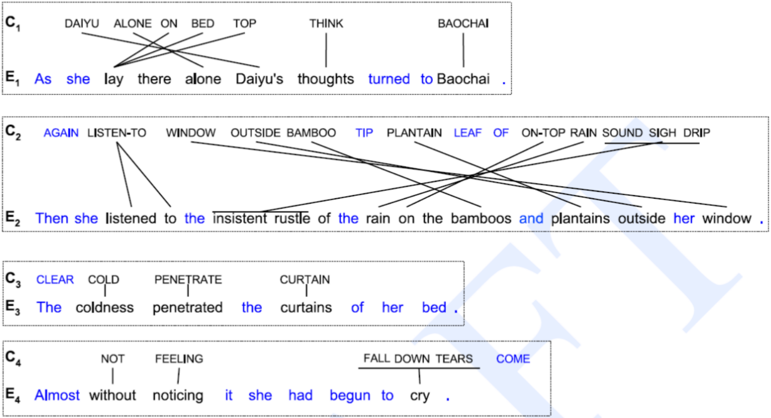


Figure 25.1 A Chinese passage from *Dream of the Red Chamber*, with the Chinese words represented by English glosses IN SMALL CAPS. Alignment lines are drawn between 'Chinese' words and their English translations. Words in blue are Chinese words not translated into English, or English words not in the original Chinese.

Understanding of **Natural Languages**
cannot be approached by brute force

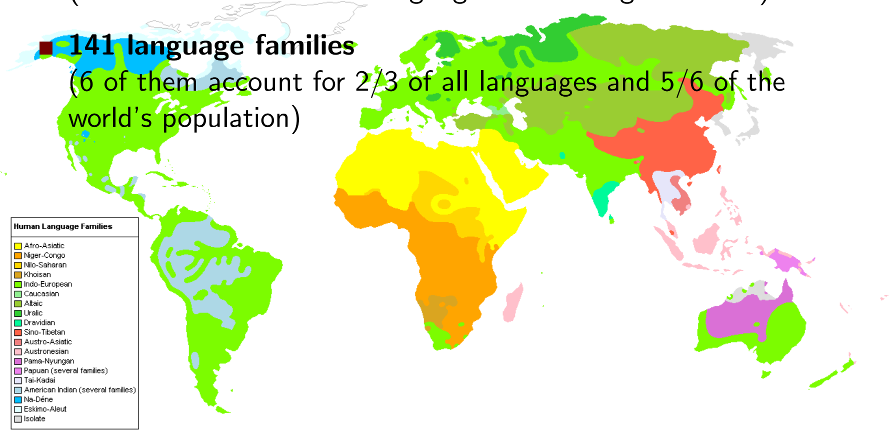
Natural Language understanding is an
AI-complete problem

(from the intro)

Natural Languages

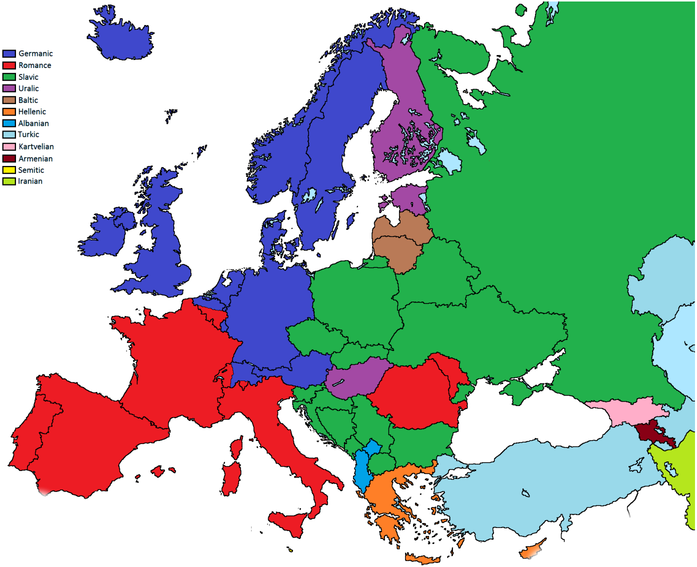
Some Numbers

- There are more than **7000 languages**
  (even if the definition of language is not straightforward!)
- **141 language families**
  (6 of them account for 2/3 of all languages and 5/6 of the world's population)



Natural Languages

Language Families in Europe

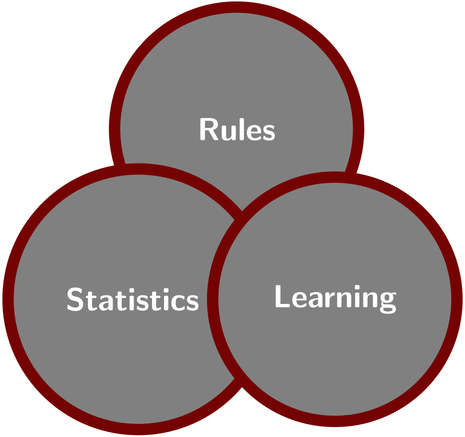


Outline

1. Artificial Intelligence
2. Natural Languages
3. Natural Language Processing
4. Neural Networks and Deep Learning
5. Tomorrow (today!) in Europe

Natural Language Processing

General Approaches



Rules in Machine Translation

```
function DIRECT_TRANSLATE_MUCH/MANY(word) returns Russian translation

if preceding word is how return skol'ko
else if preceding word is as return stol'ko zhe
else if word is much
   if preceding word is very return nil
   else if following word is a noun return mnogo
else /* word is many */
   if preceding word is a preposition and following word is a noun return mnogii
   else return mnogo
```

Figure 25.7 A procedure for translating *much* and *many* into Russian, adapted from Hutchins' (1986, pg. 133) discussion of Panov 1960. Note the similarity to decision list algorithms for word sense disambiguation.

(from Jurafsky & Martin, 2007)

Natural Language Processing

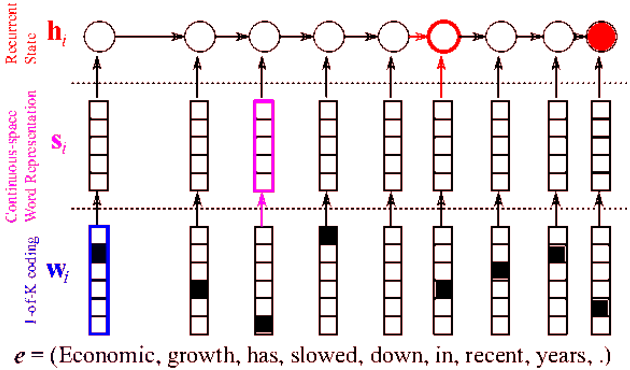
Statistics in Machine Translation

	No.	CLASSIC SOUPS	Sm.	Lg.
清燉雞湯	57.	House Chicken Soup (Chicken, Celery, Potato, Onion, Carrot)	1.50	2.75
雞飯湯	58.	Chicken Rice Soup	1.85	3.25
雞麵湯	59.	Chicken Noodle Soup	1.85	3.25
廣東雲吞	60.	Cantonese Wonton Soup	1.50	2.75
蕃茄蛋湯	61.	Tomato Clear Egg Drop Soup	1.65	2.95
雲吞湯	62.	Regular Wonton Soup	1.10	2.10
酸辣湯	63.	Hot & Sour Soup	1.10	2.10
蛋花湯	64.	Egg Drop Soup	1.10	2.10
雲蛋湯	65.	Egg Drop Wonton Mix	1.10	2.10
豆腐菜湯	66.	Tofu Vegetable Soup	NA	3.50
雞玉米湯	67.	Chicken Corn Cream Soup	NA	3.50
蟹肉玉米湯	68.	Crab Meat Corn Cream Soup	NA	3.50
海鮮湯	69.	Seafood Soup	NA	3.50

(from Josef van Genabith slides)

Natural Language Processing

Learning in Machine Translation



(from Kyunghyun Cho slides)

Standard Corpora

Language Pair	# segments (app.)	# words (app.)
En–De	$4.5 \cdot 10^6$	$222 \cdot 10^6$
En–Fr	$36.3 \cdot 10^6$	$2100 \cdot 10^6$

Standard Corpora

Language Pair	# segments (app.)	# words (app.)
En–De	$4.5 \cdot 10^6$	$222 \cdot 10^6$
En–Fr	$36.3 \cdot 10^6$	$2100 \cdot 10^6$

Books

Title	# words (approx.)
The Bible	$0.8 \cdot 10^6$
The Dark Tower series	$1.2 \cdot 10^6$
Encyclopaedia Britannica	$44 \cdot 10^6$

Named-Entity Recognition

Named-Entity Recognition

POS Tagging

Named-Entity Recognition

POS Tagging

Parsing

Named-Entity Recognition

Sentiment Analysis

POS Tagging

Parsing

Named-Entity Recognition

Sentiment Analysis

POS Tagging

Semantic Role Labeling

Parsing

Named-Entity Recognition

Sentiment Analysis

POS Tagging

Question Answering

Semantic Role Labeling

Parsing

Named-Entity Recognition

Sentiment Analysis

POS Tagging

Question Answering

Semantic Role Labeling

Machine Translation

Parsing

Named-Entity Recognition

Sentiment Analysis

POS Tagging

Question Answering

Semantic Role Labeling

Dialogue

Machine Translation

Parsing

Other NLP Tasks

Named-Entity Recognition

Sentiment Analysis

POS Tagging

Question Answering

Semantic Role Labeling

Dialogue

Text Summarisation

Machine Translation

Parsing

Other NLP Tasks

Seq2Tag

Named-Entity Recognition

Sentiment Analysis

POS Tagging

Question Answering

Semantic Role Labeling

Seq2Seq

Dialogue

Text Summarisation

Machine Translation

Parsing

Yesterday

- **Shallow Machine Learning** models —XGB, SVM, logistic regression— trained on (lots of!) features

Yesterday

- **Shallow Machine Learning** models
  —XGB, SVM, logistic regression—
  trained on (lots of!) features

Tomorrow

- (Deep) **Neural Networks** using **dense vector representations** and learned features

Outline

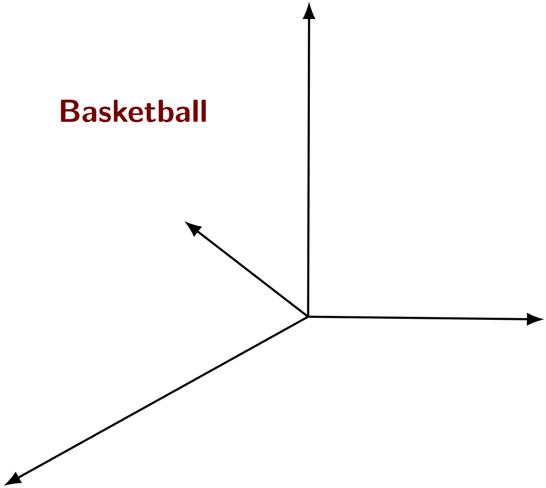
Neural Networks and Deep Learning

A World where Words are Numbers

Basketball = (0.101159, 0.550446, 0.543801, -0.973852, -0.680835, 0.417193, -0.247181, 0.209725, -1.136055, -0.059531, -0.401640, 0.171540, 0.925121, -0.143815, 0.781714, -1.482425, 0.347008, -0.112342, 0.442418, -1.020457, -0.071752, 1.873548, -0.222886, -0.729569, -0.830224, -0.868407, 0.203496, 0.469911, -0.191363, 0.565102, 0.687738, 0.480823, 0.842358, -0.173656, -0.265585, 0.685740, 0.488047, -0.359772, -0.576064, -0.802884, 0.081554, 0.046882, -0.861532, -0.461855, 0.613098, -1.534642, -0.884534, 0.207728, 1.396512, -0.242900, -0.383959, 0.570844, -0.703350, -1.368813, -1.008194, 1.534660, 0.171693, 0.640925, -0.233116, 0.324685, 0.483171, 0.337947, -0.963290, -0.400558, 0.830977, 0.913474, 0.251693, -0.589420, -0.299622, 1.047515, -0.266679, -1.247186, 1.087610, -0.549028, 1.600710, -1.538516, -1.703301, -1.393499, -0.894448, 0.717204, 0.105767, -0.189234, -0.615609, -0.658315, 0.051877, 0.014180, -0.791282, 0.150424, 1.343751, -0.464859, 0.871426, 1.542864, -1.202150, -0.767113, -1.734738, 0.073633, -1.012583, 0.747787, 0.476070, -0.454807, 0.642685, -0.854152, -0.071798, 0.233724, 0.712329, -0.097752, -0.531132, 0.323271, -0.447342, 0.657913, 1.199492, -0.107360, -0.154234, -1.131168, 1.354793, 1.721385, -0.240023, 0.655765, -0.217006, -0.801722, 0.553369, 0.213377, 0.323267, -1.516051, 2.106244, -0.134282, 0.742155, 0.426344, 0.197991, -0.806768, 0.372546, -0.160200, -1.552847, -0.286178, -0.707796, 0.527352, -0.259658, 0.230387, 0.105294, -0.194481, 0.301772, -1.022163, 0.557191, 1.096709, 0.058422, -1.036384, 0.353412, -0.623097, -0.689515, 0.091472, 0.783885, 0.184088, -0.367950, 0.952462, 0.183704, 0.677562, 0.293917, -0.214309, -0.487794, 0.934296, 0.311513, 0.286514, -0.085511, 0.777691, 1.232603, -0.309367, -0.225086, 0.005091, -0.099195, -0.293117, 1.305563, 0.595816, 0.950316, 0.568706, -0.561446, 0.911634, -0.383941, 0.758054, -0.197820, 0.506777, -0.290767, -0.356727, 1.229474, -0.156489, -0.782741, -0.210163, -0.029169, 0.602664, 0.418375, 0.148975, -0.761796, 1.322690, -0.173410, 0.204111, -1.344531, 1.081905, -0.660543, -0.225615, -0.444753, -0.929671, 0.054136, 0.052031, -0.164926, 0.159312, -1.316333, 0.837011, -1.290353, 0.958403, 1.247478, 0.442009, 0.455497, -1.856268, -0.358823, -0.230839, -0.206271, 0.227012, -0.454163, 0.747798, -1.252855, 1.436849, -0.427915, -0.810428, -0.628144, -0.288458, 0.087355, 0.356739, 0.153036, 0.516594, -0.504978, 0.814432, 1.052940, 1.094526, -0.219595, 0.722178, 0.267325, -0.087458, -1.270262, -0.039461, 0.991926, -0.112005, -0.009605, 0.149920, 0.164717, 0.280475, 0.966384, 0.327598, 0.189590, -0.208946, 0.838261, 0.051847, -0.277932, -0.788527, -0.768702, -1.688721, 0.388215, 0.170153, -0.555723, -0.529565, -0.528982, -0.659930, 0.588041, -0.368195, -0.850188, -0.004996, 0.925476, 1.046587, -0.731761, 0.519435, 0.193188, -0.709557, 0.123329, -0.454316, 1.885830, -0.201841, -0.728933, -0.953455, -0.205837, -0.724068, 0.120158, 1.765389, -0.192159, 1.062490, -0.002634, 0.125790, -0.846565, 0.548899, -1.062821, -2.146826, 0.134681, 0.570950, 0.851783, 0.436544, 0.688986, 1.229008, 1.435449, 0.118766, -0.132411, 2.527890, 0.778142)

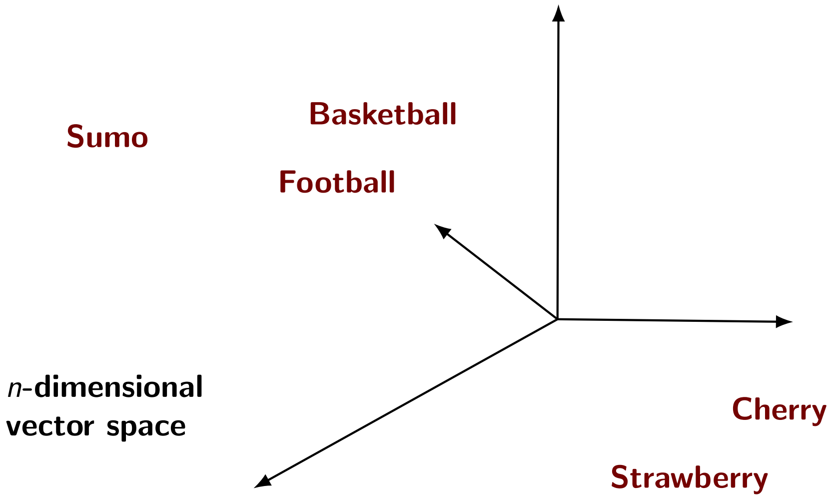
Word Vector Space

Basketball

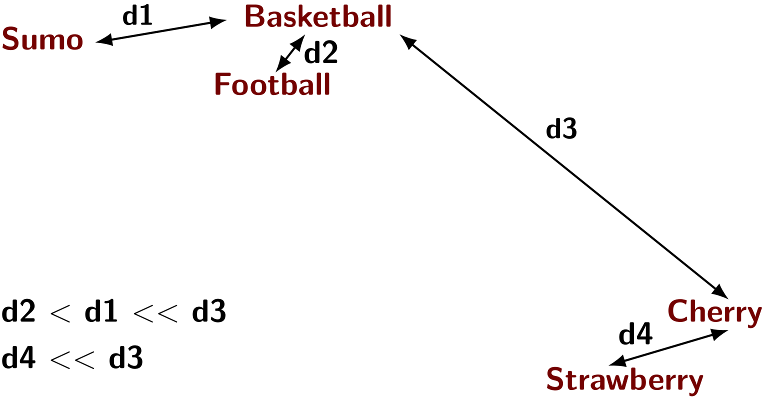


n-dimensional vector space

Word Vector Space



A World where Words are Numbers

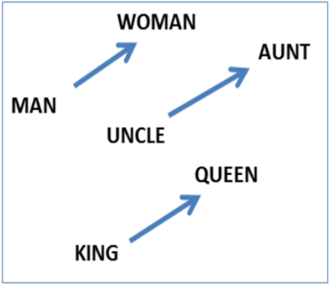


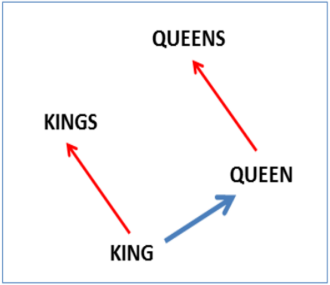
d2 $<$ d1 $<<$ d3

d4 $<<$ d3

Would that allow language understanding?

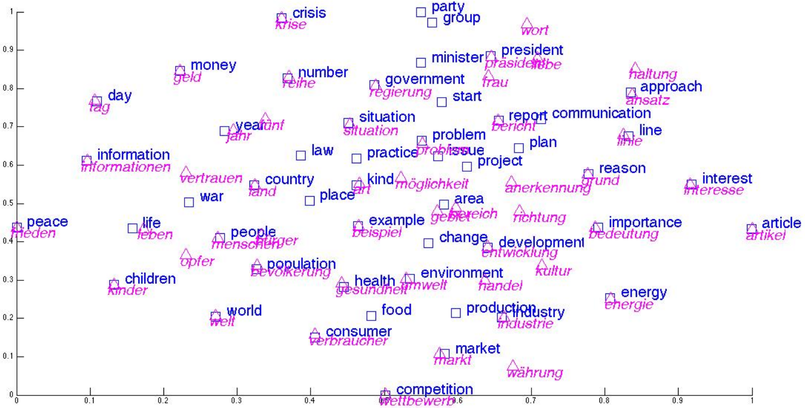
King - Man + Woman = Queen





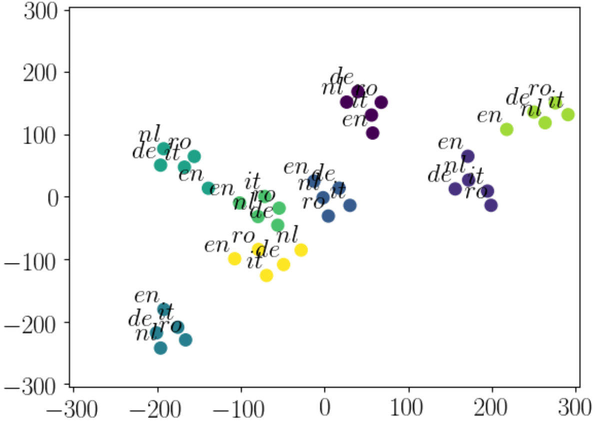
(Mikolov et al., NAACL HLT, 2013)

(Luong, Pham & Manning, NAACL, 2015)



Barnes-Hut-SNE visualisation of bilingual embeddings German/English

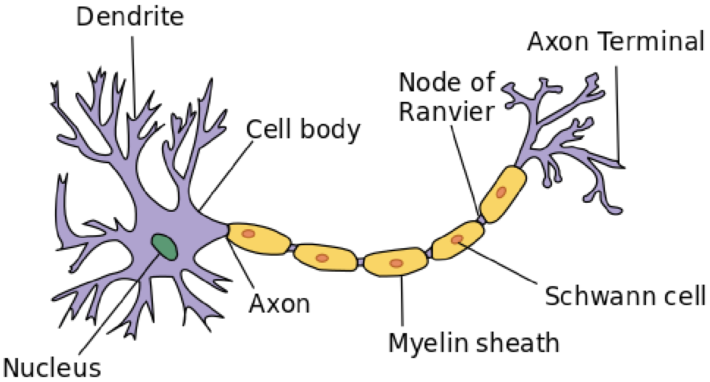
(España-Bonet & van Genabith, 2018)



ML-NMT $\{de, en, nl, it, ro\} \rightarrow \{de, en, nl, it, ro\}$ with TED talks

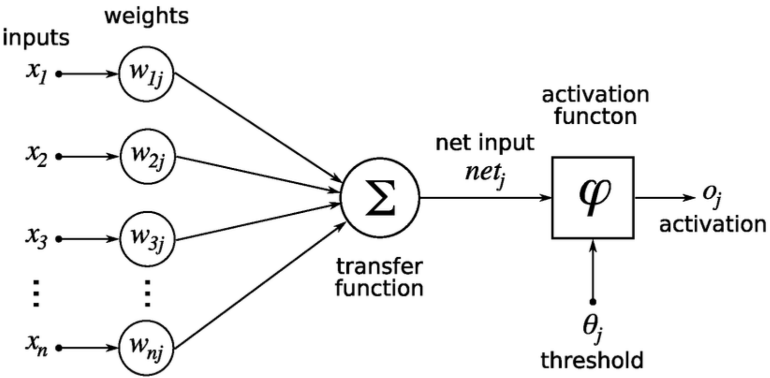
How? What for?

How? Neurons or Units?



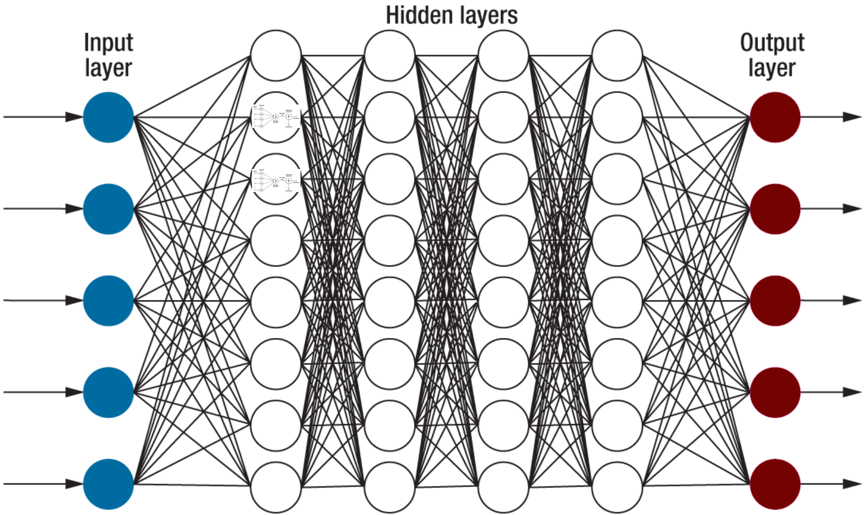
Neural Networks and Deep Learning

How? Neurons or Units?



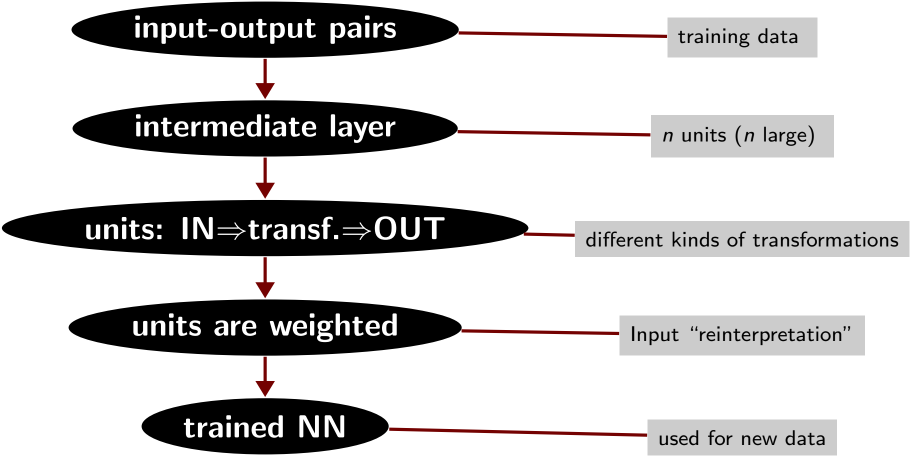
Neural Networks and Deep Learning

How? (Deep) Neural Networks

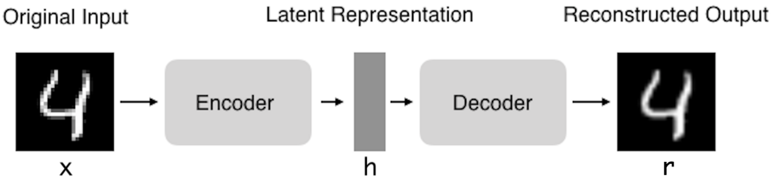


Neural Networks and Deep Learning

Neural Networks (feed-forward-like)



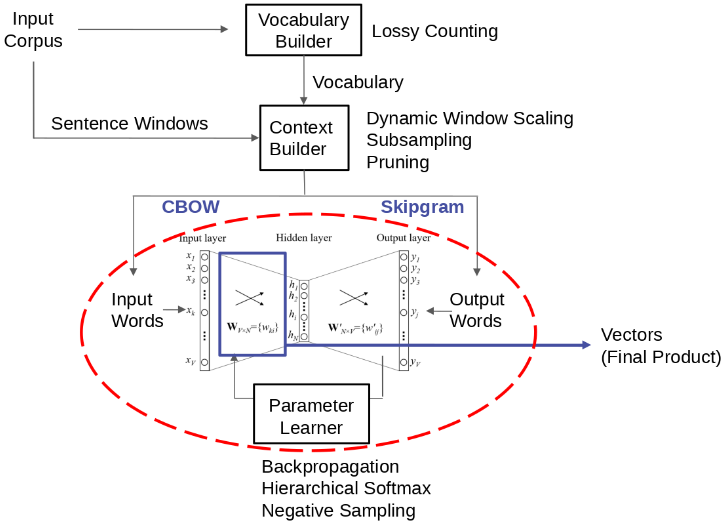
Example: Autoencoder



(https://towardsdatascience.com/deep-inside-autoencoders)

Neural Networks and Deep Learning

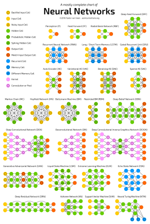
Word Embeddings, also a Reinterpretation (word2vec example)

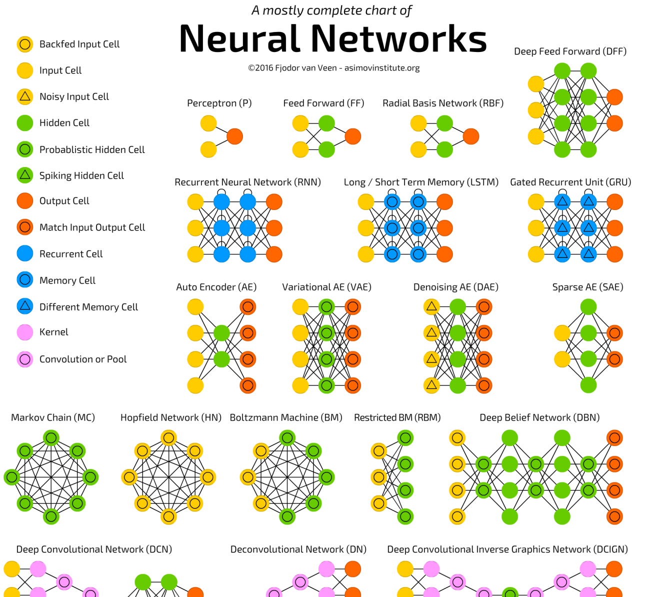


Credits: Xin Rong

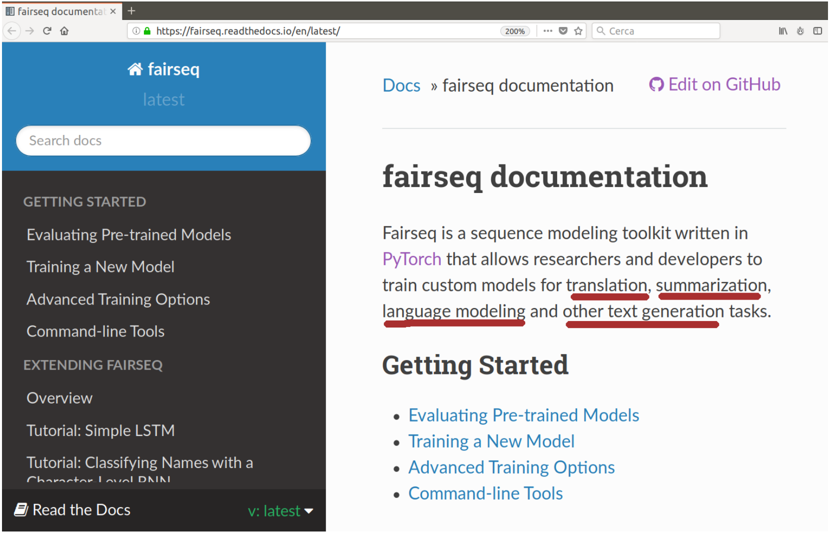
Neural Networks and Deep Learning

Neural Networks, (D)NNs, in Plural!





Seq2Seq: a single NN for Several Tasks



Neural Networks and Deep Learning

Deep Learning: why so good, why now?

NNs with n layers — definition of DL, deep vs. wide

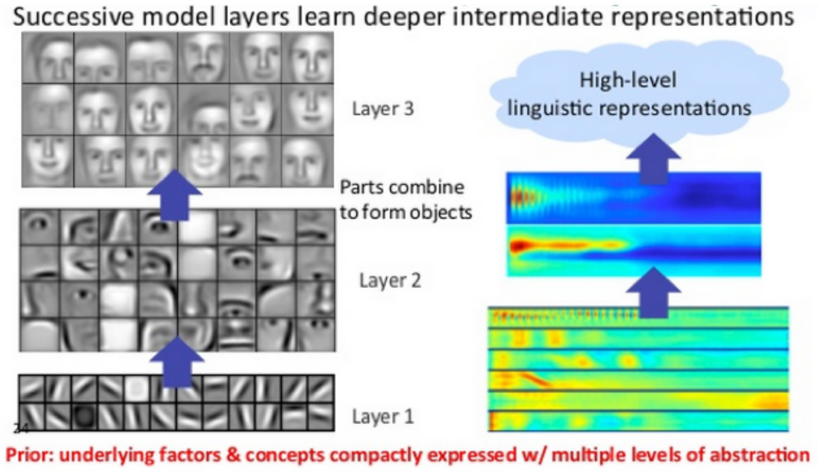
layers extract features — richness of features

non-linearities — capture difficult relations

lots of parameters — need for powerful GPUs

lots of training data — not all tasks can be tackled

Deep Learning: Features



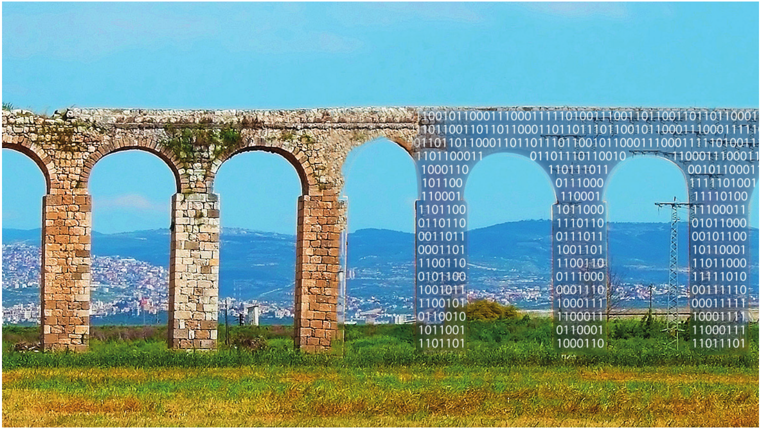
(https://skymind.ai/wiki/neural-network)

Outline

Tomorrow (today!) in Europe

Digital Service Infrastructure

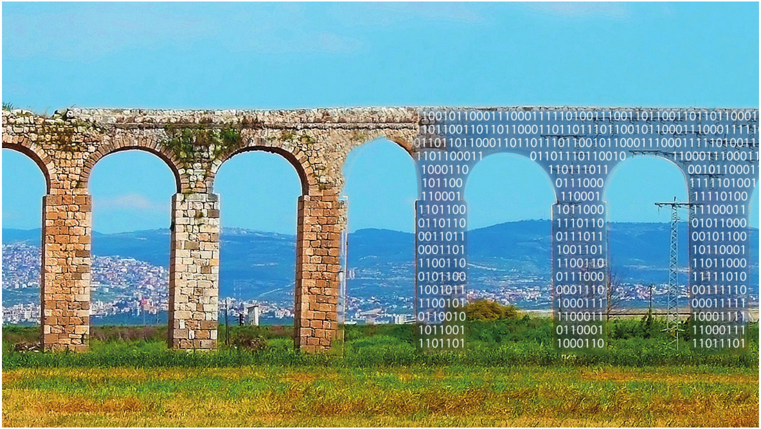
Connecting Europe through digital bridges



Tomorrow (today!) in Europe

Digital Service Infrastructure

Connecting Europe through **cross-lingual** digital bridges



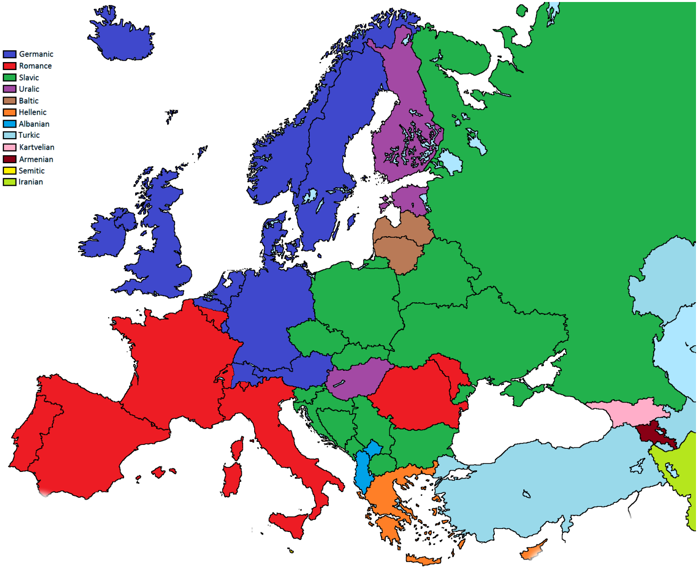
Digression

Time for a disgression on multilinguality with NNs?

let's go!

Tomorrow (today!) in Europe

Different Languages Involved



Tomorrow (today!) in Europe

The role of Artifical Intelligence within Natural Language

- Natural Language Understanding is thought to be an **AI-complete problem**
- Current advances in **computation power** have fired deep learning (DL) for NLP
- DL is achieving impressive results for several natural language tasks (**human performace**)
- In some architectures, **multilinguality** comes for free

Tomorrow (today!) in Europe

Promote All Languages Used in the EU

- DL needs huge amounts of high-quality **data** for good performance
- Lots of data for **official** EU languages
- **Non-official** languages may lack enough data
- In the DL world, languages without data **won't exist**

Basque, Catalan, Gaelic, Luxembourgish, Macedonian, Occitan, Romani, Welsh...

Thanks! And...

wait!

The role of Artifical Intelligence within Natural Language

Cristina España-Bonet

Universität des Saarlandes (UdS)

Deutsche Forschungszentrum für Künstliche Intelligenz (DFKI)

Toward Multilingual Public Services in Europe

Brussels, Belgium
17th October 2018

Similarities and differences among languages

Different cultures share basic **concepts** and **actions** and communicate them with words

- Vocabulary: We all have a mother, a father, and see trees, water...
- Grammar: nouns and verbs

Back-up slides

Cultures & Language: Non-Universal Aspects

But different cultures communicate them with words in different ways and cover different necessities

46 words for *snow* in Icelandic

https://reportsfromtherock.wordpress.com/page/2/

Differences among languages I

Besides different necessities for concepts and lexicon, languages differ in **morphology**

Western Greenlandic:
Aliikkusersuillammassuaanerartassagaluarpaalli

English:
However, they will say that he is a great entertainer but...

Differences among languages II

Declarative sentences with a **S**ubject, a **V**erb and an **O**bject can follow different **syntax** orders

SOV: Japanese, Hindi...
`Inu ga` (subject) `neko` (object) `o oikaketa` (verb)

SVO: English, French, Mandarin...
`The dog` (subject) `chased` (verb) `the cat` (object)

VSO: Irish, Arabic...
`yutarid` (verb) `alkalb` (subject) `alqut` (object)

Differences among languages III

- Omision of elements (e.g. pronouns)

- Short range order

  Adj Noun Blue house
  Noun Adj Casa blava

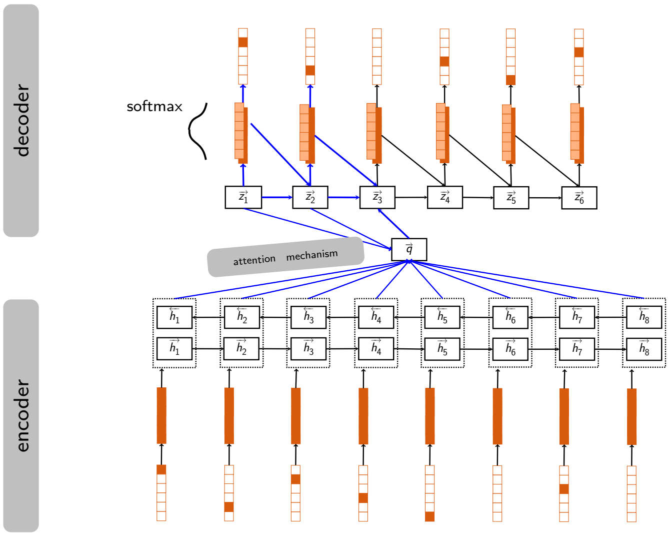
- Structure, e.g. dates

  DD/MM/YY British English
  MM/DD/YY American English
  YY/MM/DD Japanese

Multilinguality with Neural Networks

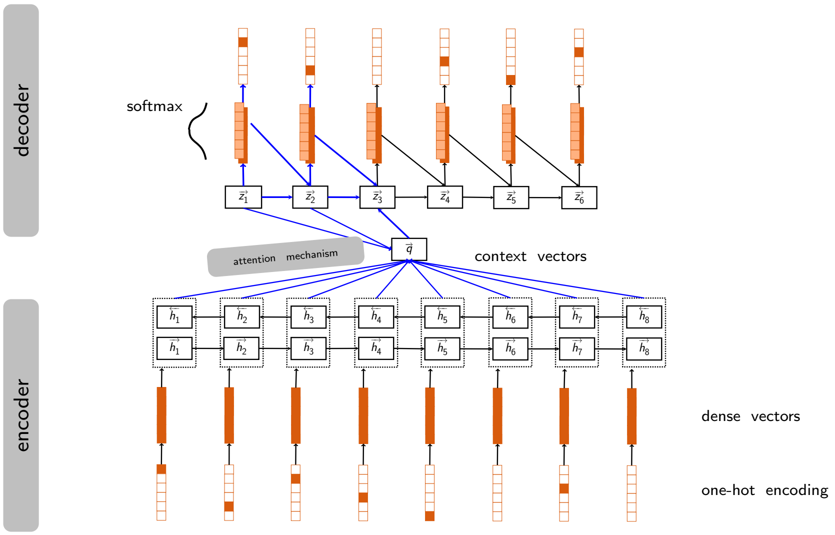
Back-up slides

Multilinguality in NMT



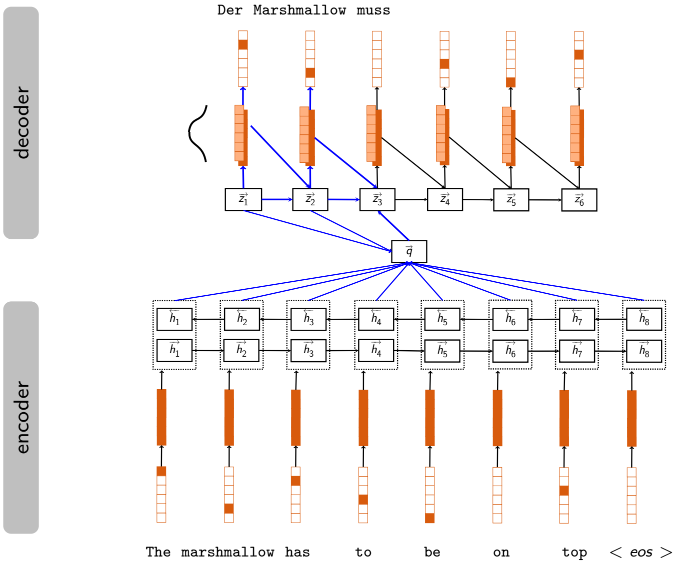
Back-up slides

Multilinguality in NMT



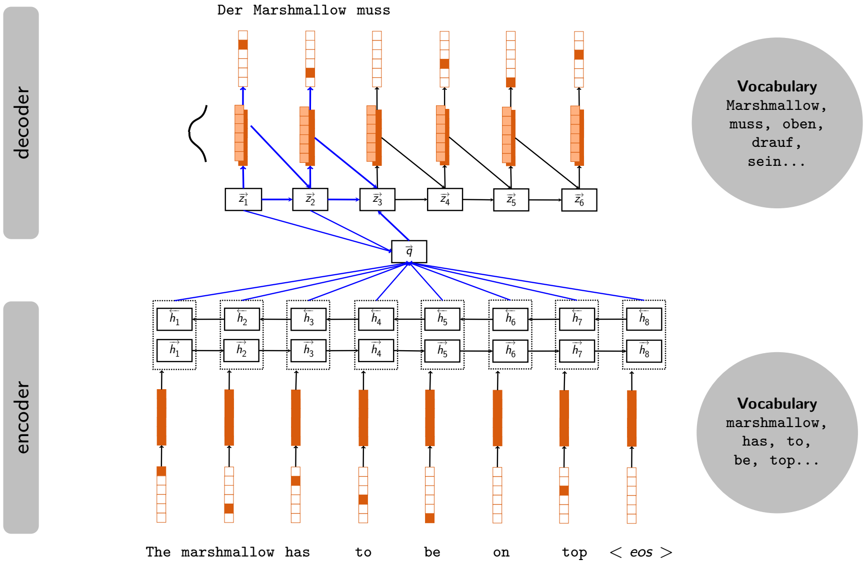
Back-up slides

Multilinguality in NMT



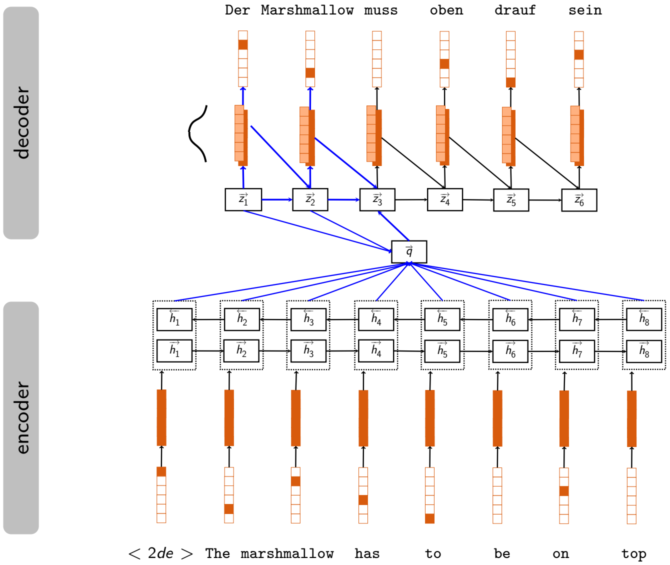
Back-up slides

Multilinguality in NMT



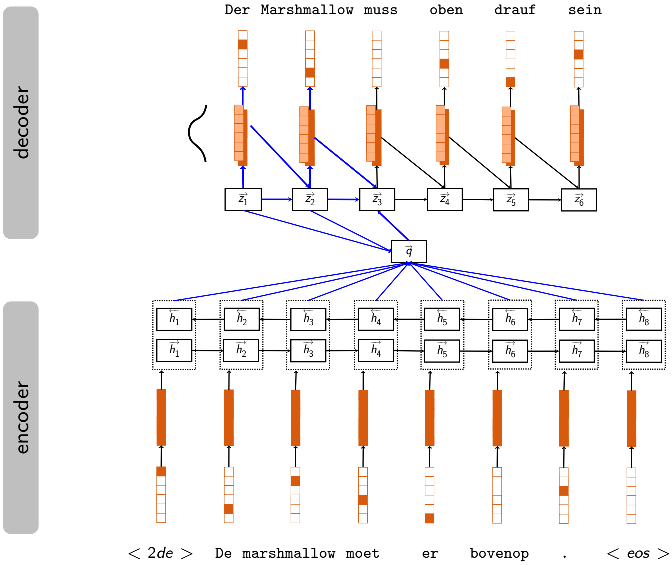
Back-up slides

Multilinguality in NMT



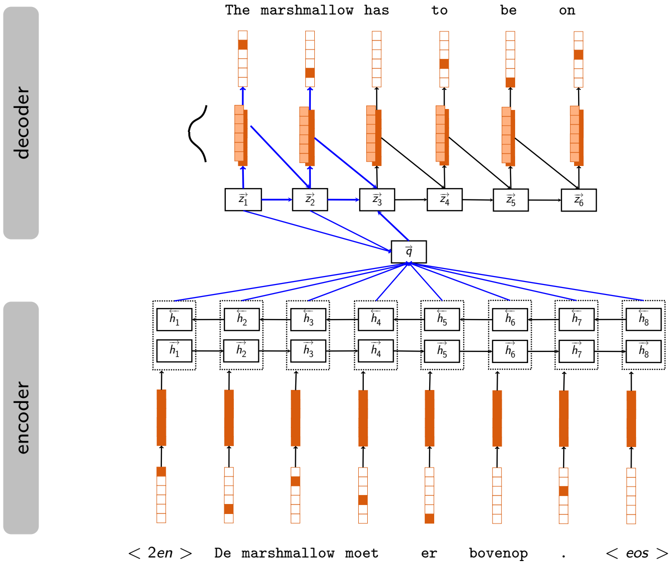
Back-up slides

Multilinguality in NMT



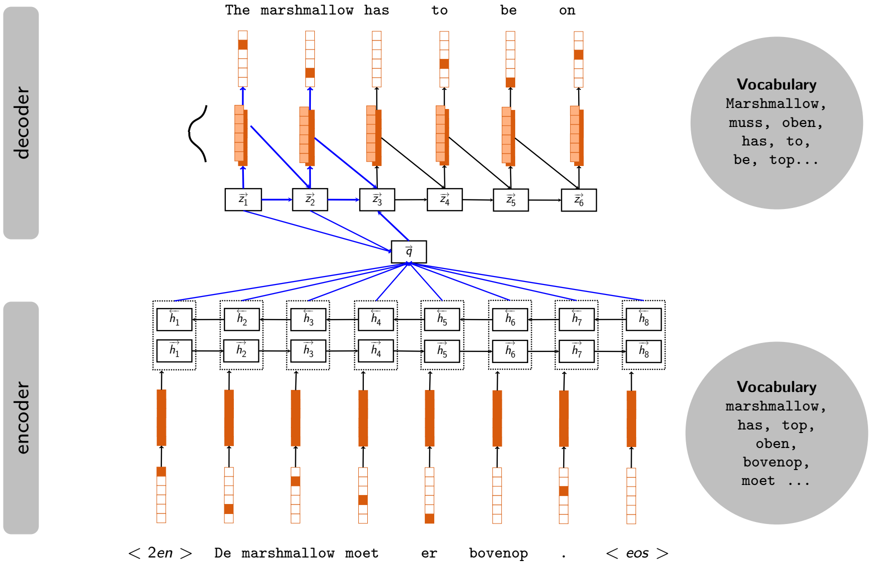
Back-up slides

Multilinguality in NMT



Back-up slides

Multilinguality in NMT



Multilinguality with Neural Networks

let's go back