Automatische Verarbeitung natürlicher Sprache
Eine Schlüsseltechnologie der Künstlichen Intelligenz

Sebastian Stüker
Importance of Language

- 5,000 to 7,000 languages exist in the world today
- “Language is the pinnacle achievement of man-kind” (Vivianne Redding, European Commission)
- Language is inherently linked with culture
- The way we speak / the way our language is structured influences the way we think (e.g., see research on cognition by Lera Boroditsky at UCLA)
- Language Diversity as important for human well-being and prosperity, as biodiversity in nature
- Language is deeply intertwined with human thinking, i.e. intelligence
Using the System

https://lecture-translator.kit.edu
Natural Language Processing

Wikipedia:

Natural language processing (NLP) is a field of computer science, artificial intelligence, and linguistics concerned with the interactions between computers and human (natural) languages. As such, NLP is related to the area of human–computer interaction.

Joint work of computer science and linguistics

Natural communication with computers

- Computers “understand”/process natural language
- Acquire knowledge from texts
- Generate text
- Overlap with Human-computer interaction/Dialog modeling
Natural Language Processing

level

beyond sentence
sentence
(ling.) phrases / chunks / n-grams
word

linguistic dimension

summarization
sentiment analysis
Co-reference resolution

parsing
NER
POS tagging

syntax

semantics

WSD

pragmatics

QA

dialog

NLU

ASR MT

summarization
sentiment analysis
Co-reference resolution

parsing
NER
POS tagging

syntax

semantics

WSD

pragmatics

QA

dialog

NLU

ASR MT

summarization
sentiment analysis
Co-reference resolution

parsing
NER
POS tagging

syntax

semantics

WSD

pragmatics

QA

dialog

NLU

ASR MT

summarization
sentiment analysis
Co-reference resolution

parsing
NER
POS tagging

syntax

semantics

WSD

pragmatics

QA

dialog

NLU

ASR MT

summarization
sentiment analysis
Co-reference resolution

parsing
NER
POS tagging

syntax

semantics

WSD

pragmatics

QA

dialog

NLU

ASR MT

summarization
sentiment analysis
Co-reference resolution

parsing
NER
POS tagging

syntax

semantics

WSD

pragmatics

QA

dialog

NLU

ASR MT

summarization
sentiment analysis
Co-reference resolution

parsing
NER
POS tagging

syntax

semantics

WSD

pragmatics

QA

dialog

NLU

ASR MT

summarization
sentiment analysis
Co-reference resolution

parsing
NER
POS tagging

syntax

semantics

WSD

pragmatics

QA

dialog

NLU

ASR MT

summarization
sentiment analysis
Co-reference resolution

parsing
NER
POS tagging

syntax

semantics

WSD

pragmatics

QA

dialog

NLU

ASR MT

summarization
sentiment analysis
Co-reference resolution

parsing
NER
POS tagging

syntax

semantics

WSD

pragmatics

QA

dialog

NLU

ASR MT

summarization
sentiment analysis
Co-reference resolution

parsing
NER
POS tagging

syntax

semantics

WSD

pragmatics

QA

dialog

NLU

ASR MT
Corpus-based Approaches to NLP - Overview

Output

Statistical Model

Word: 546; Case: Upper-case; POS: NN; …

Feature Extraction

Input Text:
Deep learning Approaches to NLP

Problems of simple statistical models

Feature engineering
- What features are important to determine the class
- Word ending
- Surrounding words
- Capitalization

Deep learning:
- Use neural networks to automatically infer features
- Better generalization
- Successfully applied to many NLP tasks
Deep Learning Approaches to NLP - Overview

Output

Statistical Model

10010010010101

Feature Extraction

Input Text:
Why is NLP hard?

Ambiguities
Ambiguities
Ambiguïtes

→ due to semantics and/or syntactic level...
Ambiguities - Examples

“I saw the man on the hill with a telescope.”

I saw the man. The man was on the hill. I was using a telescope.
I saw the man. I was on the hill. I was using a telescope.
I saw the man. The man was on the hill. The hill had a telescope.
I saw the man. I was on the hill. The hill had a telescope.
I saw the man. The man was on the hill. I saw him using a telescope.
Ambiguities - Examples

- Computer Science Lecture
  - Das Vorzeichen der Zahl

- Als Datenstruktur wird ein Keller-Automat benutzt
Ambiguities - Examples

TED
Fisch zuchtfarm anlage [fish farm sound system instead of facility]
Ambiguities - Jokes

One morning I shot an elephant in my pajamas. How he got into my pajamas, I’ll never know.
Policeman to little boy: “We are looking for a thief with a bicycle.”
Little boy: “Wouldn’t you be better using your eyes.”
Why is the teacher wearing sunglasses. Because the class is so bright.
Examples

- Lecture Translator
- Natural Language Understanding
- Summarization
- Backchannel Prediction
Institut für Anthropomatik

Simultane Vorlesungsübersetzung

Dr. Sebastian Stüker

Institut für Anthropomatik

28.05.18
Components

ASR

Segmentation

MT

... where were they ...

... Where were they? ...

... Wo waren sie ...
System Overview

Components

- ASR
- MT

Services

- Service over the network

Lecture

- Lecture 1
- Lecture 2
- Lecture 3

Recording of lectures
In the lecture hall
Web server for results

Include latest research results
Natural language understanding – Slot Filling

- Sequence labeling task
- Assign semantic class label to every word

<table>
<thead>
<tr>
<th>Sentence</th>
<th>show</th>
<th>flights</th>
<th>from</th>
<th>Boston</th>
<th>To</th>
<th>New</th>
<th>York</th>
<th>today</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slots/Concepts</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>B-dept</td>
<td>O</td>
<td>B-arr</td>
<td>I-arr</td>
<td>B-date</td>
</tr>
<tr>
<td>Named Entity</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>B-city</td>
<td>O</td>
<td>B-city</td>
<td>I-city</td>
<td>O</td>
</tr>
<tr>
<td>Intent</td>
<td>Find Flight</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Domain</td>
<td>Airline Travel</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
RNN-based Slot filling

Input: Words in one-hot encoding
Output: Label probabilities

Show flights from Boston to New York today
Summarization:

- Reduce natural language text document
- Goal:
  - Compress text by extracting the most important/relevant parts

Information overload

- High quantity of data
- Processing summary only
Applications

- Articles, news:
  - Outlines or abstracts
- Email / Email threads
- Health information
- Meeting summarization
- ...
Techniques

Extraction

Select subset of existing text segments:
E.g.:
  Sentence extraction
  Key-phrase extraction

Simpler, most focus in research in the past

Abstraction

“Understand” text
Use natural language generation to create summary
More human like
Latest trend in research
Techniques

Fourscore and seven years ago our fathers brought forth on this continent a new nation, conceived in liberty and dedicated to the proposition that all men are created equal. Now we are engaged in a great civil war, testing whether that nation or any nation so conceived and so dedicated can long endure. We are met on a great battlefield of that war. We have come to dedicate a portion of that field as a final resting-place for those who here gave their lives that that nation might live. It is altogether fitting and proper that we should do this. But in a larger sense, we cannot dedicate, we cannot consecrate, we cannot hallow this ground. The brave men, living and dead who struggled here have consecrated it far above our poor power to add or detract. The world will little note nor long remember what we say here, but it can never forget what they did here. It is for us the living rather to be dedicated here to the unfinished work which they who fought here have thus far so nobly advanced. It is rather for us to be here dedicated to the great task remaining before us — that from these honored dead we take increased devotion to that cause for which they gave the last full measure of devotion — that we here highly resolve that these dead shall not have died in vain, that this nation under God shall have a new birth of freedom, and that government of the people, by the people, for the people shall not perish from the earth.

Figure 1.1: The Gettysburg Address
Four score and seven years ago our fathers brought forth upon this continent a new nation, conceived in liberty, and dedicated to the proposition that all men are created equal. Now we are engaged in a great civil war, testing whether that nation, or any nation so conceived and so dedicated, can long endure. The brave men, living and dead, who struggled here, have consecrated it far above our poor power to add or detract.

Figure 1.3: Another 25% extract of the Gettysburg Address

This speech by Abraham Lincoln commemorates soldiers who laid down their lives in the Battle of Gettysburg. It reminds the troops that it is the future of freedom in America that they are fighting for.

Figure 1.4: A 15% abstract of the Gettysburg Address
Extractive summarization

- Three main components
  - Content selection: Which parts are important to be in the summary?
  - Information ordering: How to order summaries?
  - Sentence realization: Clean up/Simplify sentences
Sentence Extraction

Use statistic heuristics to select sentences
Do not change content and meaning

Idea:

Use measure to determine importance of sentence

TF-IDF

Supervised trained combination of several features

Rank sentence according to metric

Output sentences with heights score:

Fixed number

All sentence above threshold
Attention-based Encoder -Decoder

Encoder

Source

Sentence representation

Target

Decoder

State

Context vector

States
Attention-based Encoder-Decoder

Good quality summary output
S: a man charged with the murder last year of a british backpacker confessed to the slaying on the night he was charged with her killing, according to police evidence presented at a court hearing tuesday. ian douglas previte, ##, is charged with murdering caroline stuttle, ##, of yorkshire, england.
T: man charged with british backpacker’s death confessed to crime police officer claims
O: man charged with murdering british backpacker confessed to murder

S: following are the leading scorers in the english premier league after saturday’s matches: ## - alan shearer -lrb- newcastle united -rrb-, james beattie.
T: leading scorers in english premier league
O: english premier league leading scorers

S: volume of transactions at the nigerian stock exchange has continued its decline since last week, a nse official said thursday. the latest statistics showed that a total of ##.### million shares valued at ##.### million naira -lrb- about ##.### million us dollars -rrb- were traded on wednesday in ##.### deals.
T: transactions dip at nigerian stock exchange
O: transactions at nigerian stock exchange down
Attention-based Encoder - Decoder

Poor quality summary output

S: broccoli and broccoli sprouts contain a chemical that kills the bacteria responsible for most stomach cancer, say researchers, confirming the dietary advice that moms have been handing out for years. In laboratory tests the chemical, _<unk>_, killed Helicobacter pylori, a bacteria that causes stomach ulcers and often fatal stomach cancers.

T: for release at ### _<unk>_ mom was right broccoli is good for you say cancer researchers

O: broccoli sprouts contain deadly bacteria

S: norway delivered a diplomatic protest to russia on monday after three norwegian fisheries research expeditions were barred from russian waters. the norwegian research ships were to continue an annual program of charting fish resources shared by the two countries in the barents sea region.

T: norway protests russia barring fisheries research ships

O: norway grants diplomatic protest to russia

S: j.p. morgan chase 's ability to recover from a slew of recent losses rests largely in the hands of two men, who are both looking to restore tarnished reputations and may be considered for the top job someday. geoffrey _<unk>_, now the co-head of j.p. morgan 's investment bank, left goldman, sachs & co. more than a decade ago after executives say he lost out in a bid to lead that firm.

T: # executives to lead j.p. morgan chase on road to recovery

O: j.p. morgan chase may be considered for top job
Backchannel Prediction

Acoustic cues for the speaker
Backchannel responses used by listener in social dialog
Provide feedback to speaker (positive, negative, neutral)
Establish relationship

Artificial assistants are becoming increasingly popular
Still distinctive non-human

Make human-computer interaction more social
Neural network based approach using LSTMs
User study for evaluation

28.05.18
Post processing

Smooth output by applying low-pass filter
Use threshold to detect BC

else i tend to turn on the television at eleven o'clock just to watch th

Raw net output

Smoothed

Audio output

Threshold

Trigger

3.4. Postprocessing

Our goal is to generate an artificial audio track containing utterances such as “uh-huh” or “yeah” at appropriate times. The placement of BCs is dependent on previous BCs: If the previous BC utterance in the transcript of the listener channel as an anchor the BC utterance in the speaker channel as a trigger, we cause a BC in the listener track. We chose the beginning of the BC happening shortly is higher and vice versa. After each BC, the probability of a new BC rises over time. To accommodate backchannel prediction.

We decided not to give a BC response yet, so it is sensible to assume intuitively meaningful, because in that area the listener explicitly us an evenly balanced data set, and the negative samples are chosen areas to predict zero i.e. “no BC”. We choose the range as defined above, we also need to train the network to first learn to align its inputs. In addition to selecting the anchor the BC utterance in the transcript of the listener channel as an anchor, we also need to take context into account. We do this by using Long-term memory layers (LSTM) instead of dense feed forward networks to make the network to consider the internal state or outputs into account.

The placement of BCs is dependent on previous BCs: If the previous BC utterance in the transcript of the listener channel as an anchor the BC utterance in the speaker channel as a trigger, we cause a BC in the listener track. We chose the beginning of the BC happening shortly is higher and vice versa. After each BC, the probability of a new BC rises over time. To accommodate backchannel prediction.

We decided not to give a BC response yet, so it is sensible to assume intuitively meaningful, because in that area the listener explicitly us an evenly balanced data set, and the negative samples are chosen areas to predict zero i.e. “no BC”. We choose the range as defined above, we also need to train the network to first learn to align its inputs. In addition to selecting the anchor the BC utterance in the transcript of the listener channel as an anchor, we also need to take context into account. We do this by using Long-term memory layers (LSTM) instead of dense feed forward networks to make the network to consider the internal state or outputs into account.

3.3. Neural Network Design and Training

In 3.2. Training Area Selection, we can see that as the audio range dimension increases, the probability of a BC rises over time. To accommodate backchannel prediction.

In 3.2. Training Area Selection, we can see that as the audio range dimension increases, the probability of a BC rises over time. To accommodate backchannel prediction.

The placement of BCs is dependent on previous BCs: If the previous BC utterance in the transcript of the listener channel as an anchor the BC utterance in the speaker channel as a trigger, we cause a BC in the listener track. We chose the beginning of the BC happening shortly is higher and vice versa. After each BC, the probability of a new BC rises over time. To accommodate backchannel prediction.

We decided not to give a BC response yet, so it is sensible to assume intuitively meaningful, because in that area the listener explicitly us an evenly balanced data set, and the negative samples are chosen areas to predict zero i.e. “no BC”. We choose the range as defined above, we also need to train the network to first learn to align its inputs. In addition to selecting the anchor the BC utterance in the transcript of the listener channel as an anchor, we also need to take context into account. We do this by using Long-term memory layers (LSTM) instead of dense feed forward networks to make the network to consider the internal state or outputs into account.

3.2. Training Area Selection

3.5. Evaluation

We determined the optimal postprocess-
Example Video

[note from the newspapers and the trash and th[is]- this bottled [vocalized-noise] stu one for the regular tr]
Conclusion

Natural language processing a key technology in artificial intelligence

Large number of diverse applications, e.g.
  - Speech translation
  - Natural language understanding
  - Backchannel prediction

Many solutions nowadays use artificial neural networks

Understanding natural language processing will give insights into the making of human intelligence