Will computers ever understand us?

Zuzana Nevěřilová

Natural Language Processing Centre, Faculty of Informatics
Masaryk University, Brno, Czech Republic

xpopelk@fi.muni.cz

Workshop of Natural Language Processing Centre
May 16, 2014

Will computers ever understand us?

e may hope that machines will eventually compete
n all purely intellectual fields. But which are the best ones to
tart with? Even this is a difficult decision. Many people think
that a very abstract activity, like the playing of chess, would be
best.

It can also be maintained that it is best to provide the **machine**
with the best sense organs that money can buy, and then teach
it to **understand and speak English**. This process could follow
the normal teaching of a child. Things would be pointed out
and named, etc. Again I do not know what the right answer is,
but I think both approaches should be tried.

We can only see a short distance ahead, but we can see plenty
to be done.

Alan Turing

Computer "understanding": Use cases

- inappropriate discussion posts detection
- text summarization
- sentiment analysis
- question answering
- plagiarism detection

That BOY aint done growing and fcuking so she would be stooopid to tie HERSELF down wit a BABY and a tattoo is just as worse!!!

That BOY aint done growing and fcuking so she would be stooopid to tie HERSELF down wit a BABY and a tattoo is just as worse!!!

Inappropriate discussion posts detection

Use case: discussion forum, automatic detection of inappropriate posts

Common solution: word list

But: users use concealed words that are difficult to detect (f*king, f.u.c.k, f..k, fcuking)

Better solution: word list + concealing rules

But: users invent new words and concealing patterns

Even better solution: word list + automatically generated thesaurus + concealing rules + metarules

fcuking *(adjective)*

Lemma	Score	Frequency
phucking	0.618	67
out-an-out	0.456	22
god-damned	0.445	357
self-deceiving	0.36	133
frigging	0.21	1296
sleazebags	0.181	93
whingeing	0.162	105
bold-faced	0.158	471
complusive	0.155	69
bald-faced	0.137	870
self-deluded	0.106	401
blithering	0.105	844
bare-faced	0.1	343
barefaced	0.097	740
cking	0.089	2365

Use case: discussion forum, automatic detection of inappropriate posts

Common solution: word list

But: users use concealed words that are difficult to detect (f*king, f.u.c.k, f..k, fcuking)

Better solution: word list + concealing rules

But: users invent new words and concealing patterns

Even better solution: word list + automatically generated thesaurus + concealing rules + metarules

Sentiment analysis

The **iPhone 5** price was predictably high and continues to be so, so consumers will need to bear that in mind too when looking for their next smartphone.

. . .

Well, all of those picking up the iPhone 5 will have the same reaction: this thing is amazingly light. You've probably heard the numbers by now (20 per cent lighter than the predecessor, as well as beating most of the opposition too at 112g.)

Sentiment analysis

The **iPhone 5** price was predictably high and continues to be so, so consumers will need to bear that in mind too when looking for their next smartphone.

. . .

Well, all of those picking up the iPhone 5 will have the same reaction: this thing is amazingly light. You've probably heard the numbers by now (20 per cent lighter than the predecessor, as well as beating most of the opposition too at 112g.)
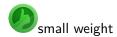
 high price

 small weight

Sentiment analysis

Use case: what are people thinking about a particular product/company/idea X?

Solution: search X, find evaluative words

But: opinions are expressed by non-evaluative words

Better solution:

- extract useful attributes of X (noise, weight, price, appearance)
- generate thesauri of evaluative words:
  low price × low quality
  thin iPhone 5 × thin tasteless burger

Do you have a bike for a 4–year-old girl?
Search results for "bike", "girl"

· · ·

Do you have a bike for a 4–year-old girl?
If she is under 110 cm tall I will recommend Maggie, Princess or Misty. If she is taller I would recommend Miss B or Kellie. If she does not insist on bike for girls I would also recommend Racer or Mr. Lightning. How tall is she?
About 105 cm.
Do you have some other constraints?
I look for something cheaper.
Then I would recommend Princess. It is a popular bike.

Use case: chatbot providing basic support

Solution: patterns, keyword detection, searching

But: no real dialogue, no real answers, just searching

Better solution: sentence structure analysis, keyword detection, coreference resolution, dialogue strategy

bike girl kellie maggie missb misty mrlightning **princess** racer → **too many products** → color **price** size **weight**

Is this real understanding?

Probably not.

We do not know what understanding is but we know how it looks like when someone understands.

Computer programs that can discover a vulgar text, summarize a text, recognize someone's opinions or answer questions **look like** they understand our language . . . .

Is this real understanding?

Probably not.

We do not know what understanding is but we know how it looks like when someone understands.

Computer programs that can discover a vulgar text, summarize a text, recognize someone's opinions or answer questions **look like** they understand our language . . . .

Is this real understanding?

Probably not.

We do not know what understanding is but we know how it looks like when someone understands.

Computer programs that can discover a vulgar text, summarize a text, recognize someone's opinions or answer questions **look like** they understand our language . . . .

Is this real understanding?

Probably not.

We do not know what understanding is but we know how it looks like when someone understands.

Computer programs that can discover a vulgar text, summarize a text, recognize someone's opinions or answer questions **look like** they understand our language . . . .