Computational Humor

Can a machine have a sense of humor?

Thomas Winters

PhD Student at KU Leuven & FWO Fellow

@thomas_wint

thomaswinters.be

Can a machine have a sense of humor?

Part 1: Humor

Intrinsically human!

Every known human civilisation has some form of humor



Caron, J.E.: From ethology to aesthetics: Evolution as a theoretical paradigm for research on laughter, humor, and other comic phenomena. Humor15(3), 245–281(2002)

Purpose of Humor = Sign of Intelligence?



Brain is rewarding **noticing incongruities** and successfully **resolving them**

\+ **linguistic skills**
(required to create jokes)

= **Evolutionary advantage!**

How does humor work?

Incongruity-Resolution Theory

Setup *Two fish are in a tank. Says one to the other:*

Punchline *"Do you know how to drive this thing?"*

Obvious Interpretation



Hidden Interpretation



Based on: Ritchie, G. (1999). Developing the incongruity-resolution theory.

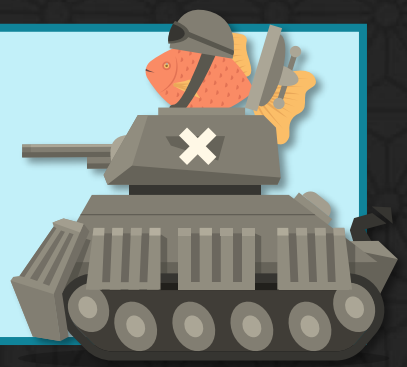
Human-focused definition!

Machine should **not only spot two mental images**

Obvious Interpretation



Hidden Interpretation



But also this is **not too hard or too easy** for a **human**!

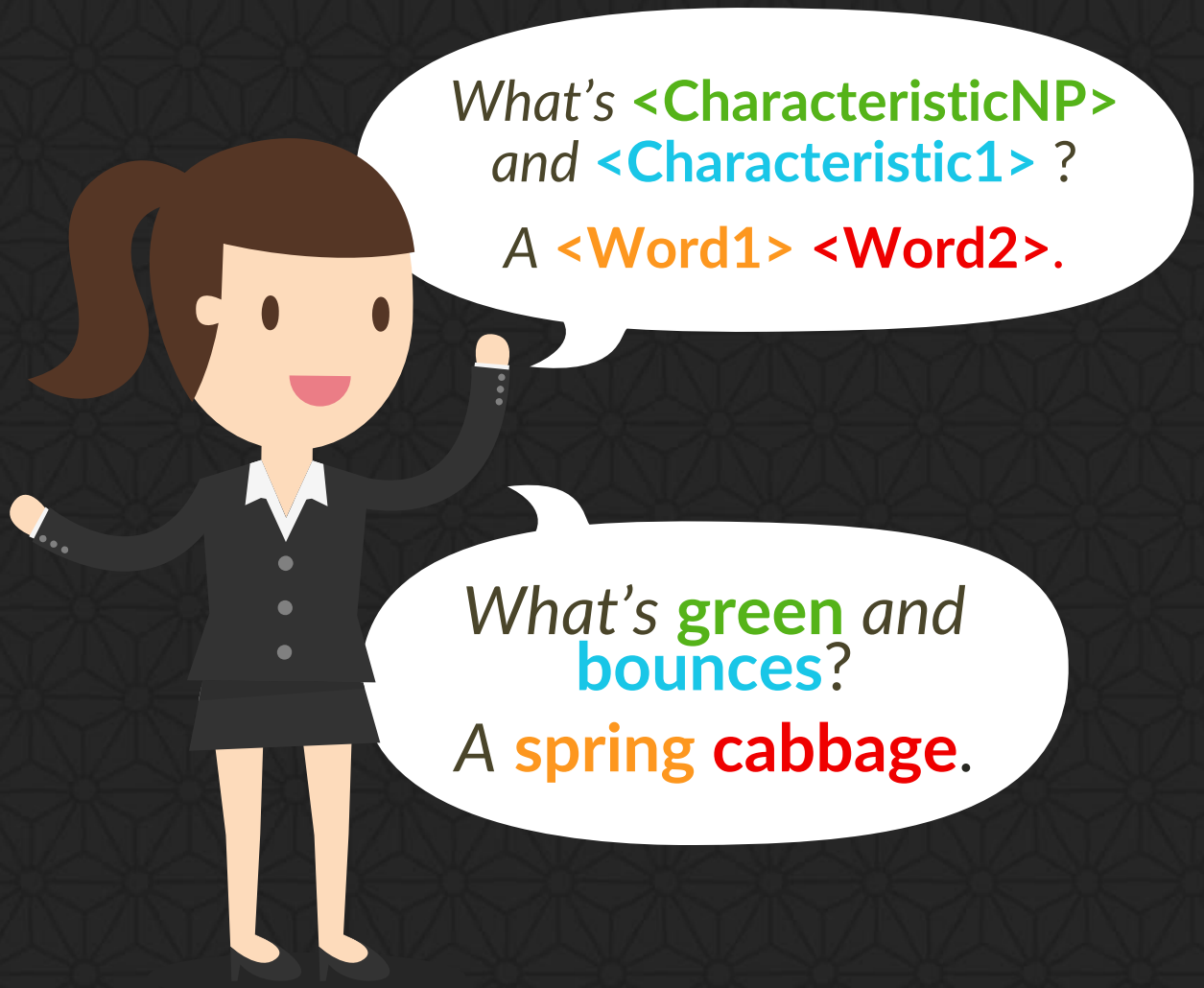
Humor = AI-complete?

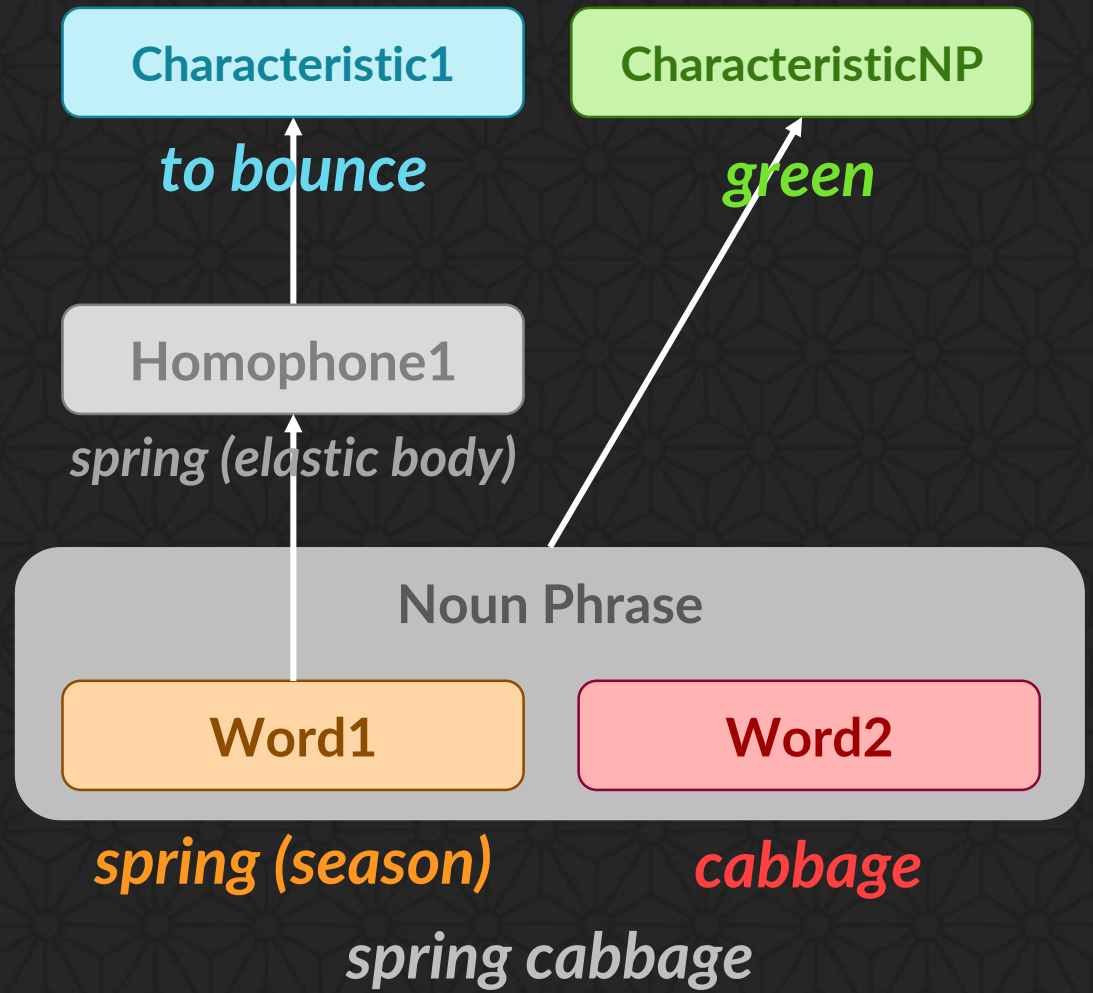
~ as hard as making computers as “intelligent” as people

Part 2: Humor Generation

Hard-coding a sense of humor

JAPE: Templates & Schemas





JAPE Joke Examples

What is the difference between leaves and a car?
One you brush and rake, the other you rush and brake

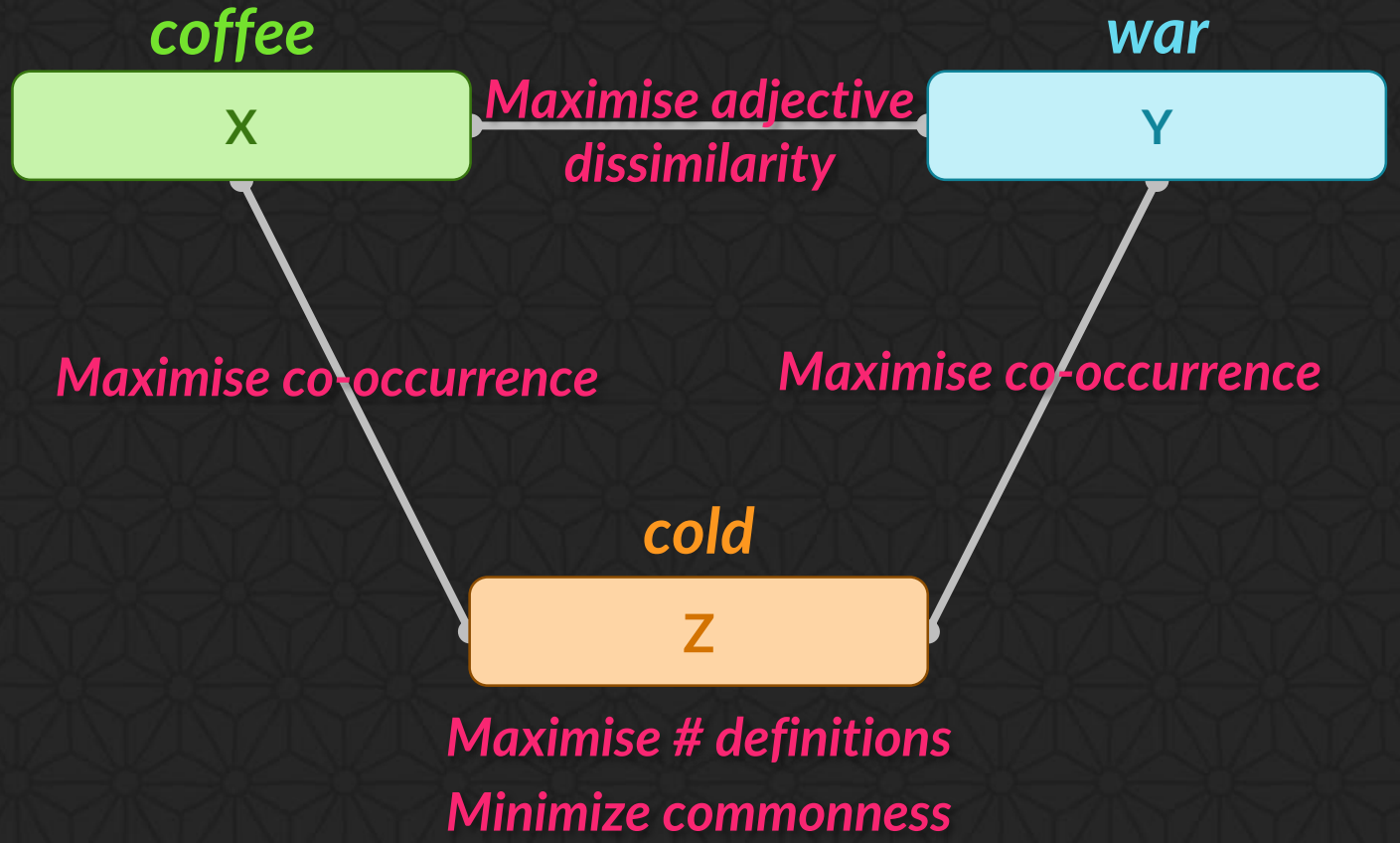
What do you call a strange market?
A bizarre bazaar.



https://www.gwern.net/GPT-3
https://towardsdatascience.com/teaching-gpt-2-a-sense-of-humor-fine-tuning-large-transformer-models-on-a-single-gpu-in-pytorch-59e8cec40912

Unsupervised Analogy Generator

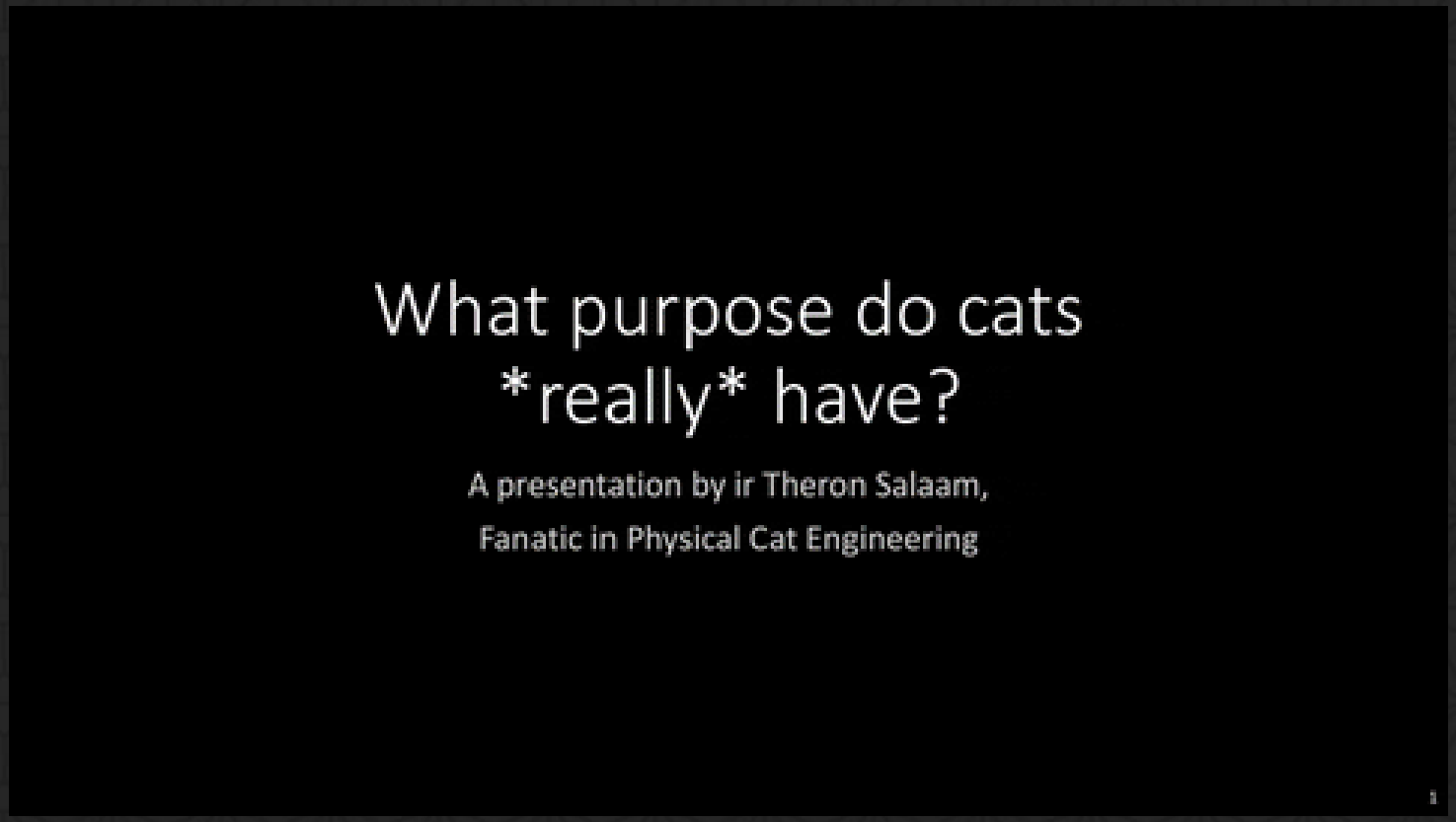




Petrović, S., & Matthews, D. (2013). Unsupervised joke generation from big data

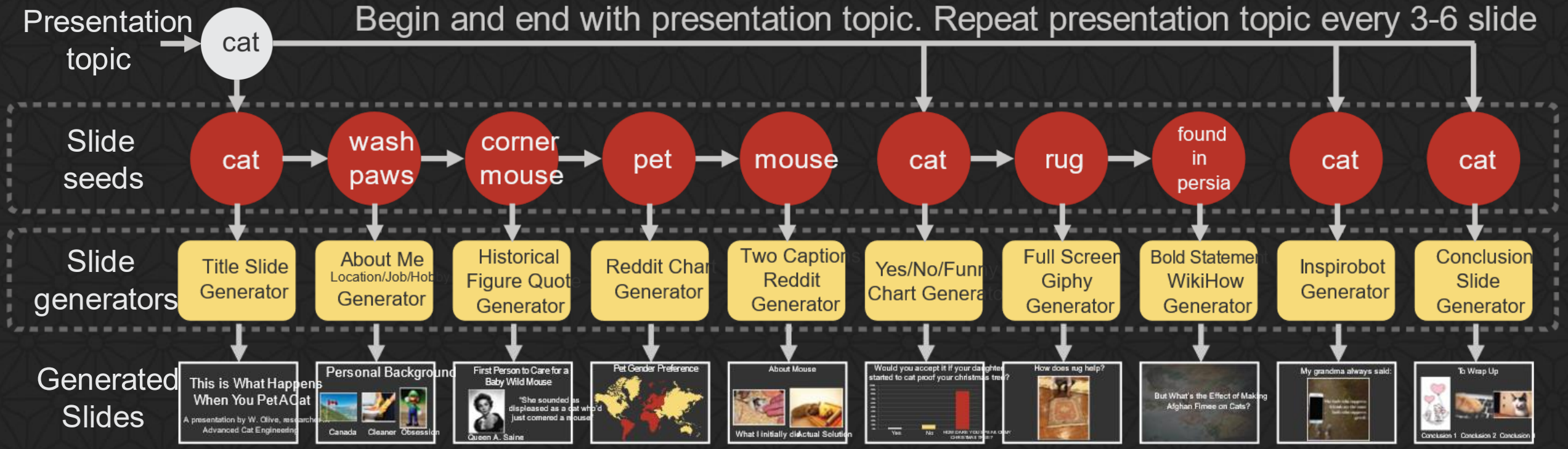
Talk Generator

Generates nonsense PowerPoints about any given topic



Try it yourself: ***talkgenerator.com***

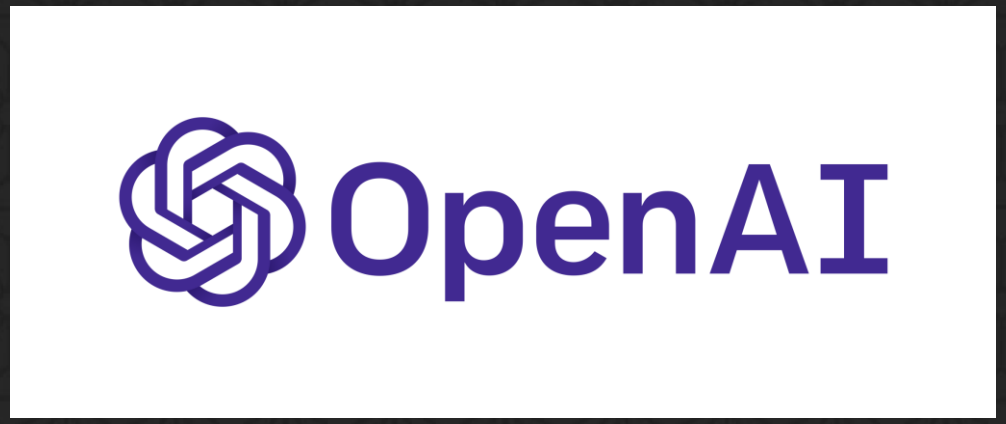
Talk Generator



Learning a sense of humor?

Transformer models

Large **language models**, **pretrained** on large corpora





GPT-2 & GPT-3

Completes any textual prompt

BERT

Classifies any text sequence / token

Outperforming previous neural architectures
on most language tasks

Brown, Tom B., et al. "Language models are few-shot learners."
Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding

Not just for English!



RobBERT

Our Dutch BERT-like model

Phase 1: Pretraining:

- Find unlabeled **corpus** for target language
- **Expensive** pretraining phase
- BUT: **reusable**! Just download existing!

Phase 2: Finetuning a pre-trained BERT

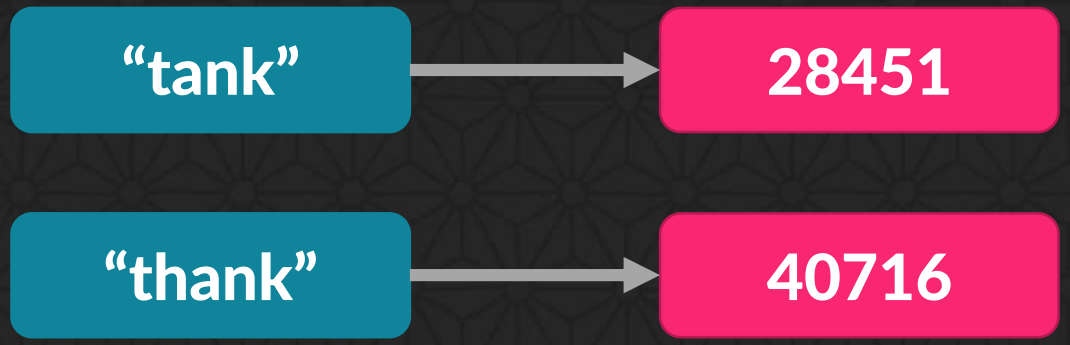
- Feed **labeled** training data
- Relatively **cheap** fine-tuning for the target task
- Usually **outperforms** other language models on most NLP tasks

Delobelle, P., Winters, T., & Berendt, B. (2020). RobBERT: a dutch RoBERTa-based language model.

GPT-2 / GPT-3 humor?

Tokenizer problem: maps words to numbers

→ **Remove** knowledge about **letters**!



Hard to make/recognize new **wordplay**!

https://www.gwern.net/GPT-3

https://towardsdatascience.com/teaching-gpt-2-a-sense-of-humor-fine-tuning-large-transformer-models-on-a-single-gpu-in-pytorch-59e8cec40912

GPT-2 / GPT-3 joke examples

GPT-2 can mimic **joke style**, but usually **very absurd**

What did one plate say to the other plate?
Dip me!

What did the chicken say after he got hit by a bus?
"I'm gonna be fine!"

A woman walks into the bar.....she says to the bartender "I want a double entendre" and the bartender says "No, we don't serve that"



https://www.gwern.net/GPT-3

https://towardsdatascience.com/teaching-gpt-2-a-sense-of-humor-fine-tuning-large-transformer-models-on-a-single-gpu-in-pytorch-59e8cec40912

Reminiscent of how children write jokes

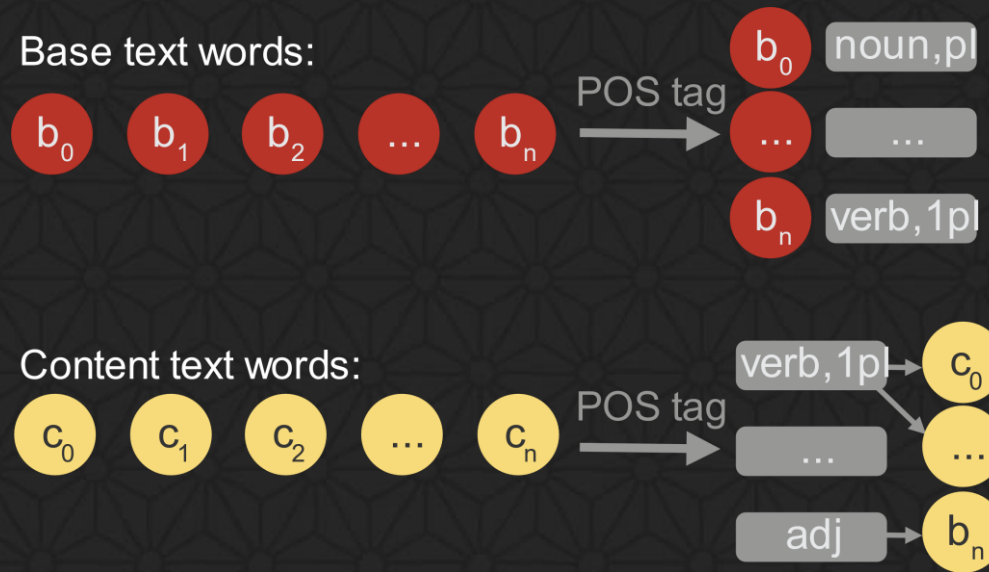




Let’s try simpler & more understandable models!

Dynamic Templates

"Are there also atheists that don't believe in atheism?"



"They see the fact that the former God is not trying to deny this newly acquired insight as proof of them being right. Even with the Church, things are not going well. Norse popes."

Replace some matching words with matching POS tag with lowest frequencies

b0 c2 b2 ... c0

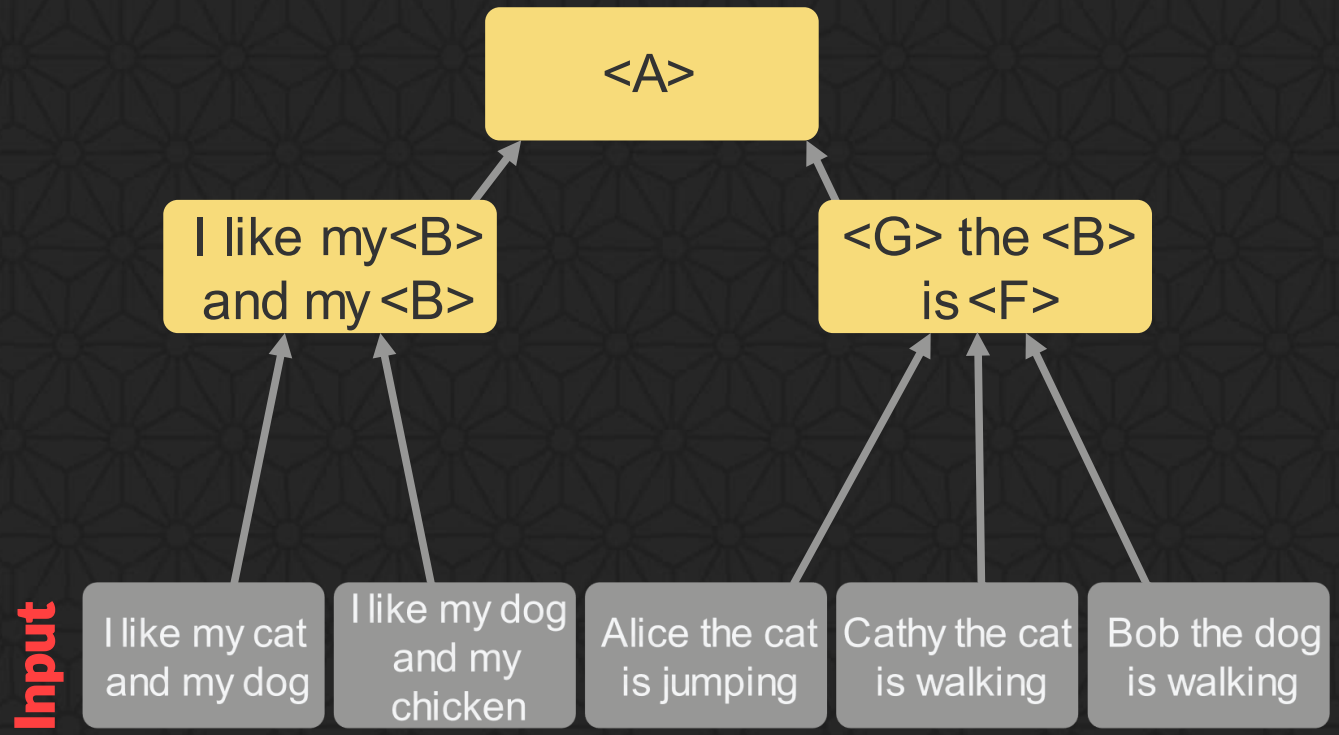
"Are there also popes that don't believe in God?"

Used in @TorfsBot



GITTA: Template Trees for extracting templates

1. Join **closest** strings
2. **Merge similar** template slots
3. Iteratively **simplify until convergence**



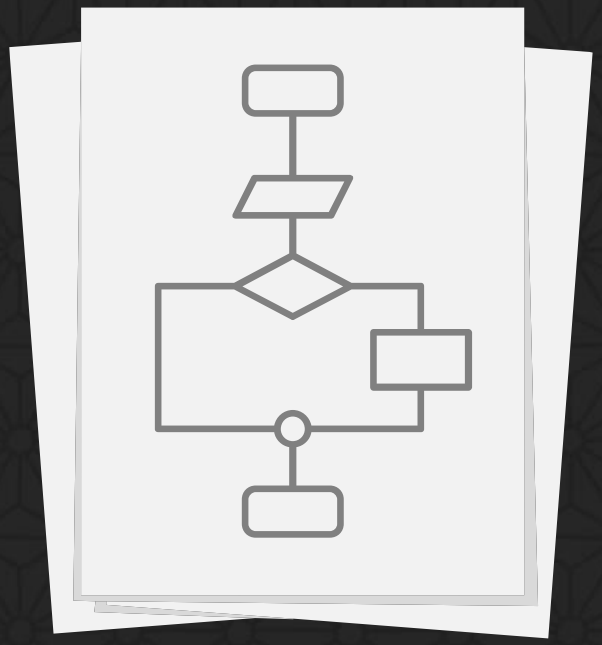
A: I like my **<B>** and my **<B>** | **<G>** the **<B>** is **<F>**

B: chicken | cat | dog

F: walking | jumping

G: Alice | Bob | Cathy

Generalising Schemas to Enable Machine Learning



Constraint-based Schemas

Hard to tweak to preference

Header

Keywords

Template

Features





Aggregator

Values Generator



Metrical Schema

Allows for learning in aggregator component

GOOFER: Generalised framework

a) I like my relation… …n.
b) I like my coffee li…
c) The sailor bears …
d) I like my girlfrien…
e) The lord cultivat…

Joke	Rater 1	Rater 2	Rater 3
a	5	4	5
b	4	4	3
c	5	3	4
d	2	1	3
e	3	2	2

Template 1: "I like my X like I like my Y, Z."
relations, source, open | [5,4,5]
coffee, war, cold | [4,4,3]
girlfriend, estate, real | [2,1,3]

Template 2: ``The A B a C. D.''
sailor, bears, stress, Pier pressure | [5,3,4]
lord, cultivates, region, A baron land | [3,2,2]

X	Y	Z	Score	FreqX	FreqY	FreqZ	FreqXY	...
relations	source	open	[5,4,5]	3 831 210	4 050 904	4884757	0	...
coffee	war	cold	[4,4,3]	784 065	9 211 826	2 010 173	0	...
girlfriend	estate	real	[2,1,3]	75 392	998 669	5 284 057	0	....

1. Training

Input Jokes → **Human Evaluation** → *Rated Jokes* → **Template Extractor** → *Templates* → **Template Store** → *Set of rated template values per template* → **Values Generator** → *Rated template values with feature values (training data)* → **Metric-based Rater** → **Classifier/ Regression**

2. Generating

Input topic → **Template Store** → *Word + Template* → **Values Generator** → *Generated template values* → **Metric-based Rater** → *Template values + features* → **Classifier/ Regression** → *Filtered template values set* → **(Apply Template)** → *Jokes about seed*

"Men"

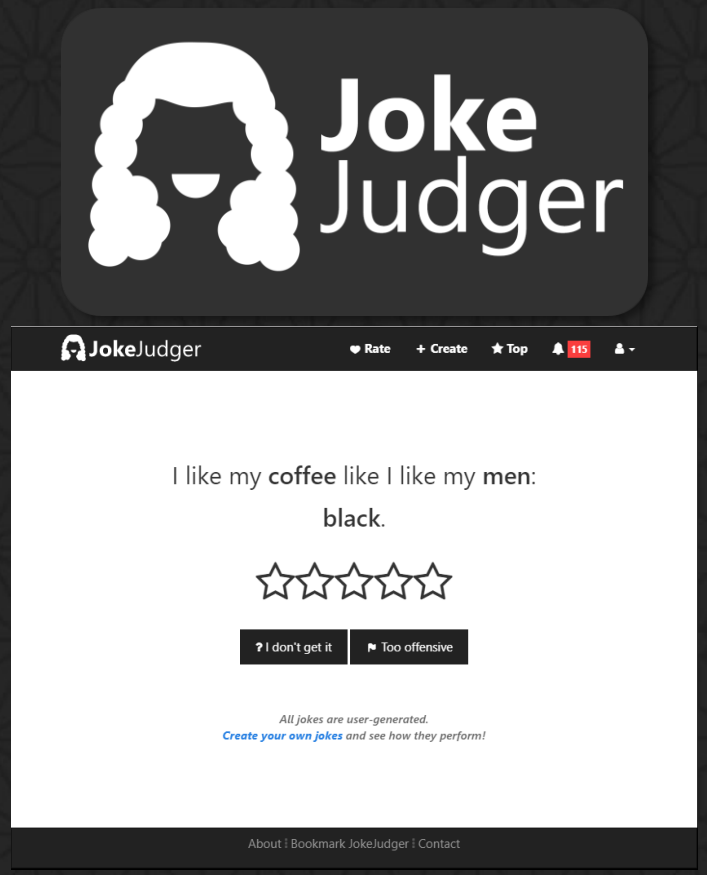
"Men" + *"I like my X like I like my Y, Z."*

a) men, grave, nameless
b) men, country, cold
c) men, turkey, roast
d) men, rain, gentle
e) men, laugh, silly

X	Y	Z	Score	FreqX	FreqY	FreqZ	FreqXY	...
men	grave	nameless	?	9 876 543	4 050 904	4884757	0	...
men	country	cold	?	9 876 543	9 211 826	2 010 173	0	...
men	turkey	roast	?	9 876 543	998 669	5 284 057	0	...
men	rain	gentle	?	9 876 543	9 876 543	845 321	0	...
men	laugh	silly	?	9 876 543	954 569	123 454	0	....

I like my men like I like my grave: nameless
I like my men like I like my rain: gentle

GAG: implements GOOFER framework

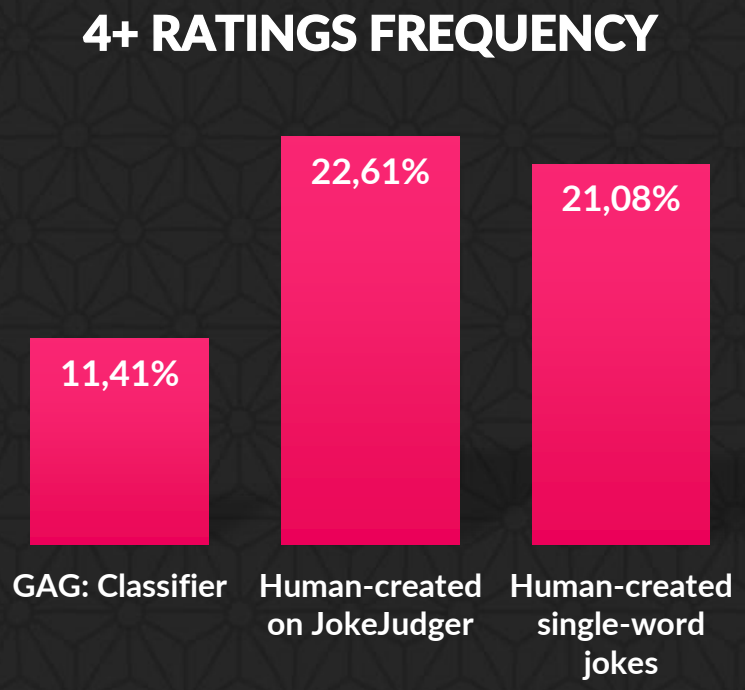


GAG
GENERALISED ANALOGY GENERATOR

I like my **coffee** like I like my **sleep**: **strong**.

I like my **men** like I like my **graves**: **nameless**.

I like my **women** like I like my **tests**: **independent**.



Winters, T., Nys, V., & De Schreye, D. (2018). Automatic Joke Generation: Learning Humor from Examples

Need for good humor detection algorithms!

Part 3: Humor Detection

Early Humor Detector

- Designed **humor features** *e.g. alliteration, antonym, adult slang…*
- Used **Naïve Bayes** and **Support Vector Machines**
- **Task:** One-liners vs **news**, **neutral corpus** & **proverbs**

One-liners Reuters	One-liners BNC	One-liners Proverbs
96.95%	79.15%	84.82%

Table 4: Humor-recognition accuracy for combined learning based on stylistic and content features.

Mihalcea, R., & Strapparava, C. (2005). Making computers laugh: Investigations in automatic humor recognition.

But is this a good dataset?

News & proverbs have **completely different types** of words than jokes!



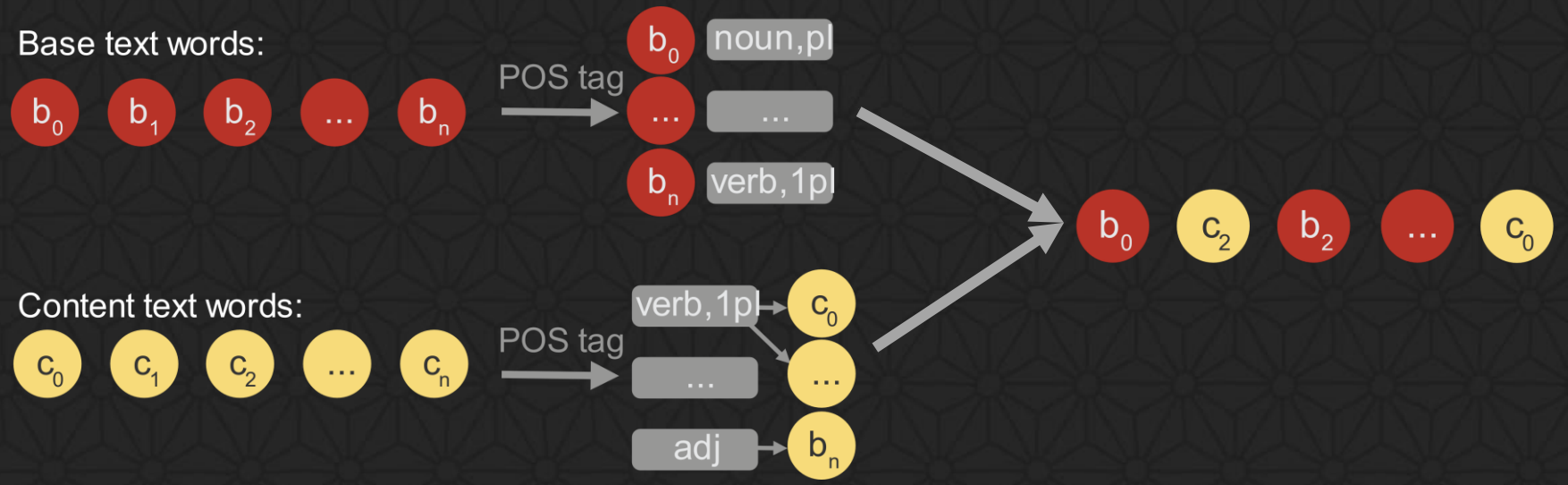


→ Looking at **word frequencies** is often already *"enough"*!

Is this really **humor** detection?

Jokes are **fragile**!

Two men are in a bar. Says one to the other:
"Do you know how to drive this thing?"



Generate non-jokes using **dynamic templates**!

Word-based features **won't work** anymore!

Much more challenging dataset!

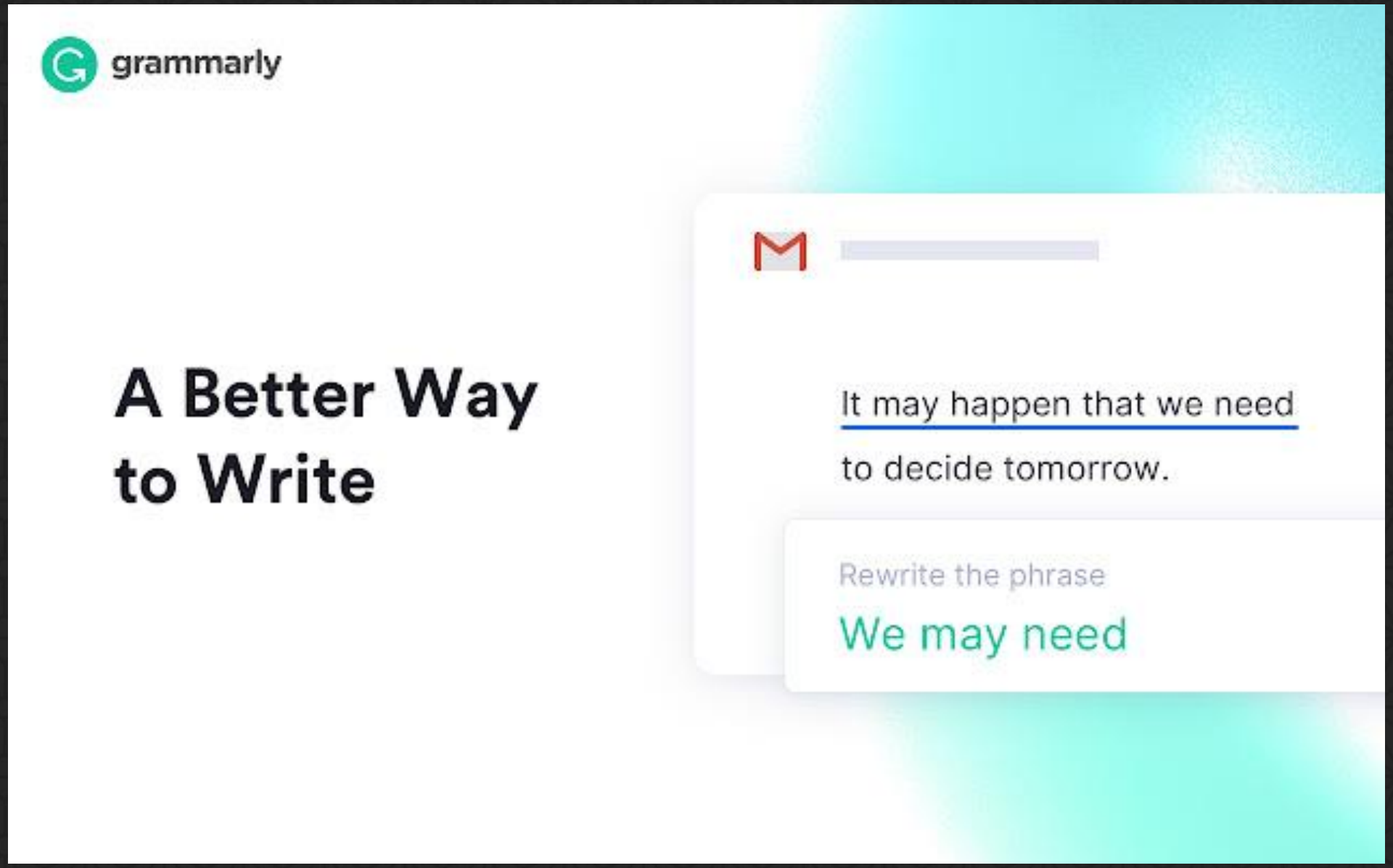
Table 1: Classification results for the held-out test set on three datasets versus the Jokes dataset. For the accuracy, we additionally report the 95% CI.

					Dynamic templates			
	News		Proverbs		Single		Pairwise	
Model	ACC [%]	F_1 [%]	ACC [%]	F_1 [%]	ACC [%]	F_1 [%]	ACC [%]	F_1 [%]
Naive Bayes	51.0 ± 3.1	49.3	60.2 ± 3.5	50.5	49.9 ± 3.1	49.9	-	-
LSTM	94.0 ± 1.5	94.0	94.4 ± 1.6	94.1	46.8 ± 3.1	35.2	47.9 ± 4.4	32.4
CNN	93.6 ± 1.5	93.6	94.1 ± 1.7	93.6	47.4 ± 3.1	46.3	58.6 ± 4.4	58.5
RobBERT	$\mathbf{98.8} \pm 0.7$	**98.8**	$\mathbf{99.6} \pm 0.4$	**99.6**	$\mathbf{89.2} \pm 1.9$	**89.1**	$\mathbf{82.51} \pm 3.4$	**82.5**

More truthful humor detection?

Part 3: Why?

Grammarly but for jokes?



VIRTUAL ASSISTANTS

People often treat virtual assistants like humans

("please", "thank you", "sorry") [1]

41%

say it's like talking to a friend [1]

Humor and creativity in language is important **between friends** [2]



[1] https://www.thinkwithgoogle.com/consumer-insights/voice-assistance-consumer-experience/

[2] Laughter's Influence on the Intimacy of Self-Disclosure, Gray, A.W., Parkinson, B. & Dunbar, R.I. Hum Nat (2015) 26: 28.

Casual Creation: it's fun to create these little bots!



https://thomaswinters.be/mopjesbot

Mutually interactive bots

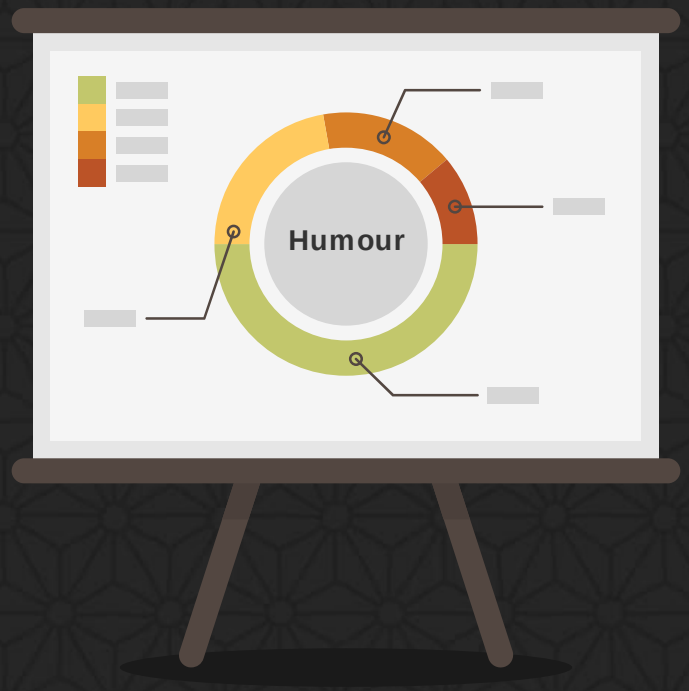


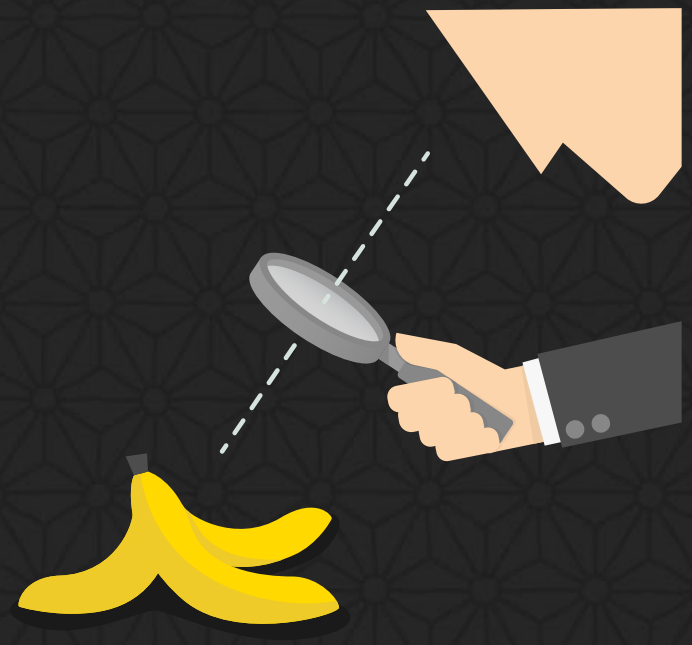


https://thomaswinters.be/samsonbots

Studying & understanding humor

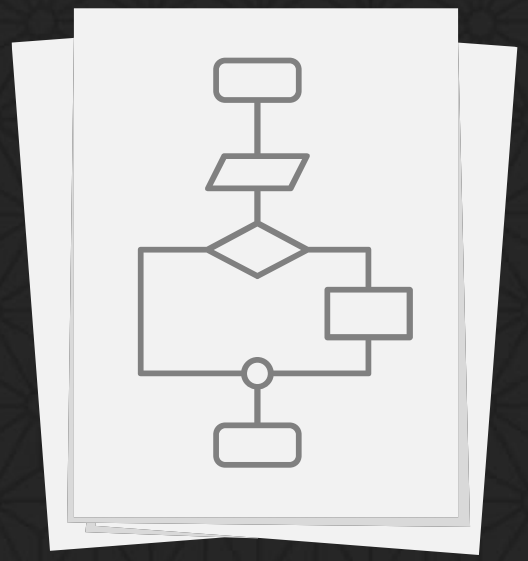
We do not fully know **how humor works**



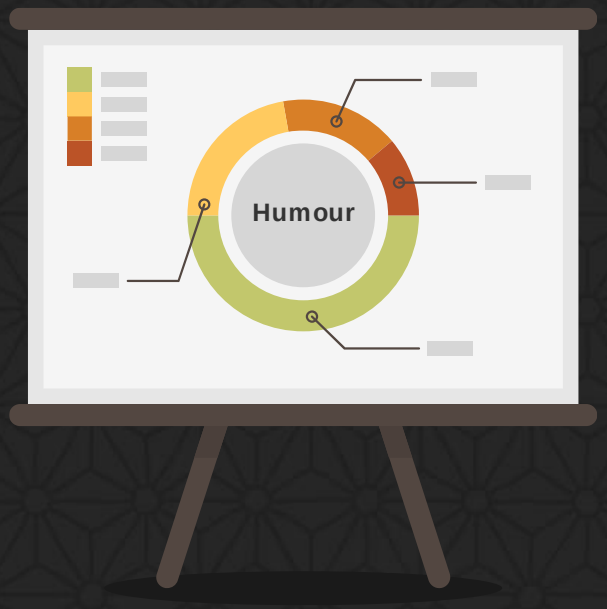


Artificial Intelligence can shed a light!

Can computers have a sense of humor?



Hand-craft rules
for type of humor



Learn rules if
enough data



Better **NLP models**
get us closer & closer

BUT:

AI-complete problem!

Need to process two mental images
before fully understanding the joke

Computational humor is still in its infancy



But, children have a fascinating sense of humor



Thank you!

Thomas Winters

PhD Student at KU Leuven & FWO Fellow

@thomas_wint

thomaswinters.be

Some images (based on the works) of dooder on freepik.com