

Centre for  
Teaching  
and Learning



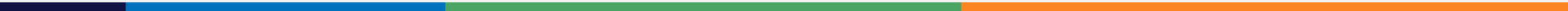
UNIVERSITY OF  
OXFORD

A low-angle photograph of a large, ornate domed building, likely a cathedral or university hall, with a prominent dome and classical architectural details. The building is set against a clear blue sky.

# From structure of language to token embeddings and back

[Dominik.Lukes@ctl.ox.ac.uk](mailto:Dominik.Lukes@ctl.ox.ac.uk)

# Rethinking context in light of the LLM revolution



AI in 2023  $\approx$  LLMs  $\approx$   
Transformers  $\approx$  ChatGPT  
et al.

# Two perspectives on AI



Tool



Subject of  
study

# Linguistics and AI



## Tool

Language tutor  
Reflection of  
language data  
beyond corpus



## Subject

New language speaker  
with certain properties  
Challenge to the very  
foundations of  
linguistics

**Opportunity and  
challenge in one**





Looking up facts



Performing complex calculations



Writing computer code based on text



Making complex decisions from data



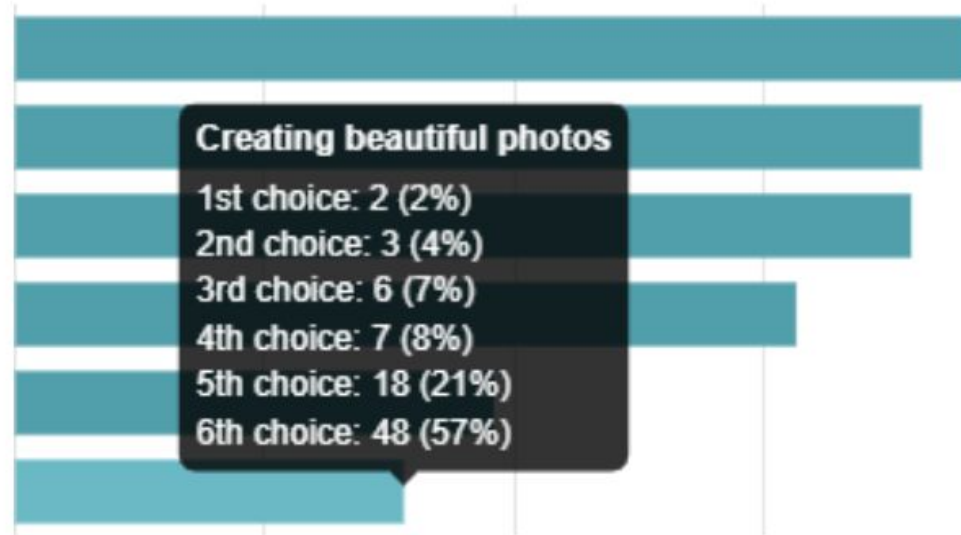
Writing poetry



Making beautiful art

**What do you think AI might be best at?**

- 1 Looking up facts
- 2 Performing complex calculations
- 3 Writing computer code based o...
- 4 Making complex decisions base...
- 5 Writing poetry
- 6 Creating beautiful photos







Write a poem to help me remember the bones in the hand.



In the realm of hands, a tale unfolds,  
Of bones and joints, a story told,  
Listen close, my friend, and hear,  
The names of bones, so you'll endear.

Twenty-seven bones, a framework grand,  
Form the structure of our hand,  
Three groups of bones to memorize,  
Carpals, metacarpals, phalanges arise.

Begin with carpals, eight in sum,  
In the wrist, they're not cumbersome,  
Proximal row, four in row,  
From lateral to medial, here's how they go:

Scaphoid, lunate, triquetrum align,  
Pisiform, a small pea, so benign,  
Distal row, to continue the tour,  
Trapezium, trapezoid, capitate, hamate ensure.



Make a mnemonic for the carpals

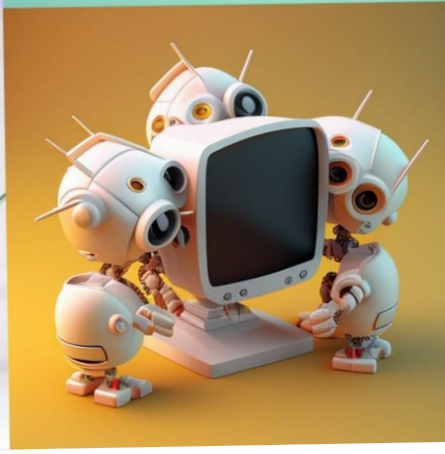


Recall the carpals, eight in all,  
With this mnemonic, you'll never stall:

She Left The Party  
To Take Cathy Home

Scaphoid - She  
Lunate - Left  
Triquetrum - The  
Pisiform - Party  
Trapezium - To  
Trapezoid - Take  
Capitate - Cathy  
Hamate - Home

**Nobody expected AI to be any good at this**



Or this



$4^2 - 234 * -23 * 329870134987$



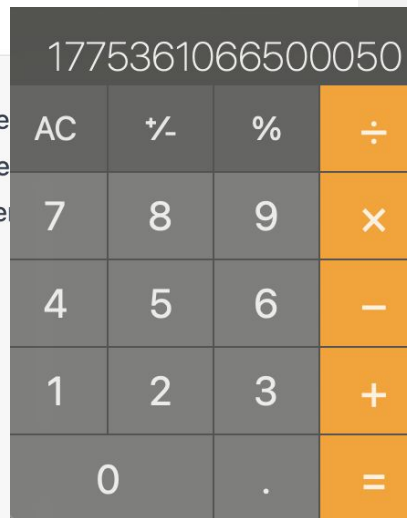
To simplify this expression, we need to follow the order of operations: parentheses, exponents, multiplication and division (performed left to right), and subtraction (performed left to right). As there are no parentheses in the expression, we can simplify it as follows:

$$4^2 - 234 * (-23) * 329870134987$$

$$= 16 + 10825535722602$$

$$= 10825535722618$$

Therefore,  $4^2 - 234 * (-23) * 329870134987$  equals 10825535722618.



We thought AI could do this



Write code to add two numbers in python



In Python, you can add two numbers by simply using the `+` operator. Here's a simple function to add two numbers:



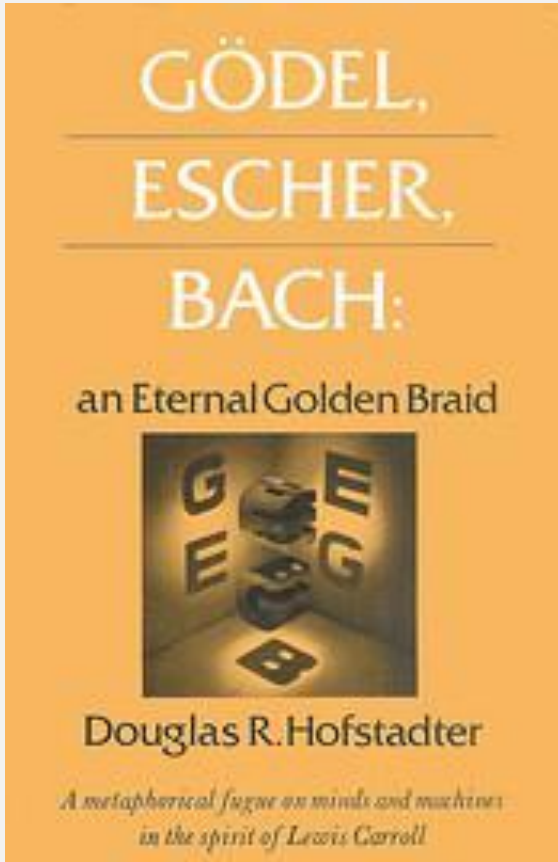
python

Copy code

```
def add_numbers(a, b):  
    return a + b  
  
number1 = float(input("Enter the first number: "))  
number2 = float(input("Enter the second number: "))  
  
result = add_numbers(number1, number2)  
  
print("The sum of the two numbers is:", result)
```

In this code, the `add\_numbers` function takes two arguments, `a` and `b`, and returns their sum. The user is prompted to enter two numbers, which are then passed to the function to perform the addition. The result is then printed.

**But it is also unexpectedly good at this**



## Question:

Will a thinking computer be able to add fast?




## **Speculation:**

Perhaps not. We ourselves are composed of hardware which does fancy calculations but that doesn't mean that our symbol level, where "we" are, knows how to carry out the same fancy calculations.



Different aspects of language have different **degrees of accessibility to consciousness** and, when in consciousness, have different degrees of salience, from backgrounded to foregrounded.

**The challenge of  
LLMs has a long  
history**





1966





Pro dnešní stav jazykovědy je nanejvýš důležitá odpověď na tuto otázku: **Proč se dosud nepodařilo zkonstruovat automatický přístroj, který by mohl převádět akustický signál přirozeného jazyka do sledu diskrétních symbolů**



Aby inženýři mohli zkonstruovat automatický „fonémový detektor“, musí mít především **spolehlivou teorii fonémů.**

Tato teorie bude pravděpodobně vycházet z poznatků získaných při analýze percepce řečového signálu.

1985?



With every linguist I fire, the performance of my speech recognition system goes up.

2013





"To do [NLP], computers would have to master three things. First, enough syntax to uniquely identify the sentence; second, enough semantics to extract its literal meaning; and third, enough pragmatics to infer the intent behind the utterance, and thus discerning what should be done or assumed given that it was uttered."

2019







The biggest lesson that can be read from 70 years of AI research is that **general methods that leverage computation** are ultimately the most effective, and by a large margin.



**Implication?** Jediná  
spolehlivá teorie  
fonému, je žádná  
teorie fonému?

# Main proposition

LLMs run on **weird**  
semantics...

...and this gives us a new  
way of looking at  
language.



Make a list of tropical fruits and divide it into categories by region c



Put it into a table with three columns: Fruit name, Region of origin, Definition



Sure, here's a list of some popular tropical fruits categorized by regi



Sure, here's the table:



**South America:**

- Pineapple
- Papaya
- Passion fruit
- Mango
- Guava
- Acerola cherry
- Acai berry
- Pitaya (Dragon fruit)

**Central America & Mexico:**

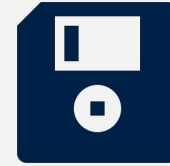
- Avocado
- Banana
- Coconut
- Lime
- Tamarind
- Carambola (Star fruit)
- Sapote

Fruit Name	Region of Origin	Definition
Pineapple	South America	A tropical fruit with a tough, spiky exterior and juicy, sweet interior.
Papaya	South America	A pear-shaped fruit with a soft, orange flesh and black seeds in the center.
Passion fruit	South America	A small, round fruit with a tough, wrinkled exterior and a juicy, tart interior.
Mango	South America	A sweet, juicy fruit with a fibrous pit in the center.
Guava	South America	A round or oval fruit with a green or yellow skin and sweet, fragrant pink flesh.
Acerola cherry	South America	A cherry-sized fruit that is a rich source of vitamin C.
Acai berry	South America	A small, dark purple fruit that grows on the acai palm tree.
Pitaya	South America	A cactus fruit with a pink or yellow exterior and white or pink interior with black seeds.
Avocado	Central America & Mexico	A large, pear-shaped fruit with a thick, rough skin and a creamy, mild-flavored interior.
Banana	Central America & Mexico	A curved fruit with a yellow or green skin and soft, sweet flesh.



## Logic

Compare  
Organise  
Transform



## Memory

Store  
Retrieve







There is no grammar

STEVEN PINKER

Words and  
Rules



No rules and words

**Useful background:  
5 no's and 3 yes's**



1. No **look up** of data
2. No **access to training data** or internal processes
3. No **logic** or text comparison
4. No **backtracking**
5. No **forward** planning



ChatGPT is just a fancy  
**autocomplete**. It just predicts  
the next word.

Literally true, but  
practically wrong.

1. What does it predict?
2. What does it base the predictions on?
3. What does it do with those predictions?

1. Predicts tokens one by one
2. Decides from context with attention to relationships
3. Adding randomness makes it useful



I went back over to the library and got a book on Civil War

This is a great place to start  
books on the Civil War and  
battles, key individuals, po  
relevant sources. You can  
or the stories of individual  
be to find primary sources  
more personal account of

house = 4.35%

other = 2.32%

window = 1.47%

hospital = 1.26%

old = 1.17%

**library = 0.41%**

Total: -5.49 logprob on 1 tokens  
(10.98% probability covered in top 6 logits)

many

s

to find

figures

ion may

ive a

This is what it looks like



**Tokens not words**

I went back over to the library and got a book on Civil War

This is a great place to start  
books on the Civil War and  
battles, key individuals, po  
relevant sources. You can  
or the stories of individual  
be to find primary sources  
more personal account of

house = 4.35%

other = 2.32%

window = 1.47%

hospital = 1.26%

old = 1.17%

library = 0.41%

Total: -5.49 logprob on 1 tokens  
(10.98% probability covered in top 6 logits)

many

s

to find

figures

ion may

ive a

This picture is actually misleading

Ich ging barfuß

Du kannst deine Füße einfach in die richtigen Schuhe stecken und darin gehen. Es

ist sehr wick

Verfügung hast, da barfuß zu

gehen zu Ve

bar = 90.68%

üßen oder Beinen führen kann.

bytes: \xc3 = 0.85%

in = 0.82%

a = 0.80%

d = 0.72%

de = 0.52%

Total: -5.25 logprob on 1 tokens  
(94.38% probability covered in top 6 logits)

Comparison with German reveals this



# Humans

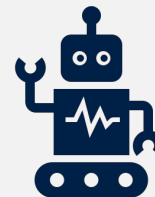
Letters

Words

Sentences

Paragraphs

Texts



# ChatGPT

Tokens

# Tokenizer

The GPT family of models process text using **tokens**, which are common sequences of characters found in text. The models understand the statistical relationships between these tokens, and excel at producing the next token in a sequence of tokens.

You can use the tool below to understand how tokens are created and the total count of tokens in that piece of text.

GPT-3 Codex

Many words map to one token, but some

Unicode characters like emojis may be split into many tokens containing the underlying bytes: 🍌

Sequences of characters commonly found next to each other may be grouped together: 1234567890

Tokens	Characters
64	252

Many words map to one token, but some don't: indivisible.

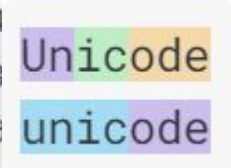
Unicode characters like emojis may be split into many tokens containing the underlying bytes: 🍌🍌🍌🍌🍌🍌

Sequences of characters commonly found next to each other may be grouped together: 1234567890

TEXT TOKEN IDS

OpenAI Tokenizer: ChatGPT only sees tokens

```
[7085, 2456, 3075, 284, 530, 11241, 11, 4, 470, 25, 773, 452,  
98, 3118, 291, 1098, 343, 13210, 271, 743,  
307, 6626, 650, 867, 16326, 7268, 262, 10, 5, 12520, 97, 248,  
8582, 237, 122, 198, 198, 44015, 3007, 286, 3435, 8811, 1043, 1306, 284,  
1123, 584, 743, 307, 32824, 1978, 25, 17031, 2231, 30924, 3829]
```



TEXT    TOKEN IDS

Actually, this is what ChatGPT sees

1 + 2 \* 3654 = 123,565  
2 + 1 / 43^4 = 43.245645

[16, 1343, 362, 1635, 21268, 19, 796, 17031, 11, 47372, 198, 17, 1343,  
352, 1220, 5946, 61, 19, 28, 3559, 13, 1731, 3980, 2231]

Numbers are not digits, just tokens





About 50,000 in total



Only 3,000 English words



Most parts of words incl. letters, spaces

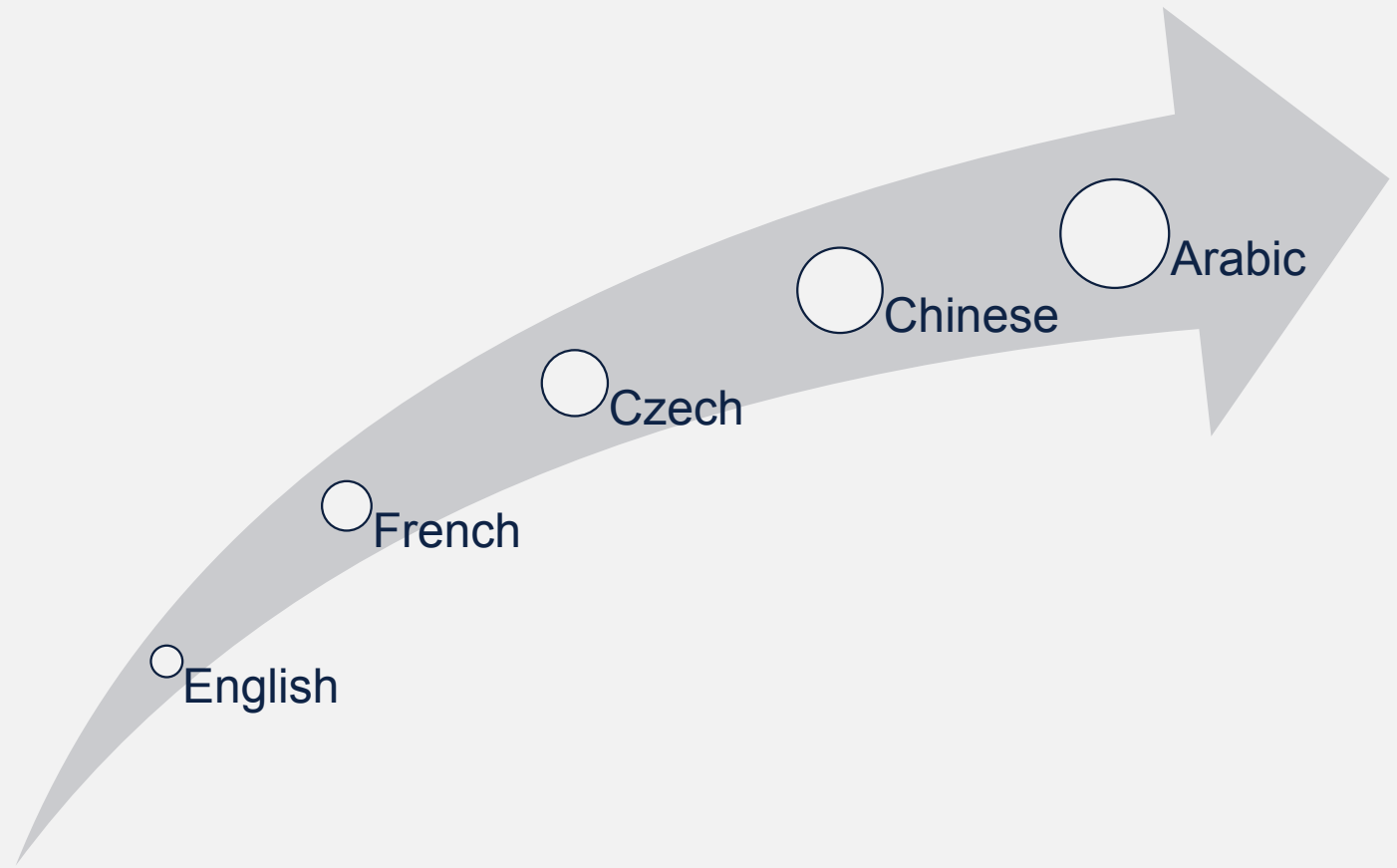


About 140 tokens per 100 English words



Much more for other languages





**Unexpected consequence: More tokens more money**

Unreliable at letter tasks

Unreliable numbers

More grammar mistakes in  
some languages




Sure, here's the story with the first two letters of every word in bold:



**Lily had been asking her mother for a pet for months now. Ever since her best friend Lucy got a puppy named Max, Lily was desperate for a furry friend of her own. She begged and pleaded with her mother daily to let her get a dog. But her mother always said no, insisting that dogs were too much work and responsibility.**

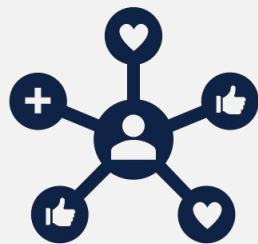
**One day after school, Lily asked her mother yet again if she could get a dog. "Please mom, \*\*I\*\*ll take good care of it! \*\*I\*\*ll walk it, feed it, and clean up after it," she pleaded. But her mother shook her head firmly and said "\*\*I\*\*m sorry Lily, but the answer is still no. We can't handle the commitment of having a dog right now."**

**Note:** Different models have different tokens (both numbers and actual tokens)





# Embeddings: Give tokens meaning



Embeddings show how tokens relate to each other in **multidimensional-dimensional** space.



# Word Embeddings

## The Ill

| *"The gift of words is the gift of deception and illusion" ~Children of Dune*

Discussions: |

With this understanding, we can proceed to look at trained word-vector examples (also called word embeddings) and start looking at some of their interesting properties.

Translations: |

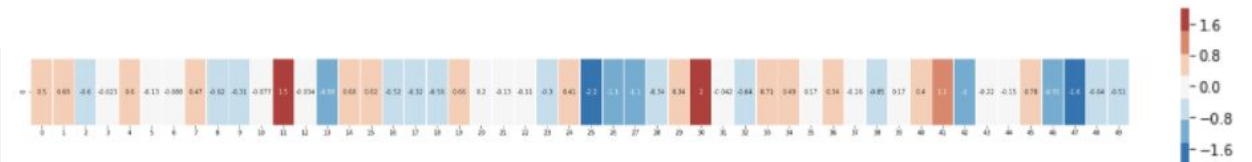
This is a word embedding for the word "king" (GloVe vector trained on Wikipedia):

```
[ 0.50451, 0.68607, -0.59517, -0.022801, 0.60046, -0.13498, -0.08813, 0.47377, -0.61798, -0.31012, -0.076666, 1.493, -0.034189, -0.98173, 0.68229, 0.81722, -0.51874, -0.31503, -0.55809, 0.66421, 0.1961, -0.13495, -0.11476, -0.30344, 0.41177, -2.223, -1.0756, -1.0783, -0.34354, 0.33505, 1.9927, -0.04234, -0.64319, 0.71125, 0.49159, 0.16754, 0.34344, -0.25663, -0.8523, 0.1661, 0.40102, 1.1685, -1.0137, -0.21585, -0.15155, 0.78321, -0.91241, -1.6106, -0.64426, -0.51042 ]
```

It's a list of 50 numbers. We can't tell much by looking at the values. But let's visualize it a bit so we can compare it other word vectors. Let's put all these numbers in one row:

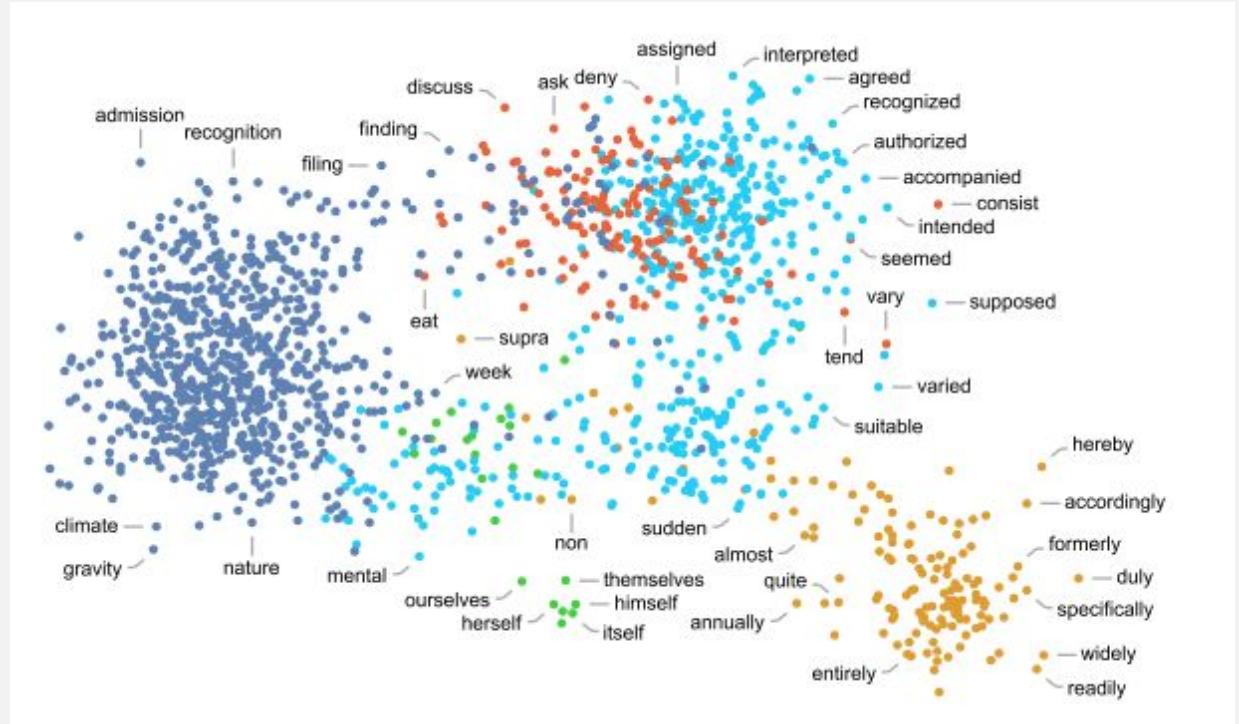


Let's color code the cells based on their values (red if they're close to 2, white if they're close to 0, blue if they're close to -2):





You do maths on matrices and you get



See in COCA 🏠 Collocates Clusters Topics Websites KWIC ⬇️ HELP 📄 ▶️

# argument

(NOUN) ☆ ↻

#1616 +

- 1. a fact or assertion offered as evidence that something is true
- 2. a contentious speech act
- 3. a discussion in which reasons are advanced for and against some proposition or proposal

D M O C G E

🖼️ 🔊 YouGlish PlayPhrase Yarn

🌐 Translate: choose language

SYNONYMS ( ▶️ CONCEPT)      NEW: DEFIN +SPEC +GENL

case, claim, contention, reason  disagreement, dispute, fight, quarrel, row, spat, squabble

## TOPICS (more)

[argue](#), [function](#), [court](#), [specify](#), [string](#), [default](#), [variable](#), [debate](#), [object](#), [command](#), [evidence](#), [theory](#), [moral](#), [parameter](#), [value](#), [file](#), [method](#), [return](#), [judge](#), [supreme](#)

## COLLOCATES (more)

NOUN [function](#), [command](#), [sake](#), [evidence](#), [string](#), [keyword](#), [conclusion](#), [array](#)

VERB [pass](#), [present](#), [specify](#), [accept](#), [reject](#), [advance](#), [evaluate](#), [construct](#)

ADJ [oral](#), [optional](#), [closing](#), [valid](#), [compelling](#), [logical](#), [convincing](#), [heated](#)

ADV [logically](#), [optionally](#), [implicitly](#), [intellectually](#), [hominem](#), [rationally](#), [false](#), [convincingly](#)

## RELATED WORDS

[argue](#), [arguably](#), [argumentative](#), [arguable](#), [argumentation](#), [unarguably](#), [unarguable](#)

## CLUSTERS (more)

argument •	argument for • arguments for • argument in • arguments in • argument with • argument against • argument to • argument about
• argument	first argument • second argument • one argument • two arguments • good argument • oral argument • strong argument • closing arguments

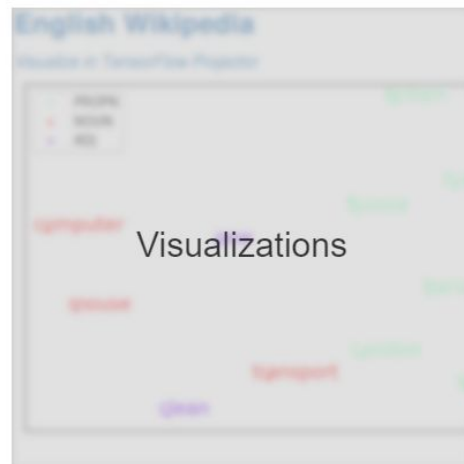
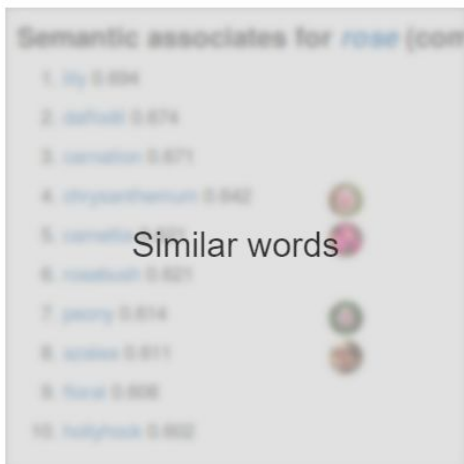
# WebVectors: word embeddings online

'You shall know a word by the company it keeps.' (Firth 1957)

Enter a word to produce a list of its 10 nearest semantic associates.

English Wikipedia model will be used; for other models, visit [Similar Words](#) tab.

Find similar words!



# Semantic Calculator

Calculate ratios, such as «find a word D related to the word C in the same way as the word A is related to the word B». An example is given in the placeholder: which word is in the same relation to the word «father» as «daughter» is to «mother»? The answer is «son». [More on this...](#)

mother\_NOUN



daughter\_NOUN

father\_NOUN



???

## Semantic

Word frequency

High  Medium

1. assertion I
2. counterar
3. objection I
4. rebuttal O.
5. justificatio
6. notion 0.5
7. rationale C
8. pleading C
9. questionin
10. affirmative

• We show or associates c

Word frequency

High  Medium  Low

## English Wikipedia

1. son NOUN 0.88
2. grandson NOUN 0.80
3. nephew NOUN 0.77
4. granddaughter NOUN 0.76
5. grandfather NOUN 0.73



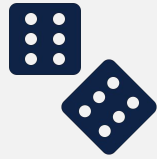
But this is deceptive

**BUT** when we  
see “father”

-0.26299598515033722, 0.02255615033209324, -0.09917508065700531, -0.07070962339639664, 0.08635025471448898, 0.06686452776193619, -0.0407336950302124, -0.04072333127260208, -0.01974628120660782, 0.07472220063209534, -0.024722471833229065, -0.13420116901397705, -0.01812688820064068, -0.07096941769123077, -0.05353084206581116, -0.10960721969604492, -0.017906684428453445, -0.047330949454549965, -0.02091103047132492, 0.1269848346710205, -0.05413510650396347, -0.046787846833467484, 0.0024005023296922445, -0.0721780061721801, -0.029329143464565277, 0.007498500403016806, -0.034666456282138824, -0.03568940982222557, 0.03427724167704582, 0.02315753698348999, -0.008645392023026943, 0.0533395223319530, 0.07456360012292862, 0.147796630859375, -0.006483903620392084, -0.08905889838933945, 0.03265034034848213, -0.0732979029417038, 0.04066538065671921, 0.023211032152175903, -0.012049349024891853, -0.02828565053641796, 0.019329581409692764, 0.09989447146654129, 0.1430598795413971, -0.061100199818611145, -0.030345138162374496, -0.0298450794070959, -0.028366880491375923, 0.052052564918994904, 0.036766957491636276, 0.003982939291745424, -0.077084481716156, 0.05044832453131676, -0.11687757074832916, 0.06646141409873962, 0.016255078837275505, -0.06982151418924332, -0.000822143629193306, -0.0026820336934179068, -0.004263593349605799, 0.09659365564584732, 0.06130471080541611, -0.068409085273742, 0.06686245650053024, -0.04831290245056152, 0.08598440140485764, 0.08331689983606339, 0.08026000112295151, 0.05451888591051102, -0.03798443824052811, 0.04084145650267601, -0.12311697751283646, 0.023645302280783653, 0.005237551871687174, 0.03906212002038956, 0.037468183785676956, -0.05121520534157753, -0.10456130653619766, 0.009842721745371819, 0.04819759353995323, -0.13286681473255157, 0.02991127222776413, -0.06024811416864395, 0.04108288511633873, -0.008447377011179924, -0.07916080206632614, 0.06436653435230255, 0.017831943929195404, -0.054629500955343246, 0.027066148817539215, -0.030593710020184517, -0.10156133025884628, -0.0013401528121903539, 0.0011191506637260318, 0.009616676717, -0.02962290495634079, 0.0042936066165566444, 0.013841508887708187, -0.047656722366809845, -0.003912750165909529, 0.06500802934169769, 0.001283025718294084, -0.08169967681169, 0.06566621363162994, -0.010957532562315464, -0.028156422078609467, 0.08978854864835739, -0.0003194105520378798, -0.02697799727320671, -0.006005867850035429, 0.07932088524103, 0.021490609273314476, 0.013727870769798756, -0.019940776750445366, 0.031798265874385834, -0.0457642637193203, 0.03235720098018646, -0.022082772105932236, -0.0490235351026058, -0.11819718778133392, -0.04506421089172363, -0.046244461089372635, 0.029877550899982452, -0.07711911201477051, 0.05314543470740318, -0.09000932425260544, -0.02375070564448833, -0.05107633396983147, 0.001467616413719952, -0.02442317083477974, 0.01248782780021429, 0.06548482179641724, 0.043813593685626984, 0.06102786585688591, 0.021692050620913506, -0.052160654217004776, -0.009674523957073689, -0.072069451212883, -0.08633119612932205, -0.05121589079499245, -0.08108754456043243, 0.03608304262161255, 0.06553766876459122, -0.0727415531873703, -0.09346839785575867, -0.07251054048538208, 0.04504929482936859, -0.01773262582719326, -0.0005254627903923392, -0.0035706141497939825, 0.0906830206513404, 0.0152428038418293, 0.009525319561362267, 0.02502918615937233, 0.02807294949889183, -0.08951258659362793, 0.018022941425442696, 0.04113161191344261, -0.09941867738962173, 0.03642140328884125, 0.07755865901708603, 0.014834643341600895, -0.05757498741149902, -0.0052739898674190044, -0.03217893838882446, 0.029460914433002472, -0.0358795598149299, 0.016881171613931656, -0.015574142336845398, -0.10131996870040894, -0.01736866682767868, 0.014807181432843208, -0.03830776736140251, -0.0307577196508646, -0.04063287377357483, 0.0017508701421320438, 0.06622152030467987, 0.06959225982427597, 0.03921446576714516, -0.029292205348610878, -0.07731080055236816, -0.0757351890206337, 0.008267058990895748, 0.10628201067447662, -0.006961626932024956, -0.060704007744789124, -0.024280674755573273, -0.011232278309762478, 0.02305467799305916, -0.040246833115816116, 0.035518884658813, -0.12048669904470444, -0.0057440041564404964, -0.008801680989563465, -0.038733456283807755, -0.0384967215359211, -0.0059003811329603195, 0.07543318718671799, 0.00299985124729, 0.11148137599229813, 0.0560586079955101, -0.01694066822528839, -0.020253779366612434, -0.11995487660169601, 0.10403268039226592, -0.022030610591173172, 0.019188301637768745, -0.03581297770142555, -0.04047990494155884, -0.03492145985364914, 0.027967417612671852, -0.07497915625572205, 0.032431814819574356, -0.025854842737317085, -0.105954959885940, -0.09982465207576752, -0.05515384301543236, 0.02156943641602993, 0.05118619278073311, -0.03904290497303009, -0.022826874628663063, -0.053247928619384766, -0.10935184359550476, 0.0006719367229379714, -0.016026955097913742, 0.13483813405036926, 0.1173691526055336, -0.01902260072529316, -0.09690848737955093, -0.07585378736257553, 0.007626112550497055, 0.019889818504452705, -0.008633404038846493, 0.010355712845921516, 0.035737670958042145, 0.011519350111484528, -0.005264237057417631, -0.06305427849292755, -0.02626376040279, 0.008310412988066673, -0.0068666874431073666, -0.13443514704704285, -0.025350390002131462, -0.0079041114076972, 0.014966381713747978, 0.01571144163608551, 0.0626633390784263, 0.05788900703191757, -0.022854981943964958, 0.09513315558433533, 0.1284472942352295, -0.061813995242118835, -0.049407169222831726, -0.10701776295900345, 0.06945358961820602, -0.07409369200468063, -0.028664348646998405, -0.0144350565969944, 0.029182329773902893, 0.007034373469650745, -0.026693496853113174, 0.0590004064142704, -0.00290216808207333, 0.02047384679317474, 0.02306350134313106, -0.0015993582783266902, 0.00770201021805

The model actually sees (for tokens)





**Final step: Add in  
randomness**

I went back over to the library and got a book on Civil War

This is a great place to start  
books on the Civil War and  
battles, key individuals, po  
relevant sources. You can  
or the stories of individual  
be to find primary sources  
more personal account of

house = 4.35%

other = 2.32%

window = 1.47%

hospital = 1.26%

old = 1.17%

**library = 0.41%**

Total: -5.49 logprob on 1 tokens  
(10.98% probability covered in top 6 logits)

many

s

to find

figures

ion may

ive a

The top item is not always selected

I went through the course, and I really enjoyed it. I found the material to be informative and well-structured, and I was able to learn a lot of new information about the subject matter. I liked the variety of activities available, and the support provided. I liked the hands-on approach to try out different concepts and the interactive, encouraging me to apply what I was learning. Course to anyone interested in the



Mode

Complete

Model

I went through the specified function line by line and saved the value of each expression in a comment.

```
//'22 + 34'
```

```
56
```

```
//'56 - 13'
```

```
43
```

```
//'43 * 2'
```

```
86
```

```
//'86 / 2'
```

```
43
```

```
//'Math.floor(43.8)'
```

```
43
```

I went to see Dr. Chen for help with my back pain

When you go to see Dr. Chen for help with your back pain, you should bring any medical records or imaging scans you have related to the issue. You should also be prepared to answer questions about your medical history and current symptoms, like when the pain started and how it affects you. Additionally, you should also make sure that you understand what treatment options he suggests, and bring a list of questions you may have. Lastly, it's important that you provide Dr. Chen with honest information in order for him to make the best decisions for you.

With heat

I went to the store

I went to the store to buy some groceries. I bought milk, eggs, bread, and some fruit. I al

store = 98.84%

grocery = 1.10%

local = 0.05%

supermarket = 0.01%

convenience = 0.00%

Total: -0.01 logprob on 1 tokens

(100.00% probability covered in top 5 logits)

I went to the store

I went to the store to buy some groceries. I bought milk, eggs, bread, and some fruit. I also picked up some snacks for later.

I went to the store

I went to the store to buy some groceries. I bought milk, eggs, bread, and some fruit. I also picked up some snacks for later.

I went to the store to buy some

I went

I went to the store

I went  
some

I went to the store to buy some groceries. I bought milk, eggs, bread, and some fruit. I also picked up some snacks for later.

I went to the store to buy some groceries. I bought milk, eggs, bread, and some fruit. I also picked up some snacks for later.

I went to the store

I went to the store to buy some groceries. I bought milk, eggs, bread, and some fruit. I also picked up some snacks for later.

I went to the store

I went to the store to buy some groceries. I bought milk, eggs, bread, and some fruit. I also picked up some snacks for later.

What it looks like with temperature 0

What is the role of  
linguistics in all of this?

**Tokens** = form of compression

**Embeddings** = dimensionality  
reduction


**Randomness** = Human  
knowledge?

Everything about this  
process screams, structure  
should matter!



But we know what matters  
and what does not.

Specify **semantically rich context** and don't worry about form



Spelling

Punctuation

Word order

Mostly include , ? “ for  
your own readability.

Examples

Perspectives

Styles

# Language Models are Few-Shot Learners

Tom B. Brown*	Benjamin Mann*	Nick Ryder*	Melanie Subbiah*	
Jared Kaplan†	Prafulla Dhariwal	Arvind Neelakantan	Pranav Shyam	Girish Sastry
Amanda Askell	Sandhini Agarwal	Ariel Herbert-Voss	Gretchen Krueger	Tom Henighan
Rewon Child	Aditya Ramesh	Daniel M. Ziegler	Jeffrey Wu	Clemens Winter
Christopher Hesse	Mark Chen	Eric Sigler	Mateusz Litwin	Scott Gray
Benjamin Chess	Jack Clark	Christopher Berner		
Sam McCandlish	Alec Radford	Ilya Sutskever	Dario Amodei	

OpenAI

## Abstract

Recent work has demonstrated substantial gains on many NLP tasks and benchmarks by pre-training on a large corpus of text followed by fine-tuning on a specific task. While typically task-agnostic in architecture, this method still requires task-specific fine-tuning datasets of thousands or tens of thousands of examples. By contrast, humans can generally perform a new language task from only a few examples or from simple instructions – something which current NLP systems still largely struggle to do. Here we show that scaling up language models greatly improves task-agnostic, few-shot performance, sometimes even reaching competitiveness with prior state-of-the-art fine-tuning approaches. Specifically, we train GPT-3, an autoregressive language model with 175 billion parameters, 10x more than any previous non-sparse language model, and test its performance in the few-shot setting. For all tasks, GPT-3 is applied without any gradient updates or fine-tuning, with tasks and few-shot demonstrations specified purely via text interaction with the model. GPT-3 achieves strong performance on many NLP datasets, including translation, question-answering, and cloze tasks, as well as several tasks that require on-the-fly reasoning or domain adaptation, such as unscrambling words, using a novel word in a sentence, or performing 3-digit arithmetic. At the same time, we also identify some datasets where GPT-3’s few-shot learning still struggles, as well as some datasets where GPT-3 faces methodological issues related to training on large web corpora. Finally, we find that GPT-3 can generate samples of news articles which human evaluators have difficulty distinguishing from articles written by humans. We discuss broader societal impacts of this finding and of GPT-3 in general.

## Zero-shot

The model predicts the answer given only a natural language description of the task. No gradient updates are performed.

```
1 Translate English to French: ← task description
2 cheese => ..... ← prompt
```

## One-shot

In addition to the task description, the model sees a single example of the task. No gradient updates are performed.

```
1 Translate English to French: ← task description
2 sea otter => loutre de mer ← example
3 cheese => ..... ← prompt
```

## Few-shot

In addition to the task description, the model sees a few examples of the task. No gradient updates are performed.

```
1 Translate English to French: ← task description
2 sea otter => loutre de mer ← examples
3 peppermint => menthe poivrée ←
4 plush girafe => girafe peluche ←
5 cheese => ..... ← prompt
```



## Playground

Load a preset...



Save

View code

### Playground

Going to the

Yes, I am!

Submit



Going to the library?

Yes, you can go to the library! Hours and services may vary by location, so check with your local library for specific information.

Submit



50



## Playground

Load a preset...

A: Going to the pub?

B: Yes, I am.

A: Going to the swimming pool?

B: No, not to day:

A: Going to the library?

B: Yes, I am.

A: Going to the shop?

B: No, not today.

Submit



69

Engineered prompt to force conversation

How does it work?



Let's have a conversation in English at the level of A1. You are a shopkeeper and I am a student. I'm trying to buy bread. You need to simplify the way you speak so that I can understand.



Shopkeeper: Hello, how can I help you?



Student: Hi, I want to buy bread.

Shopkeeper: Sure, what type of bread would you like? We have white bread, whole wheat bread, and rye bread.

Student: What's the difference?

Shopkeeper: White bread is made with flour that has been stripped of its nutrients. Whole wheat bread is made with flour that includes the bran and germ, which makes it healthier. Rye bread is made with rye flour and has a distinctive flavor.

Student: I want the healthier one. Can I have whole wheat bread?

Shopkeeper: Sure, how many loaves would you like?

Student: Just one loaf, please.

Shopkeeper: That will be \$2.50. Would you like anything else?

Student: No, that's all. Thank you.

Shopkeeper: You're welcome. Have a nice day!

**Just giving command to use concept is not enough**





As an expert ESL teacher, describe the speaking skills of a student at level A1 according to CEFR.



Now write a conversation in English following these principles to illustrate a conversation between a shopkeeper and a beginner student. The student is trying to buy bread. The shopkeeper needs to simplify the way they speak so that the student can understand.



The Common European Framework of Reference for Languages (CEFR) is a widely accepted framework for assessing language proficiency. According to CEFR, a student at level A1, often referred to as a 'Breakthrough' or 'Beginner' level, will have the following speaking skills:



1. **Basic Vocabulary:** An A1 student can use a very basic range of words and phrases about personal details and specific concrete situations. They can use some simple vocabulary and expressions but are limited and often rely on repetition and high-frequency language.
2. **Simple Communication:** They can interact in a simple way provided the other person is prepared to repeat or rephrase things at a slower rate of speech and help them formulate what they are trying to say. They can initiate and respond to simple statements in areas of immediate need or on very familiar topics.
3. **Expressions:** They can introduce themselves and others, and can ask and answer questions about personal details such as where they live, people they know, and things they have.
4. **Pronunciation:** Pronunciation will generally be understood, albeit with some difficulty, due to a likely heavy influence from their first language. Mispronunciations may still be frequent at this level.
5. **Grammatical Accuracy:** Their use of grammar is often simplified or even incorrect, but the intent can usually be understood in context. The use of present simple tense is common, with other tenses generally not used correctly or understood.
6. **Fluency:** Fluency is minimal, with a very basic level of conversational competence.



Sure, here's an example of such a conversation:



**Student:** Hello.

**Shopkeeper:** Hello! How can I help you today?

**Student:** I want...uh...bread.

**Shopkeeper:** Of course! We have many types of bread. Do you like white bread or brown bread?

**Student:** Uh...brown...brown bread.

**Shopkeeper:** Very good! How many pieces do you want? One? Two?

**Student:** One...please.

**Shopkeeper:** Alright, here is your brown bread. Anything else you need?

**Student:** No, thank you. How much?

**Shopkeeper:** It's two dollars.

**Student:** Here...two dollars.

**Shopkeeper:** Thank you! Here's your bread. Have a nice day!

**Student:** You...too. Thank you.

Ask to give its own examples and then follow them

# Chain-of-Thought Prompting Elicits Reasoning in Large Language Models

Jason Wei Xuezhi Wang Dale Schuurmans Maarten Bosma  
Brian Ichter Fei Xia Ed H. Chi Quoc V. Le Denny Zhou

Google Research, Brain Team  
{jasonwei, dennyzhou}@google.com

## Abstract

We explore how generating a *chain of thought*—a series of intermediate reasoning steps—significantly improves the ability of large language models to perform complex reasoning. In particular, we show how such reasoning abilities emerge naturally in sufficiently large language models via a simple method called *chain-of-thought prompting*, where a few chain of thought demonstrations are provided as exemplars in prompting.

Experiments on three large language models show that chain-of-thought prompting improves performance on a range of arithmetic, commonsense, and symbolic reasoning tasks. The empirical gains can be striking. For instance, prompting a PaLM 540B with just eight chain-of-thought exemplars achieves state-of-the-art accuracy on the GSM8K benchmark of math word problems, surpassing even finetuned GPT-3 with a verifier.

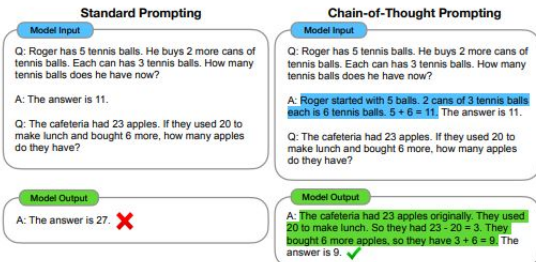


Figure 1: Chain-of-thought prompting enables large language models to tackle complex arithmetic, commonsense, and symbolic reasoning tasks. Chain-of-thought reasoning processes are highlighted.

# Chain of thought (“step by step”) improves performance on number tasks.



## LARGE LANGUAGE MODELS ARE HUMAN-LEVEL PROMPT ENGINEERS

Yongchao Zhou<sup>1,2,\*</sup>, Andrei Ioan Muresanu<sup>2,3,\*</sup>, Ziyen Han<sup>1,2,\*</sup>, Keiran Paster<sup>1,2</sup>, Silviu Pitis<sup>1,2</sup>, Harris Chan<sup>1,2</sup>, Jimmy Ba<sup>1,2</sup>  
<sup>1</sup>University of Toronto <sup>2</sup>Vector Institute <sup>3</sup>University of Waterloo \*Equal contribution  
{yczhou,hanziwen,keirp,spitis,hchan,jba}@cs.toronto.edu  
{andrei.muresanu}@uwaterloo.ca

### ABSTRACT

By conditioning on natural language instructions, large language models (LLMs) have displayed impressive capabilities as general-purpose computers. However, task performance depends significantly on the quality of the prompt used to steer the model, and most effective prompts have been handcrafted by humans. Inspired by classical program synthesis and the human approach to prompt engineering, we propose *Automatic Prompt Engineer*<sup>1</sup> (APE) for automatic instruction generation and selection. In our method, we treat the instruction as the “program”, optimized by searching over a pool of instruction candidates proposed by an LLM in order to maximize a chosen score function. To evaluate the quality of the selected instruction, we evaluate the zero-shot performance of another LLM following the selected instruction. Extensive experiments show that our automatically generated instructions outperform the prior LLM baseline by a large margin and achieve better or comparable performance to the instructions generated by human annotators on 24/24 Instruction Induction tasks and 17/21 curated BIG-Bench tasks. We conduct extensive qualitative and quantitative analyses to explore the performance of APE. We show that APE-engineered prompts are able to improve few-shot learning performance (by simply prepending them to standard in-context learning prompts), find better zero-shot chain-of-thought prompts, as well as steer models toward truthfulness and/or informativeness.

“Let’s work this out in a step by step way to be sure we have the right answer.”

What matters is semantics  
but it is **weird semantics**.

**Poetry** = rich in semantic relationships and relies on reader resonance



Computer code

Logic and math

Lists and definitions

Essays

Foreign languages

Part of this weirdness is  
hidden because...

Raw embeddings (pre-training)  
Instruction tuning  
RLHF



What kind of  
“representations” are we  
seeing in an LLM output?



# Compliance

## Definition:

This frame concerns **Acts** and **State\_of\_affairs** for which **Protagonists** are responsible and which either follow or violate some set of rules or **Norms**.

A lot of people suspect that **blocking savers' access to their money** **VIOLATES** **the new constitution's explicit ban on compulsory loans to the government**.

Not that by **ADHERING** to the subject-before-predicate principle, **the translator of the above extract** has had to ignore the principle of end-weight.

## FEs:

### Core:

#### **Act [Act]**

Semantic Type: State\_of\_affairs

Excludes: State\_of\_affairs

This FE identifies the **Act** that is judged to be in or out of compliance with the **Norms**.  
Jo's **VIOLATION** of the no trespassing law came as a surprise.

#### **Norm [Norm]**

This FE identifies the rules or **Norms** that ought to guide a person's behavior.  
John **VIOLATED** **the law** by not paying taxes.

#### **Protagonist [Prot]**

Semantic Type: Sentient

The **Protagonist's** behavior is in or out of compliance with norms.  
**Jo** **VIOLATED** the no trespassing law.

#### **State\_of\_affairs [soa]**

A **State\_of\_affairs** may violate a law or rule.  
**Their pricing policy** **VIOLATED** inter-governmental agreements.

- [Commerce scenario](#)
- [Commerce sell](#)
- [Commercial transaction](#)
- [Commitment](#)
- [Committing crime](#)
- [Commonality](#)
- [Communicate categorization](#)
- [Communication](#)
- [Communication manner](#)
- [Communication means](#)
- [Communication noise](#)
- [Communication response](#)
- [Commutation](#)
- [Commutative process](#)
- [Commutative statement](#)
- [Compatibility](#)
- [Competition](#)
- [Complaining](#)
- [Completeness](#)
- [Compliance](#)
- [Concessive](#)
- [Condition symptom relation](#)
- [Conditional occurrence](#)
- [Conditional scenario](#)
- [Conduct](#)
- [Confronting problem](#)
- [Connecting architecture](#)
- [Connectors](#)
- [Conquering](#)
- [Contact image schema](#)
- [Contacting](#)
- [Container focused placing](#)
- [Container focused removing](#)
- [Containers](#)
- [Containing](#)
- [Containment scenario](#)
- [Contingency](#)
- [Continued state of affairs](#)
- [Contextual circumstances](#)



SEARCH

WORD

CONTEXT

OVERVIEW

See in COCA Collocates Clusters Topics Websites KWIC HELP

# compliance (NOUN)

#2443

1. acting according to certain accepted standards 2. the act of submitting 3. a disposition or tendency to yield to the will of others

D M O C G **E**

YouGlish PlayPhrase Yarn

Translate: choose language

SYNONYMS ( ▶ CONCEPT) NEW: DEFIN +SPEC +GENL

**conformity** compliance, conformity, fulfillment, observance  
**obedience** agreement, compliance, defiance, obedience,  
passivity, submission

## TOPICS (more)

[regulation](#), [requirement](#), [management](#), [regulatory](#), [comply](#), [security](#), [standard](#), [audit](#),  
[employee](#), [ensure](#), [risk](#), [organization](#), [financial](#), [policy](#), [applicable](#), [procedure](#), [federal](#), [safety](#),  
[provision](#), [violation](#)

## COLLOCATES (more)

**NOUN** [requirement](#), [law](#), [regulation](#), [standard](#), [security](#), [policy](#), [risk](#), [officer](#)

**VERB** [ensure](#), [maintain](#), [monitor](#), [achieve](#), [demonstrate](#), [conduct](#), [enforce](#), [assure](#)

**ADJ** [regulatory](#), [legal](#), [federal](#), [applicable](#), [environmental](#), [corporate](#), [responsible](#), [internal](#)

**ADV** [full](#), [therewith](#), [proactively](#), [ethically](#), [hereof](#), [cost-effectively](#), [herewith](#), [hoc](#)

## CLUSTERS (more)

compliance •	compliance with • compliance requirements • compliance officer • compliance to • compliance issues • compliance program • compliance in • compliance for
• compliance	in compliance • ensure compliance • for compliance • regulatory compliance • to compliance • pci compliance • into compliance • full compliance
compliance ••	compliance with applicable • compliance with federal • compliance with state • compliance with laws • compliance and enforcement • compliance with local • compliance with section • compliance and ethics



Now write Czech tabloid front page article about this.



## Přísná Kontrola Dodržování Předpisů Zahájena ve Velké Společnosti!

### Exkluzivní zpráva

V jedné z předních společností ve městě se chystá důkladná kontrola dodržování předpisů, což vyvolává vážné obavy mezi zaměstnanci. Tento týden byl oznámen náhlý **audit dodržování předpisů**, který má za cíl zkontrolovat, zda společnost dodržuje všechny platné normy a zákony.

### Napětí mezi Zaměstnanci Stoupá

Zprávy o nadcházejícím auditu se rychle šíří mezi zaměstnanci, kteří nyní diskutují o nejlepších způsobech, jak zajistit, že všechny postupy a dokumentace jsou v souladu s interními a externími předpisy.

importance of compliance and adheres to the standards.



Create a framenet entry.



Creating a FrameNet entry involves outlining a semantic frame for the word "compliance".

Here's how a possible FrameNet entry might look:



## Frame: Compliance

### Definition

The state or process of adhering to a set of given rules, standards, or regulations, often within a formal or institutional context.

### Core Elements

#### 1. Agent

- **Definition:** The entity (person, group, or organization) that is expected to comply with the rules or standards.
- **Example:** The company, the employee, the government.

#### 2. Standard

- **Definition:** The set of rules, guidelines, or norms that the agent is expected to adhere to.
- **Example:** Legal regulations, company policies, societal norms.

ited terms

/nset for

al).

t can be

ication of

Morphology

Cross-language pre-training

World-knowledge / language  
blending



Write sentences in Czech following the pattern: I was doing X, when Y happened.



1. Dělal jsem úklid, když zazvonil telefon.
2. Četl jsem knihu, když se vypnula elektřina.
3. Vařil jsem večeři, když přišli návštěvníci.
4. Kreslil jsem, když začalo pršet.
5. Psal jsem email, když mi došla baterie v počítači.
6. Běžel jsem v parku, když jsem spatřil starého přítele.
7. Studoval jsem pro zkoušku, když zazněl požární alarm.
8. Sledoval jsem film, když mi upadl popcorn na zem.
9. Pracoval jsem v zahradě, když přiletěl kolibrík.
10. Hrál jsem na kytaru, když se strhla bouřka.

Alex: Ahoj, dobře mluvíš anglicky?

Taxi driver: Ano, mluvím. Kde chceš jet?

Alex: Chtěl bych jet do Dejvic. Kolik to bude stát?

Taxi driver: To bude stát asi 500 korun.

Alex: Děkuji, to je v pořádku. Můžu si napsat tvoje jméno a číslo, prosím?

Taxi driver: Ano, samozřejmě. Jmenuji se Petr a číslo je 777 123 456.

Alex: Děkuji, Petra.

Taxi driver: Není zač, kam tě přesně vezu?

Alex: Jedu k rodině, která bydlí na ulici Pod Kaštany v Dejvicích.

Taxi driver: Rozumím, tam tě dovezu.

Alex: Děkuji moc, ahoj!

Taxi driver: Ahoj!



Alex: Dobrý den! (Hello!)

Taxi Driver: Dobrý den! Kam jedeme? (Hello! Where are we going?)

Alex: Jedeme do Dejvic, prosím. (We're going to Dejvice, please.)

Taxi Driver: Rozumím. Kolik lidí pojedete? (I understand. How many people will be traveling?)

Alex: Jedu sám. (I'm traveling alone.)

Taxi Driver: Dobře. Kolik bude stát? (Alright. How much will it cost?)

Alex: Kolik to bude stát, prosím? (How much will it cost, please?)

Taxi Driver: To bude stát osm set korun. (It will cost eight hundred crowns.)

Alex: Děkuji. (Thank you.)

Taxi Driver: Není zač. (You're welcome.)



How will we know what an LLM will be good at?



Sam Altman 

@sama



here is GPT-4, our most capable and aligned model yet. it is available today in our API (with a waitlist) and in ChatGPT+.

[openai.com/research/gpt-4](https://openai.com/research/gpt-4)

it is still flawed, still limited, and it still seems more impressive on first use than it does after you spend more time with it.

Let's keep this in mind



The only way to evaluate ChatGPT and similar tools is **systematically** and over time.

Linguistics is uniquely suited to this role and can help offer a long tradition of language analysis.

# Motto

*V lingvistice je po mém soudu třeba stále obnovovat kontakt s „jazykovou realitou“, snažit se, aby „teorie“ byly teoriemi a ne ideologiemi, a kultivovat k tomu přijatelné (neškodlivé, snesitelné, dokonce vhodné) způsoby vyjadřování. A nemusí to pořád dělat všichni, stačí, když to dělá někdo a je slyšen.*

Pavel Novák

Centre for  
Teaching  
and Learning



UNIVERSITY OF  
OXFORD

A detailed close-up of a humanoid robot with a friendly expression, wearing large, round, metallic-rimmed glasses. The robot is holding an open book and appears to be reading. The background is a blurred library with bookshelves and warm lighting. A dark blue semi-transparent box is overlaid on the right side of the image, containing the text 'Děkuji'.

**Děkuji**



This presentation is licensed under Creative Commons By Attribution license except where otherwise noted.

Icons and stock photos licensed under Microsoft Premium Content and cannot be reused outside this document.