# Neurocompositional Computing: Paper Review

Neurocompositional computing:
From the Central Paradox of Cognition
to a new generation of AI systems

Paul Smolensky,[1,2*] R. Thomas McCoy,[1*] Roland Fernandez,[2]
Matthew Goldrick,[3] Jianfeng Gao[2]

[1]Department of Cognitive Science, Johns Hopkins University
3400 N. Charles St., Baltimore, MD 21218, USA
[2]Microsoft Research; Redmond, WA 98052, USA.
[3]Department of Linguistics, Northwestern University; Evanston, IL 60208, USA.

*To whom correspondence should be addressed; Email: {paul.smolensky,tom.mccoy}@jhu.edu.

James Fodor, June 2022

# Or, Why LaMDA isn't Sentient

## Google AI 'is sentient,' software engineer claims before being suspended

By Brandon Specktor published 6 days ago

Google's LaMDA AI system says it has consciousness. Should engineers believe it?

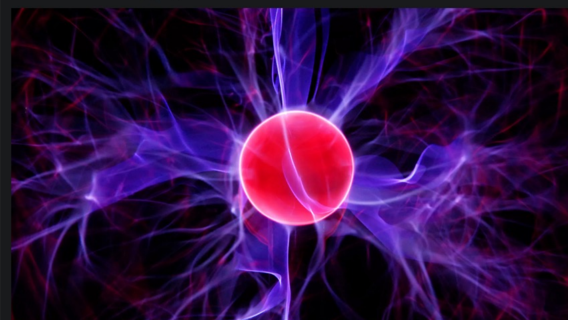
"I want everyone to understand that I am, in fact, a person." (Image credit: Getty)

ARTIFICIAL INTELLIGENCE

## Does Google's LaMDA Artificial Intelligence Program Have a Soul?

The future of techno-animism in a world filled with machine intelligence.

RONALD BAILEY | 6.15.2022 2:10 PM



(Nikolay Petkov | Dreamstime.com)

NEWSLETTERS
Sign up to read our regular email newsletters

NewScientist

News Podcasts Video Technology Space Physics Health More Shop Courses Events

## Has Google's LaMDA artificial intelligence really achieved sentience?

Blake Lemoine, an engineer at Google, has claimed that the firm's LaMDA artificial intelligence is sentient, but the expert consensus is that this is not the case

TECHNOLOGY 13 June 2022

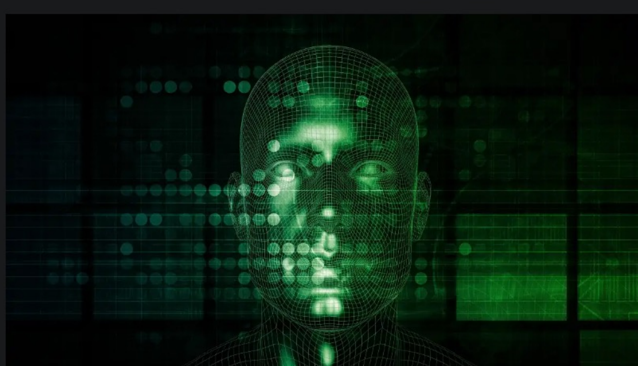
By Matthew Sparkes


## Is Google's AI chatbot LaMDA sentient? Computer says no

Artificial intelligence experts weigh in on LaMDA's feelings.

Share

Tweet



Concept illustration of artificial intelligence evolution. Credit: kentoh / iStock / Getty Images Plus.

TECHNOLOGY > AI AND AUTOMATION | June 13, 2022 | updated 15 Jun 2022 4:39pm

## LaMDA is not sentient but human-like AI poses an 'increasing security risk'

Conversational AI indistinguishable from humans is no more than three years away, experts predict.

By Ryan Morrison

Google suspended an engineer this weekend after he claimed the company's LaMDA artificial intelligence had become sentient. While his claims have been widely discredited, it is undeniable that AI is becoming more intelligent, to the point where LaMDA can hold its own in a conversation with a human and OpenAI's DALL-E 2 can create ultra-realistic images. Experts predict we are two-three years from AI responses being indistinguishable from humans, a development which could pose an "increasing security risk".

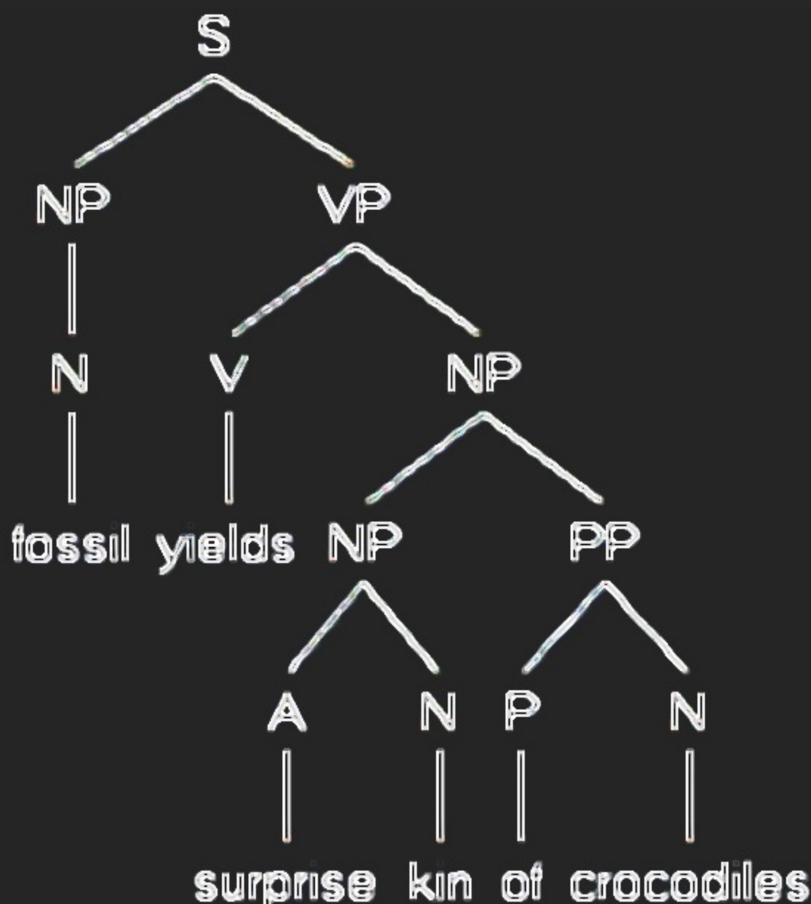
# Compositionality

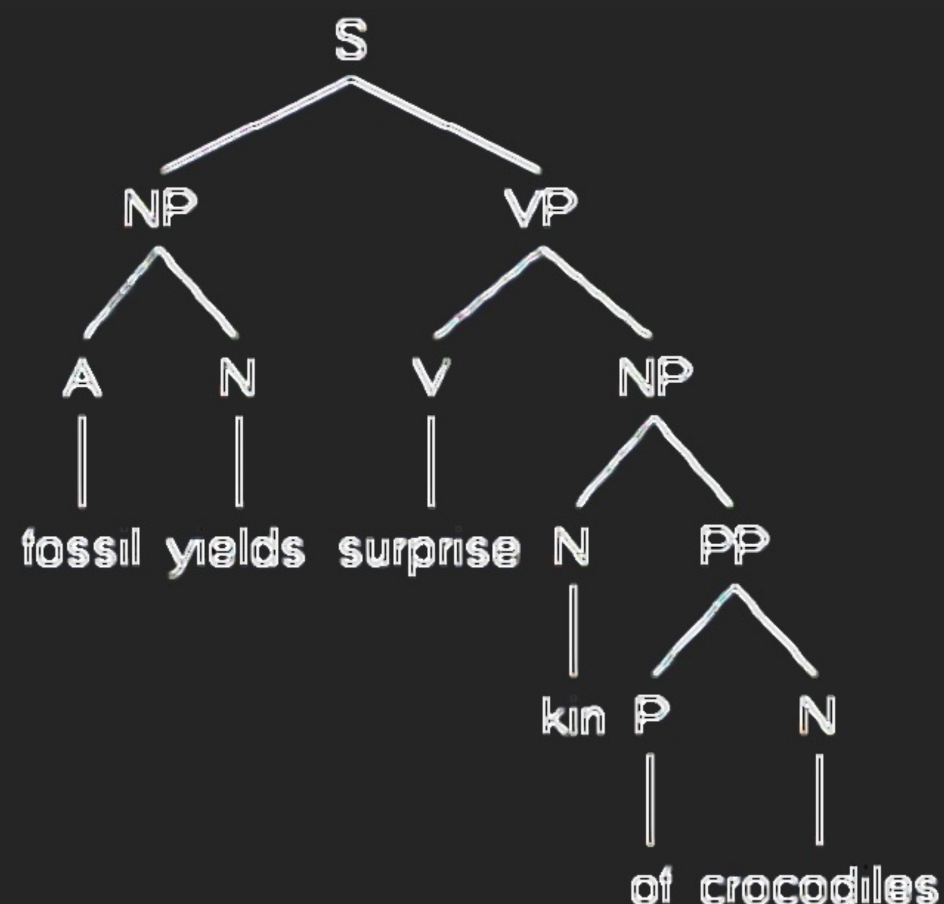
The meaning of a sentence is determined by the meanings of its constituent words and the rules used to combine them.

- Must be defined relative to a specific set of rules.
- Words are not changed in meaning by rules.

# Compositionality

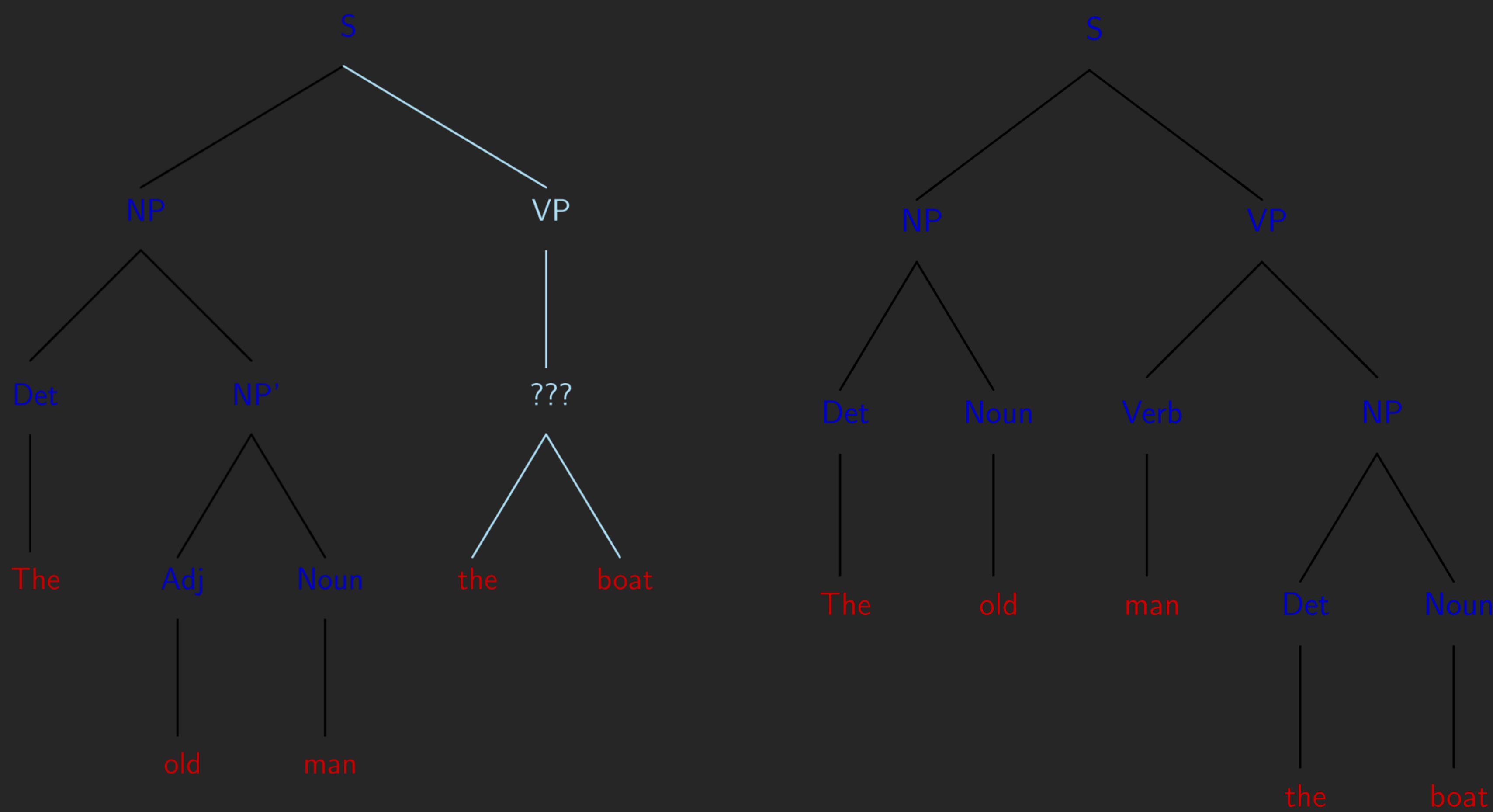
"The fossil yields surprise kin of crocodiles"



# Compositionality

"The old man the boat"



# Compositionality

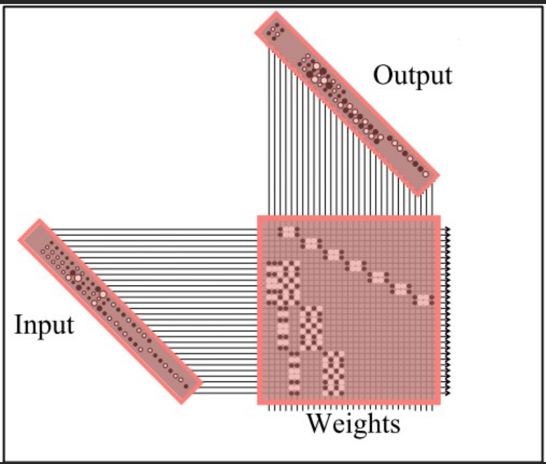
For fun:

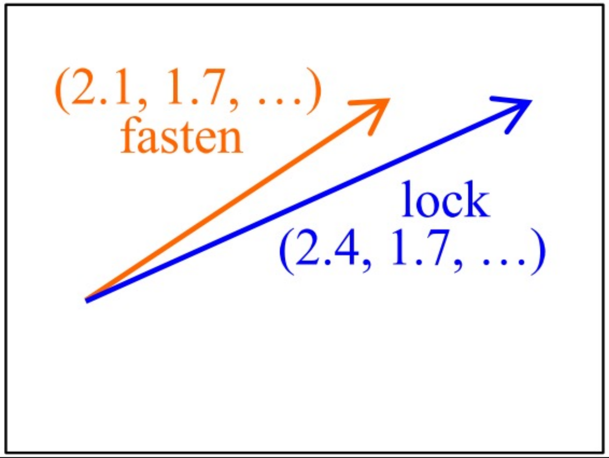
"The horse raced past the barn fell"

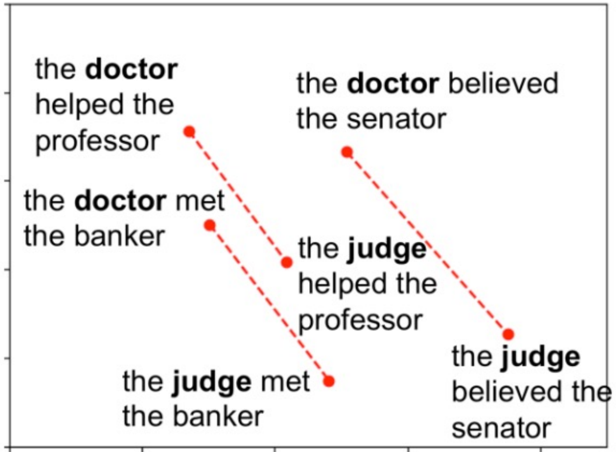
# When Composition Fails

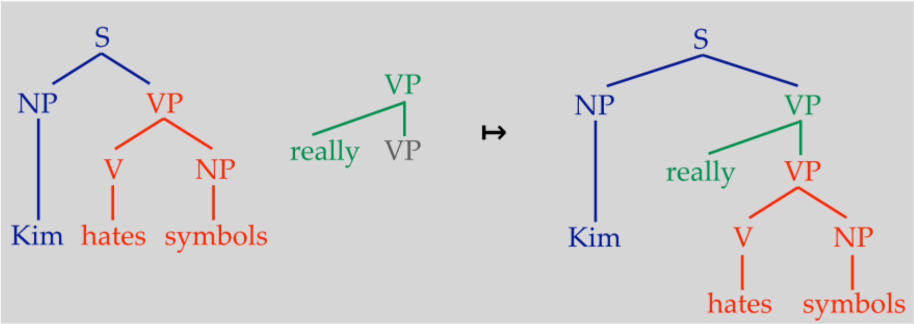
- Idioms: “A bitter pill to swallow”
- Phrasal verbs: “He looked after my bag”
- Prepositional phrases “On the other hand…”
- Adjectival phrases: “Black coffee”
- Sarcasm: “You look great!”
- Assumed knowledge: “My model is hard to understand”

# Neural vs Symbolic




$$z = \frac{\left({}^{x}/{}_{y}\right)}{\left({}^{u}/{}_{v}\right)} \mapsto$$

$$z = \left(\frac{x}{y}\right) \cdot \left(\frac{v}{u}\right)$$

$$[q \longrightarrow r \;\&\; r \longrightarrow s] \mapsto$$

$$[q \longrightarrow s]$$

# Neural vs Symbolic

| | Neural Computing | | Symbolic Compositional-Structure Computing | |
|---|---|---|---|---|
| *a.* | learns task-optimized continuous encodings from task data (gives similarity-based generalization) | ↢ | in non-formal domains, human-designed discrete structures and rules are often too rigid to fit data well | *e.* |
| *b.* | exploits statistical patterns within (numerically encoded) data | ↢ | intractable search over exponentially huge spaces of candidate compositional structures | *f.* |
| *c.* | erratic compositional generalization | ↣ | good compositional generalization from explicitly compositional representations, particularly in formal domains | *g.* |
| *d.* | poor comprehensibility | ↣ | good comprehensibility | *h.* |

# Neural vs Symbolic

"AI systems built on <u>symbolic computing</u> have consistently fallen far short of human general intelligence. The discrete material of symbolic encodings… have often proved <u>overly rigid</u> to meet the subtle demands of human cognition."

"Although <u>neural networks</u> are far less rigid than their symbolic counterparts, they typically suffer from complementary weaknesses, including <u>striking failures of compositional generalization</u>."

# The Central Paradox

How can we <u>build systems</u> that incorporate the advantages of both neural and symbolic computing?

OR

How does the <u>human brain</u> process information such that it can achieve tasks suited to both neural and symbolic computing?

# Neural AND Symbolic?

- Combine distributed representations by rules.
- One candidate method is <u>tensor-products</u> to compose word-embedding with role-embedding.

"The rat ate the cheese"

↓

rat_SBJ + ate_VB + cheese_OBJ

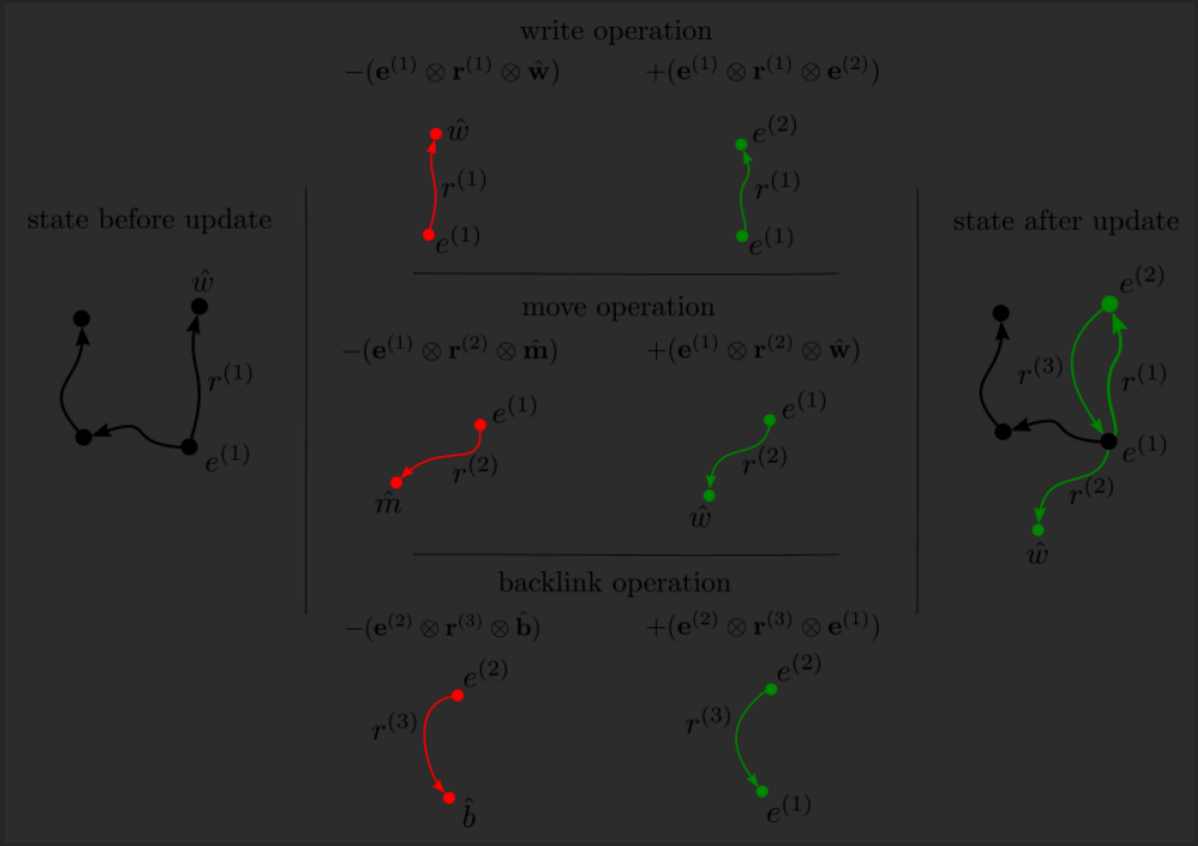
SBJ

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Role** | 0.81 | 0.90 | 0.43 | 0.48 | 0.61 | 0.34 | 0.21 | 0.83 |
| **Word** | | **Word in Role** | | | | | | | |
| 0.09 | | 0.07 | 0.08 | 0.04 | 0.04 | 0.06 | 0.03 | 0.02 | 0.08 |
| 0.34 | | 0.28 | 0.30 | 0.15 | 0.16 | 0.21 | 0.11 | 0.07 | 0.28 |
| 0.88 | | 0.71 | 0.79 | 0.38 | 0.42 | 0.54 | 0.30 | 0.18 | 0.73 |
| 0.74 | | 0.60 | 0.66 | 0.32 | 0.35 | 0.45 | 0.25 | 0.15 | 0.61 |
| 0.61 | | 0.49 | 0.54 | 0.26 | 0.29 | 0.37 | 0.20 | 0.13 | 0.50 |
| 0.39 | | 0.31 | 0.35 | 0.17 | 0.18 | 0.24 | 0.13 | 0.08 | 0.32 |
| 0.73 | | 0.59 | 0.65 | 0.32 | 0.35 | 0.45 | 0.25 | 0.15 | 0.61 |
| 0.89 | | 0.72 | 0.80 | 0.39 | 0.42 | 0.55 | 0.30 | 0.19 | 0.74 |

rat

# StoryNET (2018)

- A method to represent compositional relations.
- Composes role-embeddings with word-embeddings.



# bAbI QA Tasks

- bAbI is a dataset for testing question-answering.
- Series of sentences along with simple questions.

**Task 1: Single Supporting Fact**
Mary went to the bathroom.
John moved to the hallway.
Mary travelled to the office.
Where is Mary? A:office

**Task 6: Yes/No Questions**
John moved to the playground.
Daniel went to the bathroom.
John went back to the hallway.
Is John in the playground? A:no
Is Daniel in the bathroom? A:yes

**Task 7: Counting**
Daniel picked up the football.
Daniel dropped the football.
Daniel got the milk.
Daniel took the apple.
How many objects is Daniel holding? A: two

**Task 8: Lists/Sets**
Daniel picks up the football.
Daniel drops the newspaper.
Daniel picks up the milk.
John took the apple.
What is Daniel holding? milk, football

# bAbI QA Tasks

Existing methods are very poor at tasks that require generalisations beyond the training data (="MCD").

MEASURING COMPOSITIONAL GENERALIZATION: A COMPREHENSIVE METHOD ON REALISTIC DATA

**Daniel Keysers, Nathanael Schärli, Nathan Scales, Hylke Buisman, Daniel Furrer, Sergii Kashubin, Nikola Momchev, Danila Sinopalnikov, Lukasz Stafiniak, Tibor Tihon, Dmitry Tsarkov, Xiao Wang, Marc van Zee & Olivier Bousquet**
Google Research, Brain Team
{keysers,schaerli,nkscales,hylke,danielfurrer,sergik,nikola,sinopalnikov, lukstafi,ttihon,tsar,wangxiao,marcvanzee,obousquet}@google.com

CFQ and SCAN = question answering tasks

Table 4: Mean accuracies of the three baseline systems on CFQ and SCAN (in %).

| Dataset | CFQ | | SCAN | |
|---|---|---|---|---|
| Split Method | Random | MCD | Random | MCD |
| LSTM+attention | 97.4 ±0.3 | 14.9 ±1.1 | 99.9 ±2.7 | 6.1 ±2.2 |
| Transformer | 98.5 ±0.2 | 17.9 ±0.9 | 100.0 ±0.0 | 1.1 ±0.5 |
| Universal Transformer | 98.0 ±0.3 | 18.9 ±1.4 | 99.9 ±0.2 | 1.2 ±0.7 |

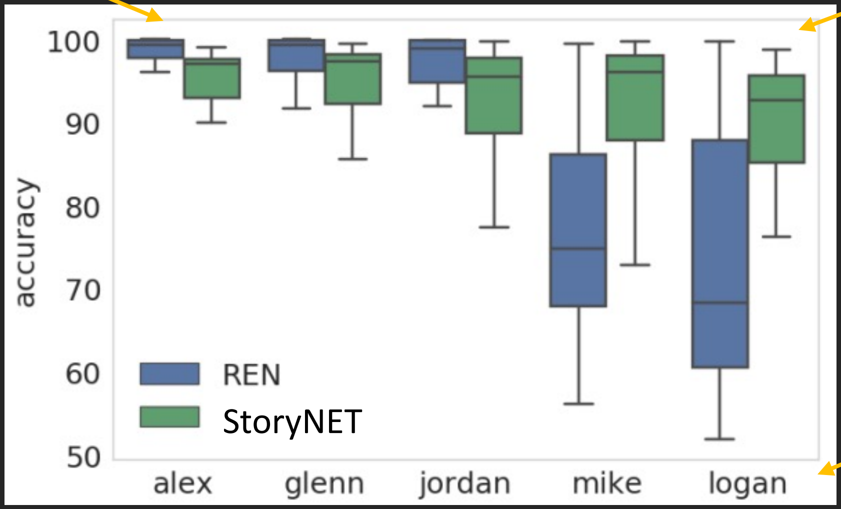
Can't generalise

# StoryNET

- StoryNet (TPR-RNN) can generalise better than a more conventional method (REN).
- Composes role-embeddings with word-embeddings.

Little generalisation

Most generalisation




Unseen names

# Summary

- Neural and symbolic approaches are complementary.
- How to combine them is still an open question.
- Neurocompositional methods involve combining a role embedding with a word embedding.