Can cognitive neuroscience inform neuro-symbolic architectures?

Shashank

Una-May O'Reilly, ALFA lab

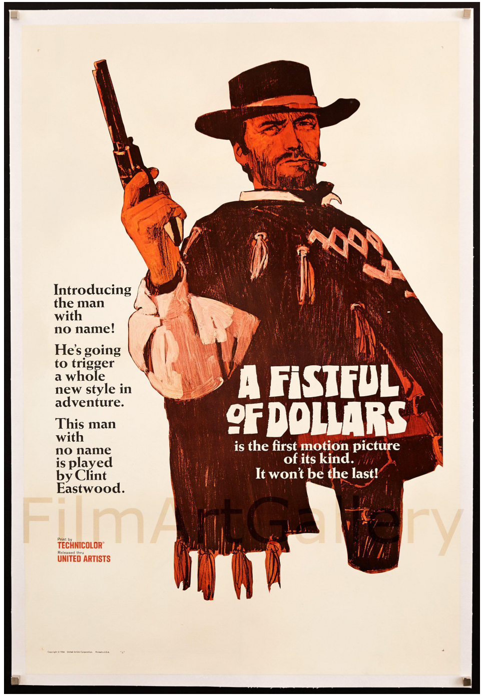
CSAIL, MIT

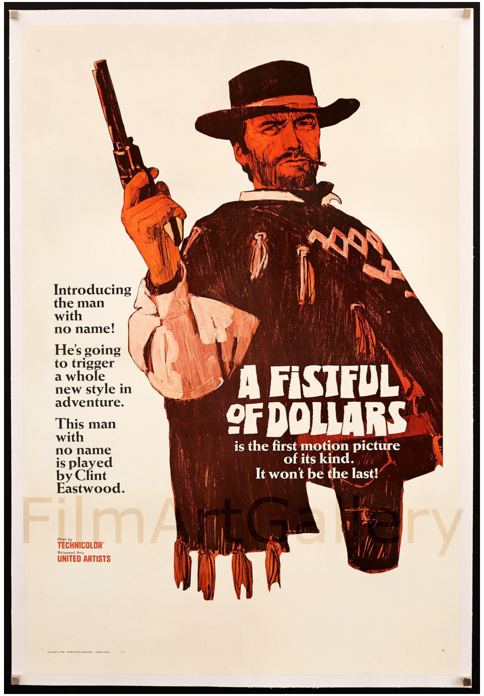
shash@mit.edu

6 May 2021

@ShashankSrikant

A tale of three stories





Human performance on NL inference tasks

Inherent Disagreements in Human Textual Inferences

Ellie Pavlick
Brown University
ellie_pavlick@brown.edu

Tom Kwiatkowski
Google Research
tomkwiat@google.com

Abstract

We analyze human's disagreements about the validity of natural language inferences. We show that, very often, disagreements are not dismissible as annotation ''noise'', but rather persist as we collect more ratings and as we vary the amount of context provided to raters. We further show that the type of uncertainty captured by current state-of-the-art models for natural language inference is not reflective of the type of uncertainty present in human disagreements. We discuss implications of our results in relation to the recognizing textual entailment (RTE)/natural language inference (NLI) task. We argue for a refined evaluation objective that requires models to explicitly capture the full distribution of plausible human judgments.

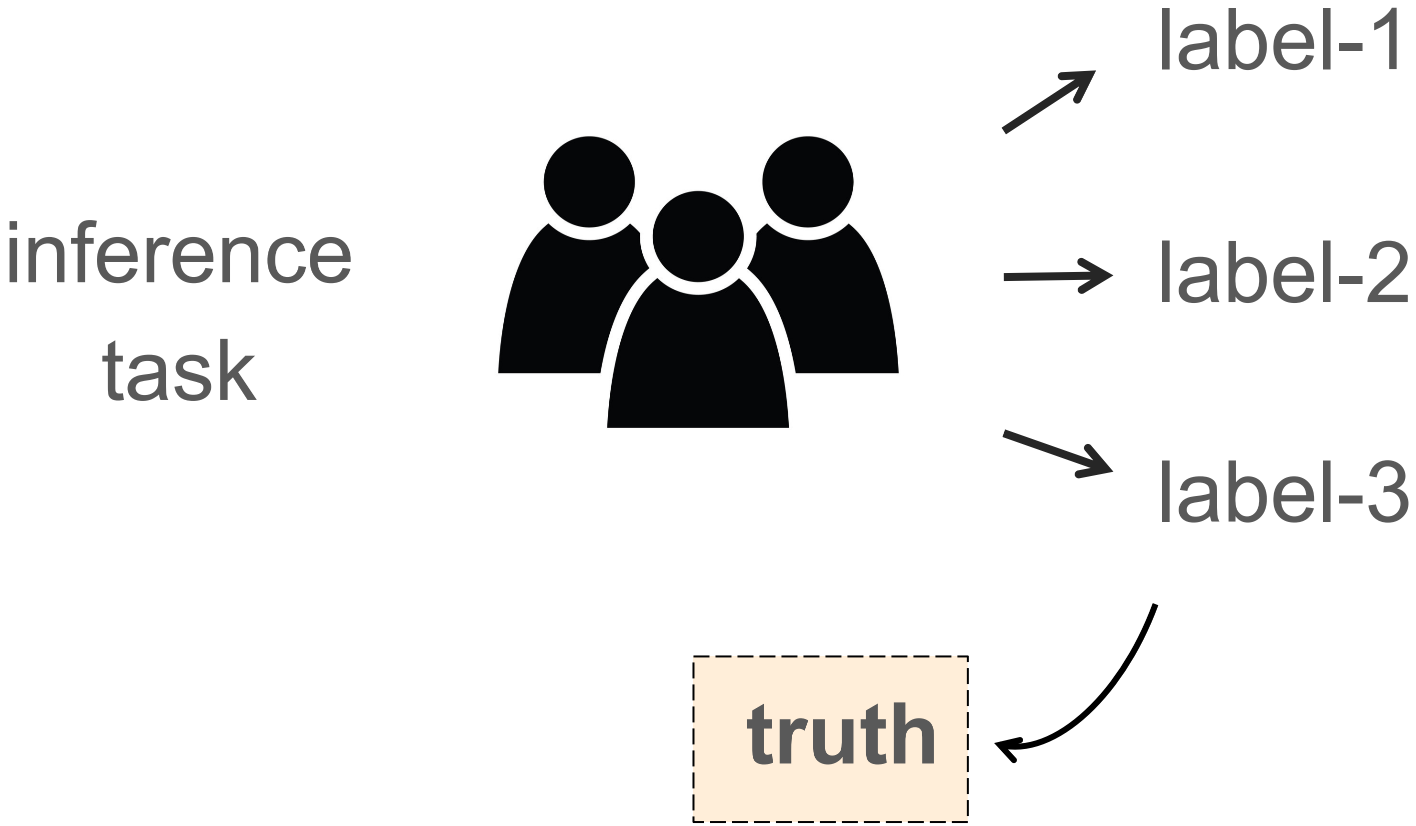
and then seeking some consensus among them. For example, having raters choose among discrete labels and taking a majority vote (Dagan et al., 2006; Bowman et al., 2015; Williams et al., 2018), or having raters use a continuous Likert scale and taking an average (Pavlick and Callison-Burch, 2016a; Zhang et al., 2017). That is, the prevailing assumption across annotation methods is that there is a single ''true'' inference about h given p that we should train models to predict, and that this label can be approximated by aggregating multiple (possibly noisy) human ratings as is typical in many other labelling tasks (Snow et al., 2008; Callison-Burch and Dredze, 2010).

Often, however, we observe large disagreements among humans about whether or not h can be inferred from p (see Figure 1). The goal of this study is to establish whether such disagree-

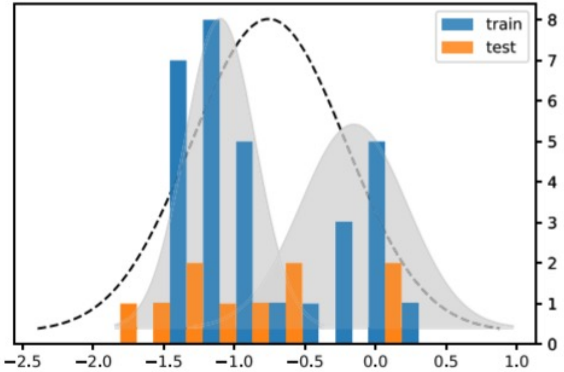
Entailment tasks

Three dogs on a sidewalk. $\rightarrow$ There are more than one dog here.
A red rally car taking a slippery turn in a race. $\rightarrow \neg$ The car is stopped at a traffic light.

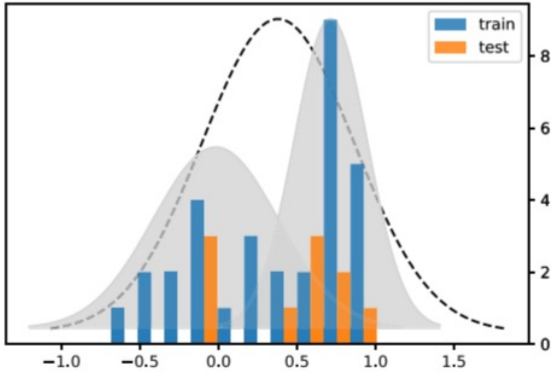
Setup



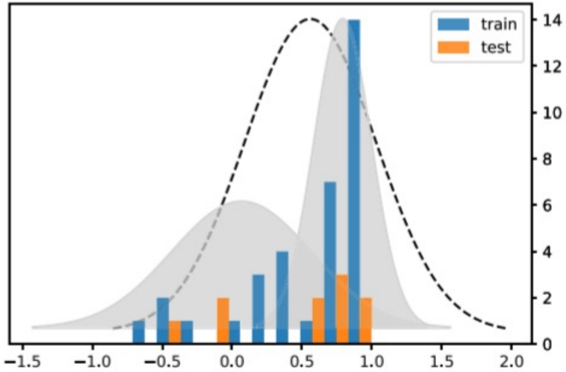
p: A homeless man being observed by a man in business attire.
h: Two men are sleeping in a hotel.



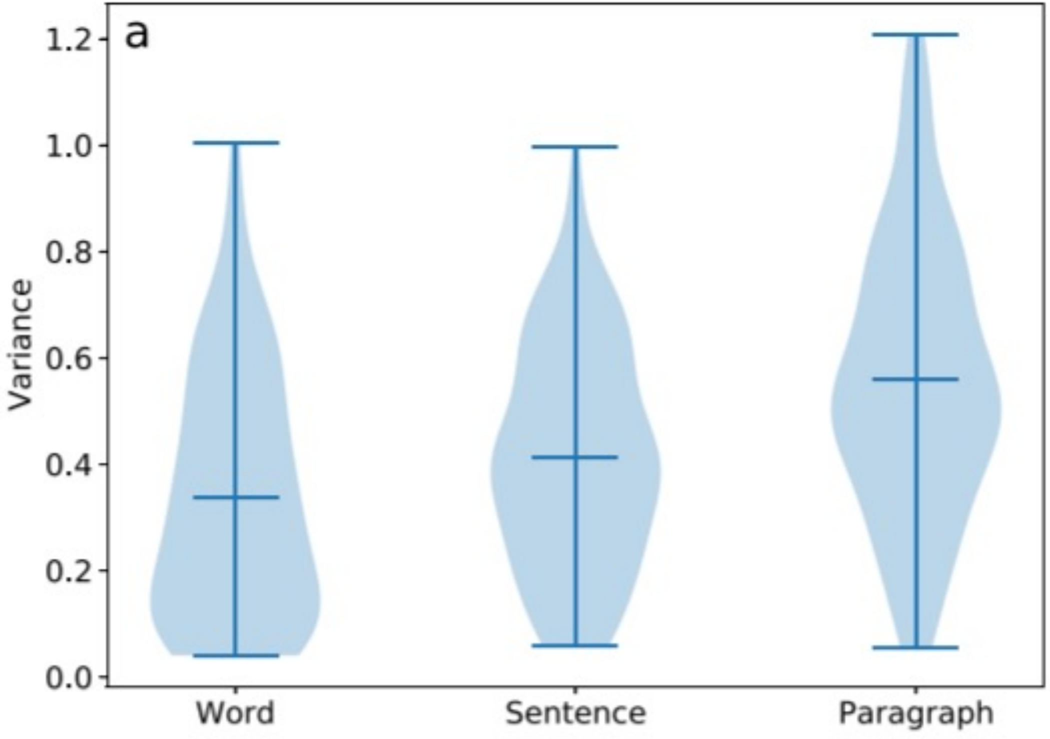
p: Paula swatted the fly.
h: The swatting happened in a forceful manner.



p: Someone confessed that a particular thing happened.
h: That thing happened.



Does context play a role then?



Do NL models display similar ambiguity?

- softmax, cross-entropy heads naturally provide a distribution, and not a point estimate

- Perhaps then human judgment mimics these heads?

Do NL models display similar ambiguity?

- softmax, cross-entropy heads naturally provide a distribution, and not a point estimate

- Perhaps then human judgment mimics these heads?

NO

Well then?

- How humans infer – not straightforward

- Making NL models infer like humans – not straightforward

Something more that we’re missing?

Understanding the human brain

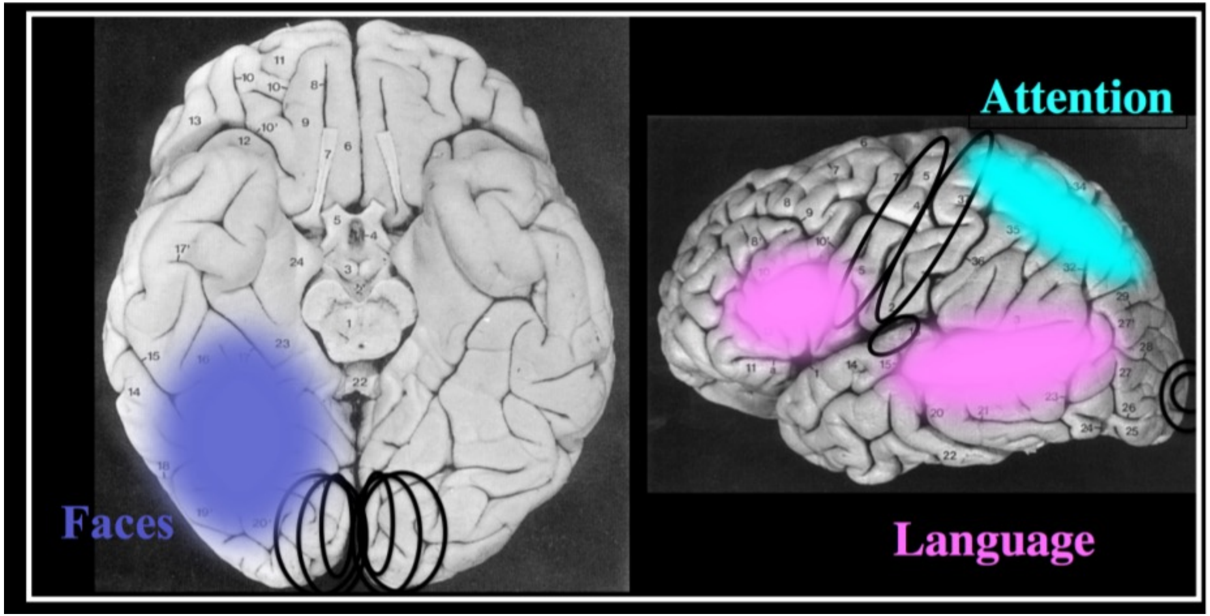
Understanding the human brain

Broad functions

- Vision
- Audio
- Motor control and dexterity
- Emotions
- Language

Understanding the human brain

Early 90s



Slide adapted from Nancy Kanwisher's course on The Human Brain (9.17, 2019)

Understanding the human brain

Broad functions

- Vision
- Audio
- Motor control and dexterity
- Emotions
- Language

Understanding the human brain

Broad functions

- **Vision**
- Audio
- Motor control and dexterity
- Emotions
- Language

Faces

Color

Places

Words/letters

Bodies

Motion

Shape

Understanding the human brain

Broad functions

- Vision
- Audio
- Motor control and dexterity
- Emotions
- **Language**

- Responds to both comprehension and production
- Across modalities (speech, written, ASL)
- Responds to typologically diverse languages
- Causally important for language

Understanding the human brain

Broad functions

- Vision
- Audio
- Motor control and dexterity
- Emotions
- Language
- **Multiple Demand system**

Broadly recruited in math, logic, reasoning, learning like tasks

JNeurosci
THE JOURNAL OF NEUROSCIENCE

Research Report: Regular Manuscript

The domain-general multiple demand (MD) network does not support core aspects of language comprehension: a large-scale fMRI investigation

https://doi.org/10.1523/JNEUROSCI.2036-19.2020

Cite as: J. Neurosci 2020; 10.1523/JNEUROSCI.2036-19.2020

Received: 20 August 2019
Revised: 2 March 2020
Accepted: 5 April 2020

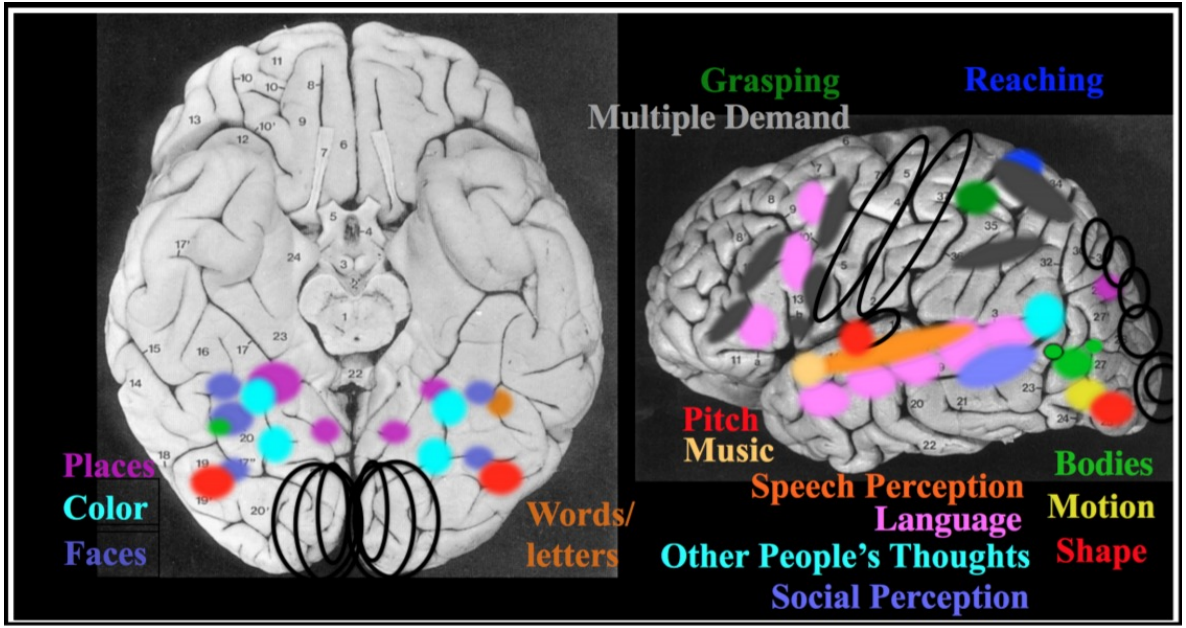
Authors
Evgeniia Diachek*[1], Idan Blank*[2], Matthew Siegelman*[3], Josef Affourtit[4], and Evelina Fedorenko[4,5,6]

Affiliations
[1]Department of Psychology, Vanderbilt University, Nashville, TN 37203, USA
[2]Department of Psychology, UCLA, Los Angeles, CA 90095, USA
[3]Department of Psychology, Columbia University, New York, NY 10027, USA
[4]Department of Brain and Cognitive Sciences, MIT, Cambridge, MA 02139, USA
[5]McGovern Institute for Brain Research, MIT, Cambridge, MA 02139, USA
[6]Department of Psychiatry, MGH, Charlestown, MA 02129, USA

*Equal contributors

Understanding the human brain

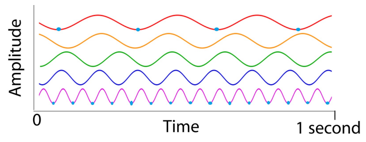
Current understanding



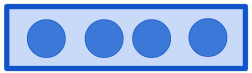
Slide adapted from Nancy Kanwisher's course on The Human Brain (9.17, 2019)

Computer programs and the human brain

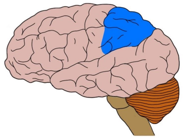
Neural representation



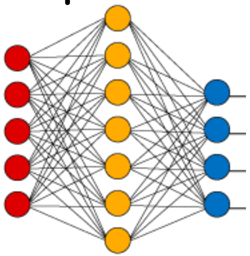
Machine representation



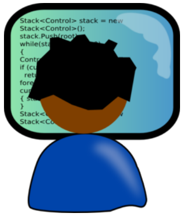
COGNITIVE NEUROSCIENCE



MACHINE LEARNING



Understanding programs



HelloWorld — vim helloworld.c — 43×8

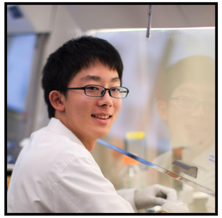
```
#include <stdio.h>

int main(int argc, char **argv) {
        printf("Hello, World!\n");
        return 0;
}

                           7,0-1          All
```

MIT CSAIL

brain+cognitive sciences

Tufts UNIVERSITY
Eliot-Pearson Department of Child Study and Human Development

Anna Ivanova



Yotaro Sueoka

Hope Kean

Riva Dhamala

Ev Fedorenko

Una-May O’Reilly

Marina Bers

Some slides adapted from Nancy Kanwisher’s course on The Human Brain (9.17, 2019)

Where to look?

Broad functions

- Vision
- Audio
- Motor control and dexterity
- Emotions
- Language
- Multiple Demand system

Understanding code?

Where to look?

Broad functions

- Vision
- Audio
- Motor control and dexterity
- Emotions
- Language
- Multiple Demand system

Understanding code

1. Vision system activated

Where to look?

Broad functions

- Vision
- Audio
- Motor control and dexterity
- Emotions
- Language
- Multiple Demand system

Understanding code

1. Vision system activated
2. Recognize characters, tokens to form statements and blocks.

Where to look?

Broad functions

- Vision
- Audio
- Motor control and dexterity
- Emotions
- Language
- Multiple Demand system

Understanding code

1. Vision system activated
2. Recognize characters, tokens to form statements and blocks.
3. Understand what the code does

Where to look?

Broad functions

- Vision
- Audio
- Motor control and dexterity
- Emotions
- Language
- Multiple Demand system

Understanding code

1. Vision system activated
2. Recognize characters, tokens to form statements and blocks.
3. Understand what the code does
4. Mentally trace it/debug it and calculate output.

Where to look?

Broad functions

- Vision
- Audio
- Motor control and dexterity
- Emotions
- Language
- Multiple Demand system

Understanding code

1. Vision system activated
2. Recognize characters, tokens to form statements and blocks.
3. Understand what the code does
4. Mentally trace it/debug it and calculate output.

Where to look?

Broad functions

- Vision
- Audio
- Motor control and dexterity
- Emotions
- Language
- Multiple Demand system

Understanding code

1. Vision system activated
2. Recognize characters, tokens to form statements and blocks.
3. Understand what the code does
4. Mentally trace it/debug it and calculate output.

Where to look?

Broad functions

- Vision
- Audio
- Motor control and dexterity
- Emotions
- Language
- Multiple Demand system

Code simulation

1. Vision system activated
2. Recognize characters, tokens to form statements and blocks.
3. Understand what the code does
4. Mentally trace it/debug it and calculate output.

Where to look?

Broad functions

- Vision
- Audio
- Motor control and dexterity
- Emotions
- Language
- Multiple Demand system

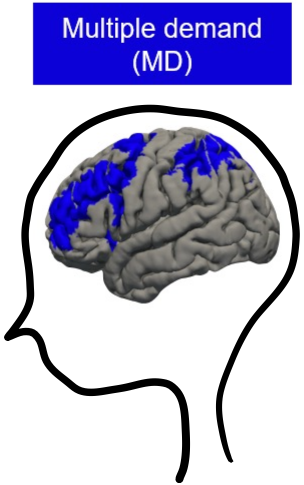
Code reading

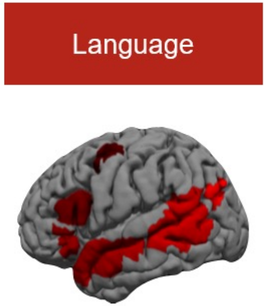
1. Vision system activated
2. Recognize characters, tokens to form statements and blocks.
3. Understand what the code does
4. Mentally trace it/debug it and calculate output.

Where to look?

Broad functions

- Vision
- Audio
- Motor control and dexterity
- Emotions
- Language
- Multiple Demand system





Disambiguate reading and simulation

code

```
big_num, small_num = 64, 16

if big_num % small_num == 0:
    print(1)
else:
    print(0)
```

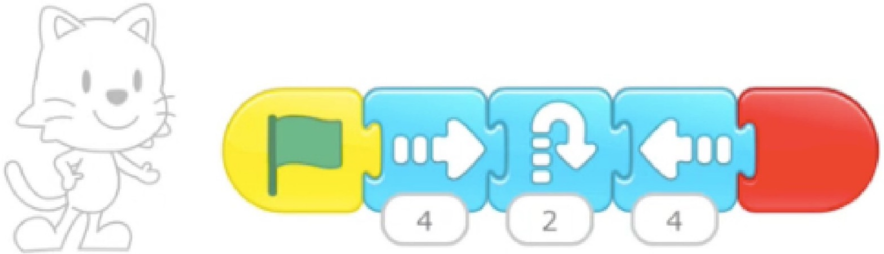
```
filename = "alphabet.java"
modified = filename.split(".")

print(modified[-1])
```

sent

You are given two numbers 64 and 16. If the remainder when the first number is divided by the second number is 0, you perform one good deed. Otherwise, you perform no good deeds. How many good deeds will you perform?

A file is named "alphabet.java". You split the name at the dot character. What is the last part of the resulting split?

python



	MD system	**Language system**
PYTHON	✓	✓
SCRATCH JR	✓	✗
	Strong, generalizable responses to code	**Moderate, task/language dependent (?) responses to code**

Understanding code

1. Vision system activated
2. Recognize characters, tokens to form statements and blocks.
3. Understand what the code does
4. Mentally trace it/debug it and calculate output.

So..

Modeling the distinction between MD and language system may be worth exploring in computational models

What’s the human baseline in NL tasks? Are we as good as we want our models to be? Worth exploring, and perhaps NS systems may help model that uncertainty.

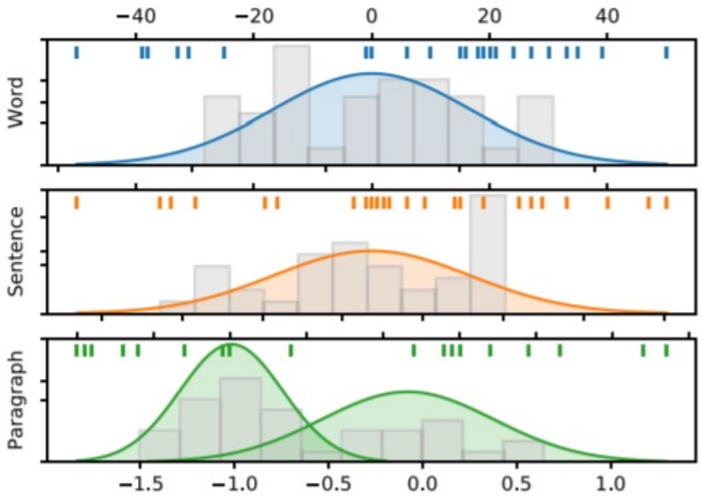
shash@mit.edu

@ShashankSrikant

https://shashank-srikant.github.io/

A watercolor painting celebrating that event hangs today in the Chenango Museum in Norwich. The canal itself was also utilized for recreation. In the summer months it supported swimming, **boating** and fishing. In the winter months, after the surface froze over, ice skating and even horse racing became favorite pastimes. Before the Chenango Canal was built, much of the Southern Tier and Central New York was still considered to be frontier.

In the summer months it supported swimming , **picnicking** and fishing .

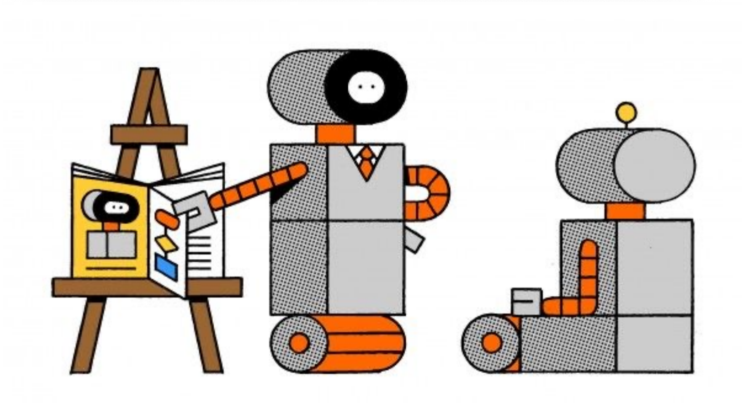


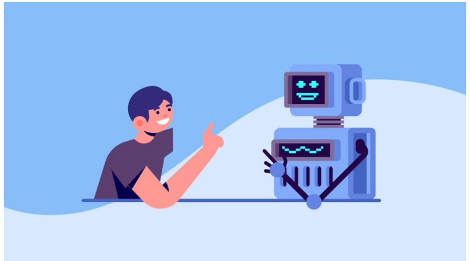
Programming language

HelloWorld — vim helloworld.c — 43×8

```c
#include <stdio.h>

int main(int argc, char **argv) {
        printf("Hello, World!\n");
        return 0;
}
```

7,0-1 All





Teaching machines to read code