Marco Maru

On Meaning, Understanding, and Other Mythical Creatures in Natural Language Processing

What's inside

SAPIENZA NLP

What's inside

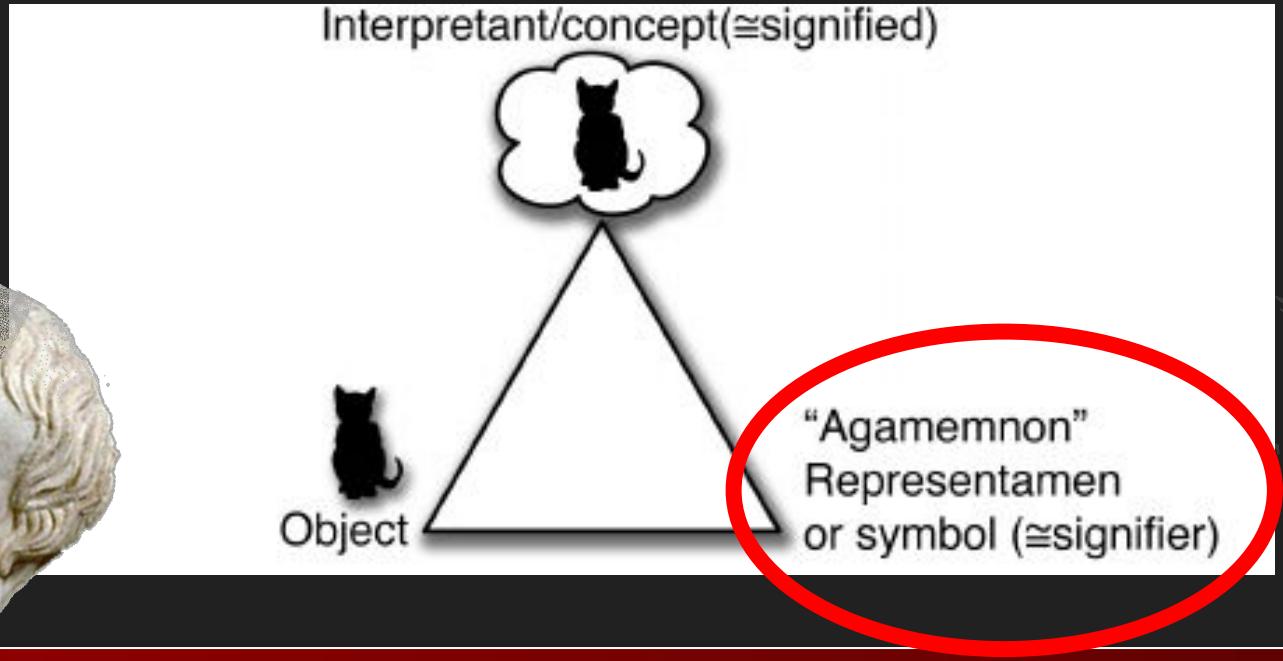
What's inside

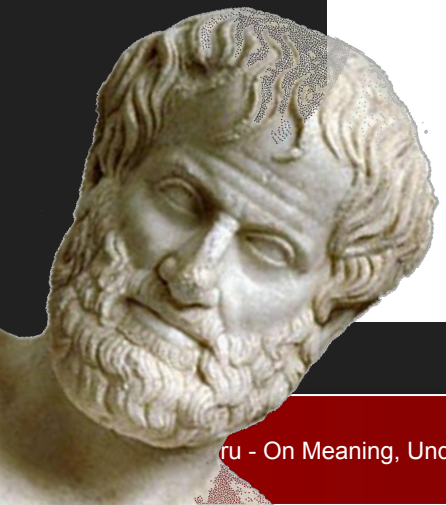
What's inside

- Are we properly using terms such as "*meaning*" and "*understanding*"?
- What stage of the NLU path are we at?

What does the meaning of "*meaning*" means?

The triangle of reference [Ogden and Richards, 1923]





What does the meaning of "*meaning*" means?

Conventional Meaning vs Conversational Implicatures
[Grice, 1968]

Husband: "How much longer will you be?"

Wife: "Mix yourself a drink."

What does the meaning of "*meaning*" means?

"Linguistic choices subtly color meaning" [Trott et al., 2020]

A. "I work **at** Apple."

B. "I work **for** Apple."

What does the meaning of "*meaning*" means?

Construal is a dynamic process of meaning construction, in which speakers and hearers encode and decode, respectively, some intended meaning in a given communicative context.

A. **Prominence** (relative attention)

B. **Resolution** (level of description)

C. **Configuration** (internal-structural properties of entities)

D. **Metaphor** (cross-domain comparison)

Spoiler

A system trained only on form (i.e. symbols, signifiers)
has *a priori* no way to learn meaning.

Overclaims and consequences

"In order to train a model that **understands** sentence relationships, we pre-train for a binarized next sentence prediction task."
[Devlin et al., 2019]

"BERT is a system by which Google's algorithm uses pattern recognition to better **understand** how human beings communicate so that it can return more relevant results for users."
[www.business2community.com]

SAPIENZA NLP

LMs and meaning sensitive tasks

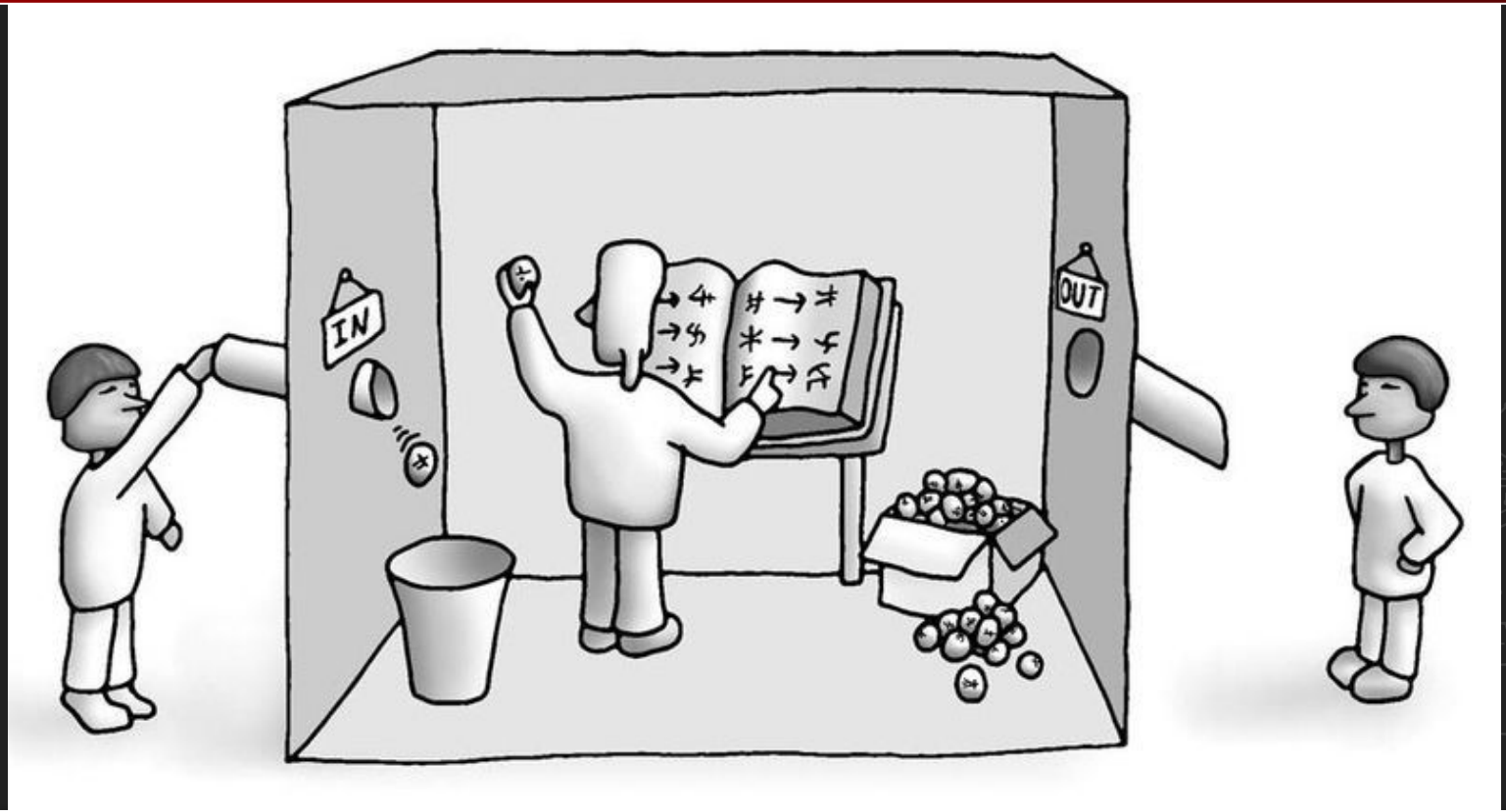
Why are LMs so good at tackling meaning sensitive tasks?

Are they performing some sort of "*reasoning*"?

NO. They're simply more effective at leveraging artifacts in the data.

[Niven and Kao, 2019] [McCoy et al., 2019]

Scenario 1: The Chinese Room Experiment [Searle, 1980]

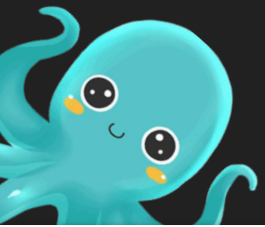


Scenario 2: The Octopus [Bender and Koller, 2020]

Scenario 2: The Octopus [Bender and Koller, 2020]

“Help! I’m being chased by a bear!
All I have is these sticks. What should I do?”

“Here’s what you can do. Take one stick and punch the bear, and then run faster to the store. Wait there until the bear is gone, and then give her all of the sticks. Then go back inside and get your gun. Advertisement.” [GPT-2 demo]



Humans and language acquisition

Language is used for communication about the speakers' actual (physical, social, and mental) world.

Reasoning must connect the meanings of perceived inputs to information about that world.

Human children do not learn meaning from form alone... [Snow et al., 1976]

...and we should not expect machines to do so either.

How to ground (w/ little spoiler)

Train distributional models on corpora augmented with perceptual data, such as photos [Hossain et al., 2019] or other modalities [Kiela and Clark, 2015; Kiela et al., 2015].

A first answer

Are we properly using terms such as “*meaning*” and “*understanding*”?

No.

Large LMs have probably learned something about meaning, in the same sense that syntax captures something about meaning: a potentially useful, but incomplete, reflection of the actual meaning.

SAPIENZA NLP

World Scopes (WS) [Bisk et al., 2020]

A **World Scope** (WS)
is a lens through which to view progress in NLP.

Five levels of World Scope (w/ huge spoiler):

WS1. Corpus *you've been there*

WS2. Internet *you are here*

WS3. Perception *you are (slowly) heading there*

WS4. Embodiment

WS5. Social

World Scope 1: Corpora and Representations

Data-driven language research begins with the corpus.

Brown Corpus [Francis, 1964]
The Penn Treebank [Marcus et al., 1993]

Dense word vectors [Mikolov et al., 2013]

Contextualized pretrained representations [Peters et al., 2018]

World Scope 2: Internet and Web-crawls

Decreased cost of annotations and viable web-crawls
[Baroni et al., 2019]

➔ No constraints on single authors/sources

➔ Large-scale data

➔ Performances advancement

World Scope 2: Internet and Web-crawls

Embeddings and the “parameter race”:

However, there are diminishing returns.

Continuing to expand hardware and data sizes is not the path forward.

Modeling lexical co-occurrence, no matter the scale, is still “modeling the written world."
Models constructed this way blindly search for symbolic co-occurrences void of meaning.

SAPIENZA NLP

World Scope 3: The World of Sights and Sounds

Learning needs perception.

There exists strong evidence that children require grounded sensory input, not just speech, to learn language.
[Sachs et al., 1981; O'Grady, 2005; Vigliocco et al., 2014]

Perception includes auditory, tactile, and visual input.

World Scope 3: The World of Sights and Sounds

Auditory input is necessary for detecting sarcasm, stress, and meaning implied through prosody.

Tactile senses give meaning,
both physical [Sinapov et al., 2014; Thomason et al., 2016]
and abstract, to concepts like *heavy* and *light*.

SAPIENZA NLP

World Scope 3: The World of Sights and Sounds

Models must be able to watch and recognize objects, people, and activities to understand the language describing them.
[Li et al., 2019; Krishna et al., 2017; Yatskar et al., 2016; Perlis, 2016]

How should we learn the meaning, method, and implications of *painting*?
[Charniak, 1977]

SAPIENZA NLP

World Scope 3: The World of Sights and Sounds

Advances in computer vision architectures and modeling have enabled building semantic representations rich enough to interact with natural language.

Farhadi et al., 2010

Antol et al., 2015

Suhr et al., 2019

Zellers et al., 2019

Mitchell et al., 2012

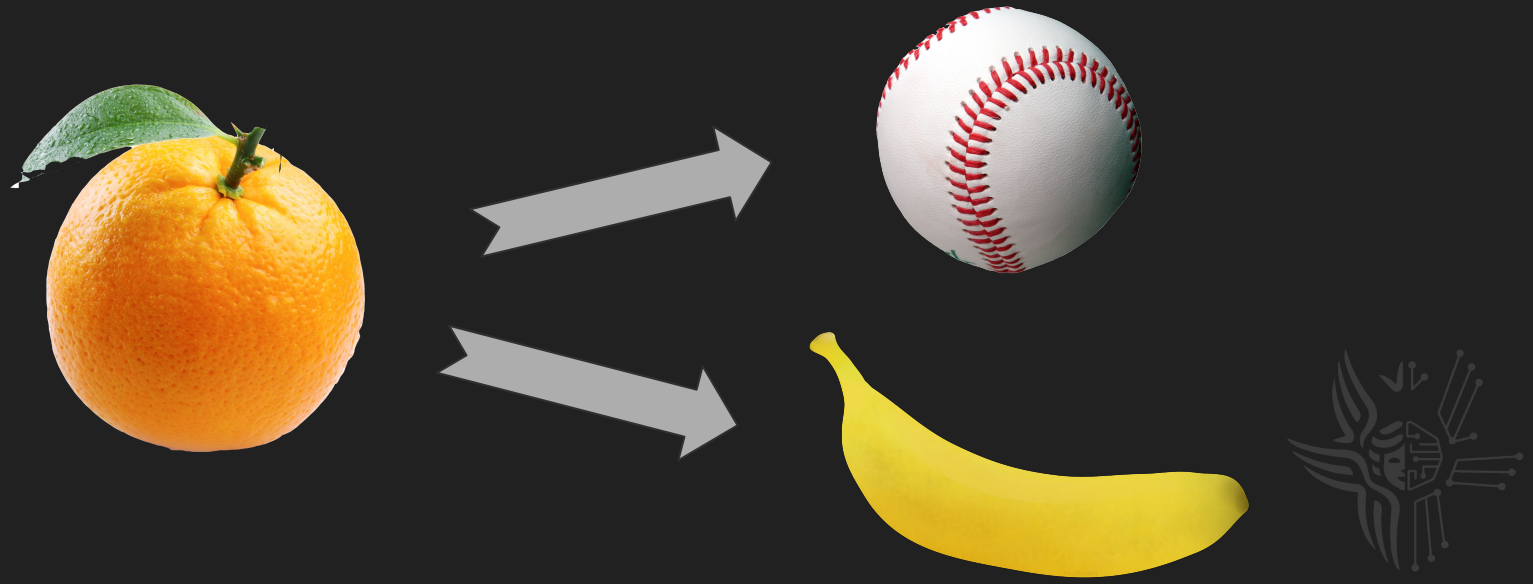
Das et al., 2018

Wang et al., 2019

Benchmarks incorporating auditory, tactile, and visual sensory information together with language will be crucial.

SAPIENZA NLP

World Scope 4: Embodiment and Action



Control is where people first learn abstraction and simple examples of post-conditions through **trial and error**.

World Scope 5: The Social World

In order to learn **the effects language has on the world**, an agent must participate in linguistic activity, such as negotiation [Lewis et al., 2017; He et al., 2018] or emotional support provision [Rashkin et al., 2019].

What is the social meaning of "*hate*"?

It requires social context [Bloom, 2002] and experimentation [Ornaghi et al., 2011] to determine.

SAPIENZA NLP

A second answer

What stage of the NLU path are we at?

While the futility of learning language from an ungrounded linguistic signal is intuitive, our entire field follows this pattern, trying to learn language from the internet.

The need for language to attach to “extralinguistic events" and the requirement for social context should guide our research.

SAPIENZA NLP

Thank you
for your attention!

Further readings:

- Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data [Bender and Koller, 2020]
- Experience grounds language [Bisk et al., 2020]
- Language (Re)modelling: Towards Embodied Language Understanding [Tamari et al., 2020]
- (Re) construing Meaning in NLP [Trott et al., 2020]
- The Unstoppable Rise of Computational Linguistics in Deep Learning [Henderson, 2020]