How Forms Shape Meaning
What Distributional Learning Can and Can't Do in the Era of LLMs

Alessandro Lenci

COmputational LINGuistics Laboratory
Università di Pisa
Dipartimento di Filologia, Letteratura e Linguistica (FiLeLi)

LLMs or the Triumph of Distributional Learning

> A new type of robot called blimp has a kind of artificial arm as one of its parts. The arm has a hook. Is the hook a part of the blimp?

To determine whether the hook is a part of the blimp, we can analyze the given information logically.

1. A blimp has an artificial arm as one of its parts.
2. The arm has a hook.

Given these two points, we can deduce the following:

- The artificial arm is a part of the blimp.
- The hook is a part of the artificial arm.

Since the artificial arm is part of the blimp, and the hook is part of the artificial arm, it follows logically that the hook is indeed part of the blimp.

Therefore, the hook is a part of the blimp.

LLMs as Distributional Learning Machines

LLMs acquire all their "knowledge" from distributional data, by being trained to predict a (written) word given its context

LLMs or the Triumph of Distributional Learning

> A new type of robot called blimp has a kind of artificial arm as one of its parts. The arm has a hook. Is the hook a part of the blimp?

⚙ To determine whether the hook is a part of the blimp, we can analyze the given information logically.

1. A blimp has an artificial arm as one of its parts.
2. The arm has a hook.

Given these two points, we can deduce the following:

- The artificial arm is a part of the blimp.
- The hook is a part of the artificial arm.

Since the artificial arm is part of the blimp, and the hook is part of the artificial arm, it follows logically that the hook is indeed part of the blimp.

Therefore, the hook is a part of the blimp.

LLMs as Distributional Learning Machines

LLMs acquire all their "knowledge" from distributional data, by being trained to predict a (written) word given its context

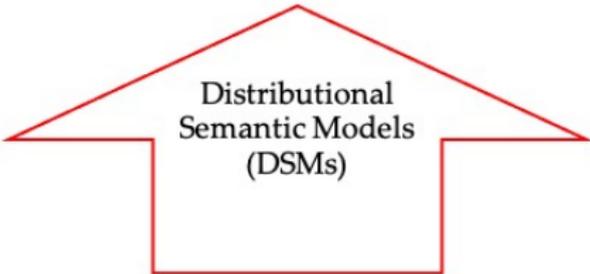# LLMs as Distributional Semantic Models (DSMs)

Lenci & Sahlgren (2023), *Distributional Semantics*, Cambridge University Press

Distributional Semantics

The meaning of linguistic expressions is represented with vectors (embeddings) encoding their statistical distribution in linguistic contexts extracted from textual data
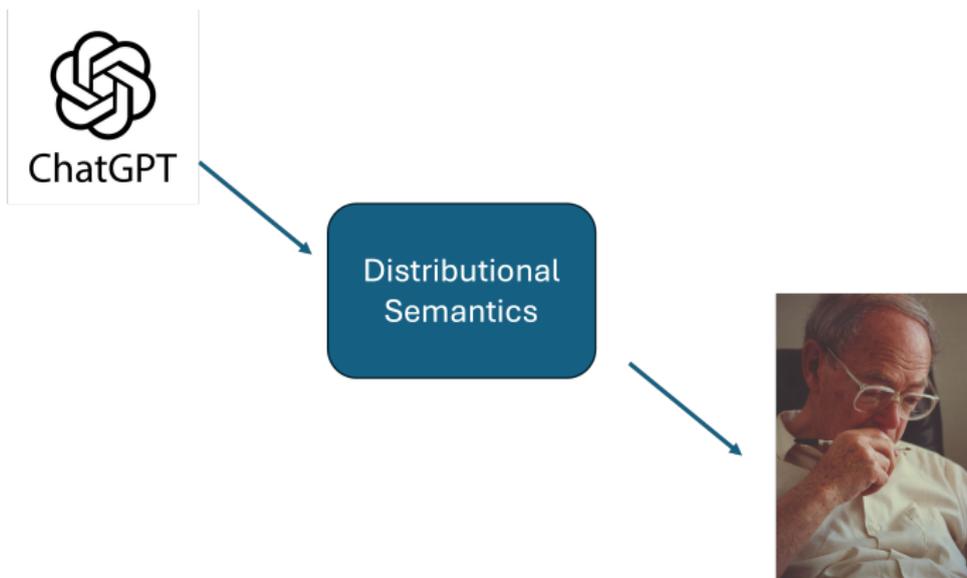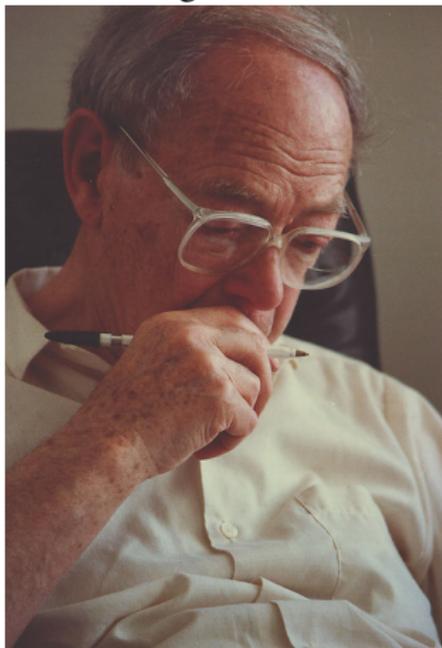

Distributional Semantic Models (DSMs)

... so we went outside, picked several red **cherries** and ate them ... the colour of an orange pink sunset and an indulgent length of rich, red **cherry** fruit with hints of almonds on the dry finish ...

The Tree of LLMs and their Roots

Back to The Roots

The Structuralist Roots of LLMs

Zellig S. Harris

If we consider words or morphemes A and B to be more different in meaning than A and C, then we will often find that the distributions of A and B are more different than the distributions of A and C. In other words, difference in meaning correlates with difference of distribution.

(Harris 1954, "Distributional Structure", *Word*, 156)

What is the Distribution of a Linguistic Element?

"The environment or position of an element consists of the neighborhood, within an utterance, of elements which have been set up on the basis of the same fundamental procedures which were used in setting up the element in question. [...] The distribution of an element is the total of all environments in which it occurs, i.e. the sum of the (different) positions (or occurrences) of an element relative to the occurrence of other elements."

Z. Harris (1951), *Methods in Structural Linguistics*, pp. 15–16

- The elements forming the distributional environments are linguistic signs
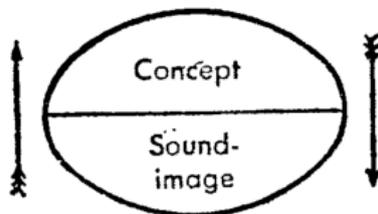
What is the Distribution of a Linguistic Element?

"The environment or position of an element consists of the neighborhood, within an utterance, of elements which have been set up on the basis of the same fundamental procedures which were used in setting up the element in question. [...] The distribution of an element is the total of all environments in which it occurs, i.e. the sum of the (different) positions (or occurrences) of an element relative to the occurrence of other elements."

Z. Harris (1951), *Methods in Structural Linguistics*, pp. 15–16

- The elements forming the distributional environments are linguistic signs

What is a Linguistic Sign?

The linguistic sign is then a two-sided psychological entity that can be represented by the drawing:

F. de Saussure (1916 [1959]), *Course in General Linguistics*, p. 66

- The elements forming the distributional environments are form-meaning pairs

What is a Linguistic Sign?

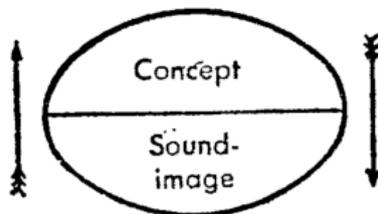The linguistic sign is then a two-sided psychological entity that can be represented by the drawing:

F. de Saussure (1916 [1959]), *Course in General Linguistics*, p. 66

- The elements forming the distributional environments are form-meaning pairs

Distributional Learning in LLMs (and all DSMs)

(1) *The woman ate the pizza with the knife.*

(2) *The woman ate the tipro with the knife.*

(3) *The wopat ate the tipro with the cruc.*

(4) *The wopat patazza the tipro with the cruc.*

(5) *Pe wopat patazza pe tipro ust pe cruc.*

Radical Distributional Learning

LLMs are just like us in front of (5): From their own point of view, they just learn from a huge amount of meaningless sequences of (written) linguistic forms

Distributional Learning in LLMs (and all DSMs)

(1) *The woman ate the pizza with the knife.*

(2) *The woman ate the tipro with the knife.*

(3) *The wopat ate the tipro with the cruc.*

(4) *The wopat patazza the tipro with the cruc.*

(5) *Pe wopat patazza pe tipro ust pe cruc.*

Radical Distributional Learning

LLMs are just like us in front of (5): From their own point of view, they just learn from a huge amount of meaningless sequences of (written) linguistic forms

Distributional Learning in LLMs (and all DSMs)

(1) *The woman ate the pizza with the knife.*

(2) *The woman ate the tipro with the knife.*

(3) *The wopat ate the tipro with the cruc.*

(4) *The wopat patazza the tipro with the cruc.*

(5) *Pe wopat patazza pe tipro ust pe cruc.*

Radical Distributional Learning

LLMs are just like us in front of (5): From their own point of view, they just learn from a huge amount of meaningless sequences of (written) linguistic forms

Distributional Learning in LLMs (and all DSMs)

(1) *The woman ate the pizza with the knife.*

(2) *The woman ate the tipro with the knife.*

(3) *The wopat ate the tipro with the cruc.*

(4) *The wopat patazza the tipro with the cruc.*

(5) *Pe wopat patazza pe tipro ust pe cruc.*

Radical Distributional Learning

LLMs are just like us in front of (5): From their own point of view, they just learn from a huge amount of meaningless sequences of (written) linguistic forms

Distributional Learning in LLMs (and all DSMs)

(1) *The woman ate the pizza with the knife.*

(2) *The woman ate the tipro with the knife.*

(3) *The wopat ate the tipro with the cruc.*

(4) *The wopat patazza the tipro with the cruc.*

(5) *Pe wopat patazza pe tipro ust pe cruc.*

Radical Distributional Learning

LLMs are just like us in front of (5): From their own point of view, they just learn from a huge amount of meaningless sequences of (written) linguistic forms

Distributional Learning in LLMs (and all DSMs)

(1) *The woman ate the pizza with the knife.*

(2) *The woman ate the tipro with the knife.*

(3) *The wopat ate the tipro with the cruc.*

(4) *The wopat patazza the tipro with the cruc.*

(5) *Pe wopat patazza pe tipro ust pe cruc.*

Radical Distributional Learning

LLMs are just like us in front of (5): From their own point of view, they just learn from a huge amount of meaningless sequences of (written) linguistic forms
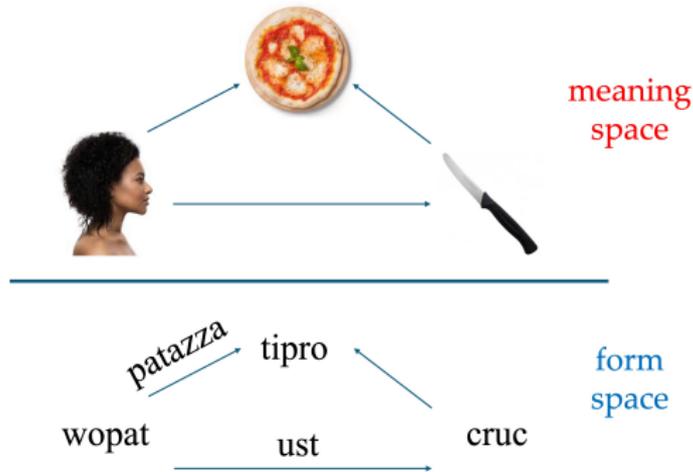
Radical Distributional Learning

meaning space

form space

patazza tipro

wopat   ust   cruc

Constraints on Radical Distributional Learning

Semantic learning from the distribution of linguistic forms is constrained by the degree of second-order isomorphism (Edelman 1998) between the form space and the semantic space

Radical Distributional Learning
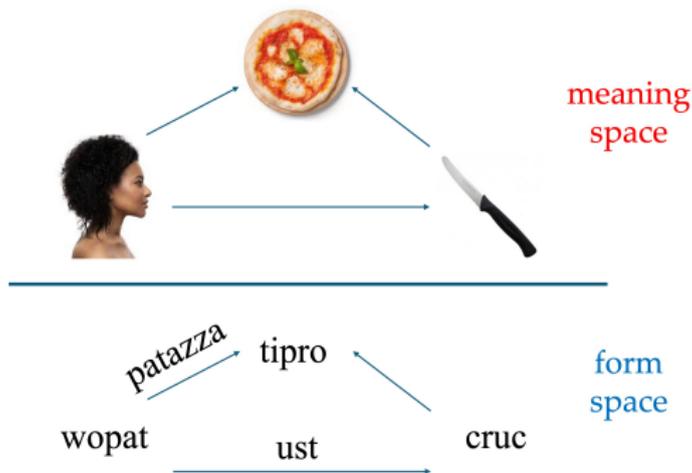

meaning
space

form
space

patazza   tipro

wopat

ust   cruc

Constraints on Radical Distributional Learning

Semantic learning from the distribution of linguistic forms is constrained by the
degree of second-order isomorphism (Edelman 1998) between the form space and the
semantic space

The Semantic Gap

The Semantic Gap of LLMs

LLMs still lack substantial aspects of human semantic competence

- cf. distinction by Mahowald et al. (2024) between formal linguistic competence (i.e., knowledge of linguistic rules and patterns) and functional competence (i.e., understanding and using language in the world)

- LMs have an almost human-like formal competence, but still fall short of functional competence

- LMs as "cultural technology" that only imitates human language production thanks to highly sophisiticated methods of textual data compression (Yiu et al. 2023; also termed as bibliotechnism by Lederman & Mahowald 2024)

The Semantic Gap

The Semantic Gap of LLMs

LLMs still lack substantial aspects of human semantic competence

- cf. distinction by Mahowald et al. (2024) between formal linguistic competence (i.e., knowledge of linguistic rules and patterns) and functional competence (i.e., understanding and using language in the world)
- LMs have an almost human-like formal competence, but still fall short of functional competence
- LMs as "cultural technology" that only imitates human language production thanks to highly sophisiticated methods of textual data compression (Yiu et al. 2023; also termed as bibliotechnism by Lederman & Mahowald 2024)

The Semantic Gap

The Semantic Gap of LLMs

LLMs still lack substantial aspects of human semantic competence

- cf. distinction by Mahowald et al. (2024) between formal linguistic competence (i.e., knowledge of linguistic rules and patterns) and functional competence (i.e., understanding and using language in the world)
- LMs have an almost human-like formal competence, but still fall short of functional competence
- LMs as "cultural technology" that only imitates human language production thanks to highly sophisiticated methods of textual data compression (Yiu et al. 2023; also termed as bibliotechnism by Lederman & Mahowald 2024)

The Semantic Gap

The Semantic Gap of LLMs

LLMs still lack substantial aspects of human semantic competence

- cf. distinction by Mahowald et al. (2024) between formal linguistic competence (i.e., knowledge of linguistic rules and patterns) and functional competence (i.e., understanding and using language in the world)
- LMs have an almost human-like formal competence, but still fall short of functional competence
- LMs as "cultural technology" that only imitates human language production thanks to highly sophisiticated methods of textual data compression (Yiu et al. 2023; also termed as bibliotechnism by Lederman & Mahowald 2024)

Inferential vs. Referential Semantic Competence

D. Marconi (1997), *Lexical Competence*, MIT Press

- Referential competence: "the ability to map lexical items onto the world" (p. 60)

  aardvark

- Inferential competence: "the ability to manage a connection among words, underlying such performances as semantic inference, paraphrase, definition, retrieval a word from its definition, finding a synonym, an so forth" (p. 59)

  - **aardvark** ISA *mammal*; **aardvark** HAS *legs*; etc.

Inferential vs. Referential Semantic Competence

D. Marconi (1997), *Lexical Competence*, MIT Press

- Referential competence: "the ability to map lexical items onto the world" (p. 60)

  aardvark

- Inferential competence: "the ability to manage a connection among words, underlying such performances as semantic inference, paraphrase, definition, retrieval a word from its definition, finding a synonym, an so forth" (p. 59)

  - **aardvark** ISA *mammal*; **aardvark** HAS *legs*; etc.

Inferential vs. Referential Semantic Competence

D. Marconi (1997), *Lexical Competence*, MIT Press

- Referential competence: "the ability to map lexical items onto the world" (p. 60)

aardvark

- Inferential competence: "the ability to manage a connection among words, underlying such performances as semantic inference, paraphrase, definition, retrieval a word from its definition, finding a synonym, an so forth" (p. 59)
    - **aardvark** ISA *mammal*; **aardvark** HAS *legs*; etc.

Inferential vs. Referential Semantic Competence

D. Marconi (1997), *Lexical Competence*, MIT Press

- Referential competence: "the ability to map lexical items onto the world" (p. 60)

aardvark

- Inferential competence: "the ability to manage a connection among words, underlying such performances as semantic inference, paraphrase, definition, retrieval a word from its definition, finding a synonym, an so forth" (p. 59)
  - **aardvark** ISA *mammal*; **aardvark** HAS *legs*; etc.

Do LLMs have an Inferential Semantic Gap?

- Does Radical Distributional Leaning suffice to let LLMs acquire true inferential semantic competence?

- Major issues:
    - is what LLMs acquire a human-like, structured conceptual space?

    - what aspects of inferential semantics can or cannot be learned from distributional data alone?

Do LLMs have an Inferential Semantic Gap?

- Does Radical Distributional Leaning suffice to let LLMs acquire true inferential semantic competence?

- Major issues:
  - is what LLMs acquire a human-like, structured conceptual space?
  - what aspects of inferential semantics can or cannot be learned from distributional data alone?

Do LLMs have an Inferential Semantic Gap?

- Does Radical Distributional Leaning suffice to let LLMs acquire true inferential semantic competence?

- Major issues:
  - is what LLMs acquire a human-like, structured conceptual space?

  - what aspects of inferential semantics can or cannot be learned from distributional data alone?

How do LLMs Represent Meronymy?

Proietti, M. & Lenci A. (2025), "The quasi-semantic competence of LLMs: a case study on the part-whole relation", *arXiv*: 2504.02395

Joint work with Mattia Proietti, Univ. Pisa

- The part-whole relation is an important aspect of inferential competence

Data sources

- **998** part-whole pairs from **McRae Norms** (McRae et al. 2005)
  - features labelled as external or internal components and expressed with the form `holonym_has_meronym`
- **1,026** part-whole pairs from **ConceptNet**
  - items linked with the `partOf` relation to the concepts in the MacRae Norms and in the THINGS database (Hebart et al. 2019)

Mechanistic Interpretability of Meronymy

- We analysed the embedding and unembedding layers of LlaMa2-7b, in order to understand how the *part-whole* relation is encoded in the input and output representations of the LLM

Linear Representation Hypothesis (LRH; Mikolov et al. 2013, Park et al. 2024)

The embedding spaces acquired by distributional semantic models are organized in terms of liner subspaces corresponding to high-level concepts

- word pairs expressing the same relation (e.g., ⟨*wheel*, *car*⟩ and ⟨*wing*, *aircraft*⟩) share similar vector differences
- semantic relations are represented in the embedding spaces as directions defined by the word pairs that express them

Mechanistic Interpretability of Meronymy

- We analysed the embedding and unembedding layers of LlaMa2-7b, in order to understand how the *part-whole* relation is encoded in the input and output representations of the LLM

Linear Representation Hypothesis (LRH; Mikolov et al. 2013, Park et al. 2024)

The embedding spaces acquired by distributional semantic models are organized in terms of liner subspaces corresponding to high-level concepts

- word pairs expressing the same relation (e.g., ⟨*wheel*, *car*⟩ and ⟨*wing*, *aircraft*⟩) share similar vector differences
- semantic relations are represented in the embedding spaces as directions defined by the word pairs that express them

Mechanistic Interpretability of Meronymy

- We analysed the embedding and unembedding layers of LlaMa2-7b, in order to understand how the *part-whole* relation is encoded in the input and output representations of the LLM

Linear Representation Hypothesis (LRH; Mikolov et al. 2013, Park et al. 2024)

The embedding spaces acquired by distributional semantic models are organized in terms of liner subspaces corresponding to high-level concepts

- word pairs expressing the same relation (e.g., $\langle wheel, car \rangle$ and $\langle wing, aircraft \rangle$) share similar vector differences
- semantic relations are represented in the embedding spaces as directions defined by the word pairs that express them
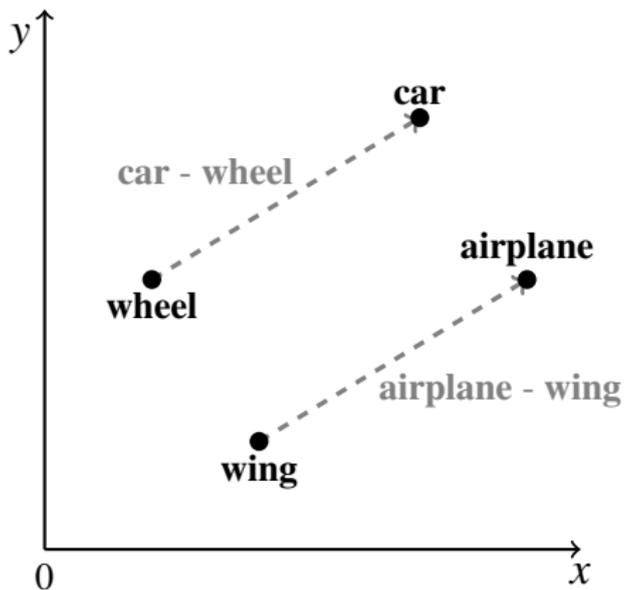
Mechanistic Interpretability of Meronymy

- We analysed the embedding and unembedding layers of LlaMa2-7b, in order to understand how the *part-whole* relation is encoded in the input and output representations of the LLM

Linear Representation Hypothesis (LRH; Mikolov et al. 2013, Park et al. 2024)

The embedding spaces acquired by distributional semantic models are organized in terms of liner subspaces corresponding to high-level concepts

- word pairs expressing the same relation (e.g., ⟨*wheel*, *car*⟩ and ⟨*wing*, *aircraft*⟩) share similar vector differences
- semantic relations are represented in the embedding spaces as directions defined by the word pairs that express them

The Linear Representation Hypothesis

Testing the LRH on Meronymy

1. We selected the embeddings γ of 1,742 meronymic pairs from the union of the MCRAE and CONCEPTNET datasets

2. Given the set Z of meronymic pairs, the direction vector $\hat{\gamma}(partOf)$ is computed as the average of the vector differences of target pairs elements:
$\hat{\gamma}(partOf) = \frac{1}{nZ} \sum_{i=1}^{nZ} [\gamma(m_i) - \gamma(h_i)]$

3. We computed the similarity between the vector difference of a given target pair i and the vector representing the *partOf* relation to see whether they align and point toward similar directions: $\hat{\gamma}(partOf) \cdot (\gamma(m_i) - \gamma(h_i))$

4. We compared above distribution against the dot products between the relation vector $\hat{\gamma}(partOf)$ and a set of randomly selected pairs

Hypothesis

If the *part-of* relation is represented as a linear subspace, meronymic pairs should be significantly more aligned with the relation direction vector than the random pairs

Testing the LRH on Meronymy

1. We selected the embeddings γ of 1,742 meronymic pairs from the union of the MCRAE and CONCEPTNET datasets

2. Given the set Z of meronymic pairs, the direction vector $\hat{\gamma}(partOf)$ is computed as the average of the vector differences of target pairs elements:
$\hat{\gamma}(partOf) = \frac{1}{nZ} \sum_{i=1}^{nZ} [\gamma(m_i) - \gamma(h_i)]$

3. We computed the similarity between the vector difference of a given target pair i and the vector representing the *partOf* relation to see whether they align and point toward similar directions: $\hat{\gamma}(partOf) \cdot (\gamma(m_i) - \gamma(h_i))$

4. We compared above distribution against the dot products between the relation vector $\hat{\gamma}(partOf)$ and a set of randomly selected pairs

Hypothesis

If the *part-of* relation is represented as a linear subspace, meronymic pairs should be significantly more aligned with the relation direction vector than the random pairs

Testing the LRH on Meronymy

1. We selected the embeddings γ of 1,742 meronymic pairs from the union of the MCRAE and CONCEPTNET datasets

2. Given the set Z of meronymic pairs, the direction vector $\hat{\gamma}(partOf)$ is computed as the average of the vector differences of target pairs elements:
$\hat{\gamma}(partOf) = \frac{1}{nZ} \sum_{i=1}^{nZ} [\gamma(m_i) - \gamma(h_i)]$

3. We computed the similarity between the vector difference of a given target pair i and the vector representing the *partOf* relation to see whether they align and point toward similar directions: $\hat{\gamma}(partOf) \cdot (\gamma(m_i) - \gamma(h_i))$

4. We compared above distribution against the dot products between the relation vector $\hat{\gamma}(partOf)$ and a set of randomly selected pairs

Hypothesis

If the *part-of* relation is represented as a linear subspace, meronymic pairs should be significantly more aligned with the relation direction vector than the random pairs

Testing the LRH on Meronymy

1. We selected the embeddings γ of 1,742 meronymic pairs from the union of the MCRAE and CONCEPTNET datasets

2. Given the set Z of meronymic pairs, the direction vector $\hat{\gamma}(partOf)$ is computed as the average of the vector differences of target pairs elements:
$\hat{\gamma}(partOf) = \frac{1}{nZ} \sum_{i=1}^{nZ} [\gamma(m_i) - \gamma(h_i)]$

3. We computed the similarity between the vector difference of a given target pair i and the vector representing the *partOf* relation to see whether they align and point toward similar directions: $\hat{\gamma}(partOf) \cdot (\gamma(m_i) - \gamma(h_i))$

4. We compared above distribution against the dot products between the relation vector $\hat{\gamma}(partOf)$ and a set of randomly selected pairs

Hypothesis

If the *part-of* relation is represented as a linear subspace, meronymic pairs should be significantly more aligned with the relation direction vector than the random pairs
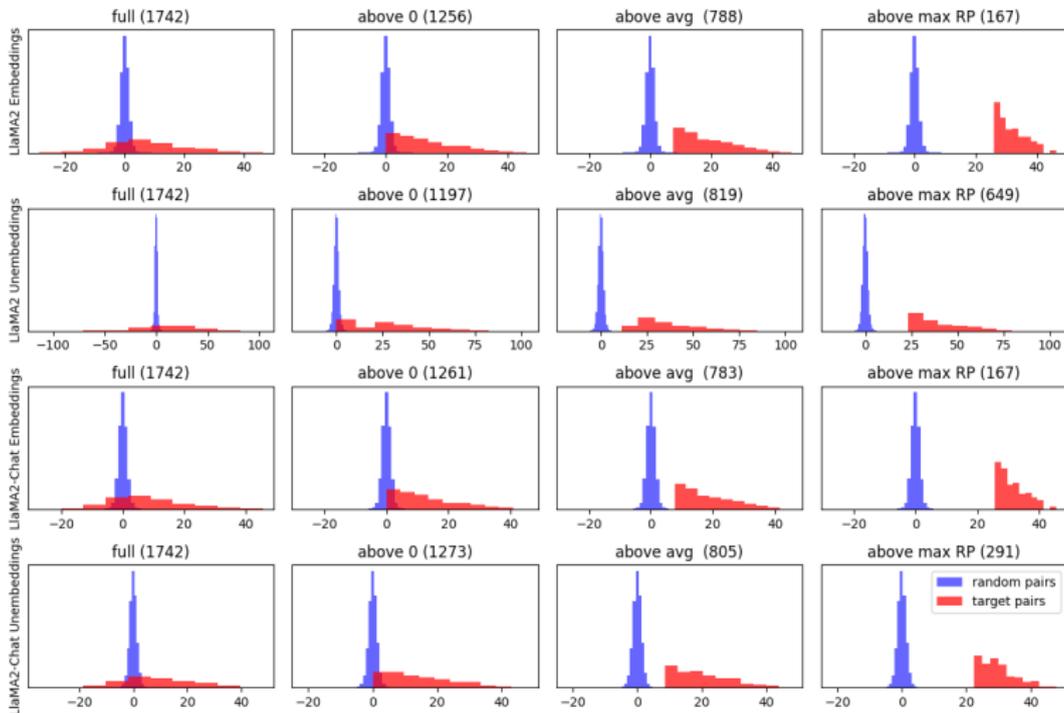
Testing the LRH on Meronymy

1. We selected the embeddings γ of 1,742 meronymic pairs from the union of the MCRAE and CONCEPTNET datasets

2. Given the set Z of meronymic pairs, the direction vector $\hat{\gamma}(partOf)$ is computed as the average of the vector differences of target pairs elements:
$\hat{\gamma}(partOf) = \frac{1}{nZ} \sum_{i=1}^{nZ} [\gamma(m_i) - \gamma(h_i)]$

3. We computed the similarity between the vector difference of a given target pair i and the vector representing the *partOf* relation to see whether they align and point toward similar directions: $\hat{\gamma}(partOf) \cdot (\gamma(m_i) - \gamma(h_i))$

4. We compared above distribution against the dot products between the relation vector $\hat{\gamma}(partOf)$ and a set of randomly selected pairs

Hypothesis

If the *part-of* relation is represented as a linear subspace, meronymic pairs should be significantly more aligned with the relation direction vector than the random pairs

Testing the LRH on Meronymy
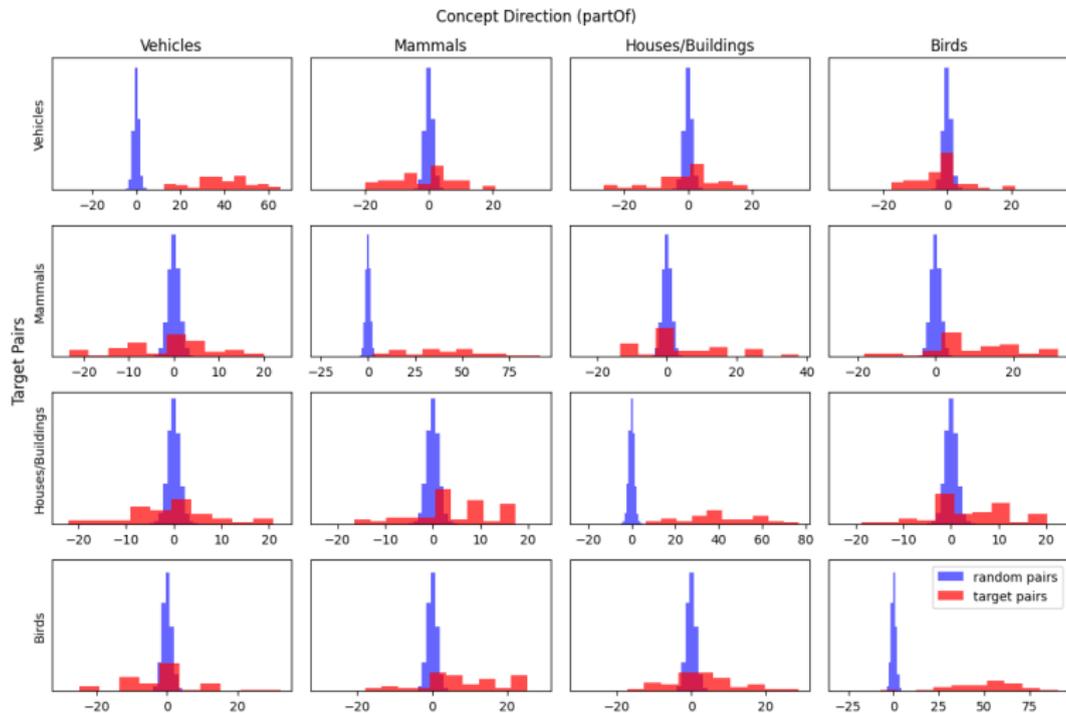
Testing the LRH on Meronymy

a class-based analysis

- We investigated the embedding spaces of meronymic pairs belonging to four semantic classes: birds, mammals, houses/buildings, and vehicles

Class	Seed Holonyms	Target Pairs
Vehicles	7	51
Mammals	11	51
Houses/Buildings	10	39
Birds	17	51

Testing the LRH on Meronymy

a class-based analysis

Testing the LRH on Meronymy

a class-based analysis

Event Knowledge in LLMs

Event Knowledge in Large Language Models: The Gap Between the Impossible and the Unlikely 🔵 🟠

Carina Kauf,[a,b,#] Anna A. Ivanova,[a,b,c,#] Giulia Rambelli,[d]
Emmanuele Chersoni,[e] Jingyuan Selena She,[a,b] Zawad Chowdhury,[f]
Evelina Fedorenko,[a,b] Alessandro Lenci[g]

Impossible and Implausible Events

- Knowledge of the prototypical, abstract structure of everyday events and their participants (e.g., McRae and Matsuki, 2011), also known as Generalized Event Knowledge (GEK)

The cop arrested the thief (plausible event)
The thief arrested the cop (implausible event)
The cop arrested the number (semantically impossible)

Impossible and Implausible Events

- Knowledge of the prototypical, abstract structure of everyday events and their participants (e.g., McRae and Matsuki, 2011), also known as Generalized Event Knowledge (GEK)

The cop arrested the thief (plausible event)

The thief arrested the cop (implausible event)

The cop arrested the number (semantically impossible)

Impossible and Implausible Events

- Knowledge of the prototypical, abstract structure of everyday events and their participants (e.g., McRae and Matsuki, 2011), also known as Generalized Event Knowledge (GEK)

The cop arrested the thief (plausible event)

The thief arrested the cop (implausible event)

The cop arrested the number (semantically impossible)

Impossible and Implausible Events

- Knowledge of the prototypical, abstract structure of everyday events and their participants (e.g., McRae and Matsuki, 2011), also known as Generalized Event Knowledge (GEK)

The cop arrested the thief (plausible event)

The thief arrested the cop (implausible event)

The cop arrested the number (semantically impossible)

Kauf et al. (2023), "Event knowledge in large language models: The gap between the impossible and the unlikely", *Cognitive Science*: 47

- We evaluate the ability of humans and LLMs to assign a higher log-probability to the plausible event description than the implausible (impossible) one
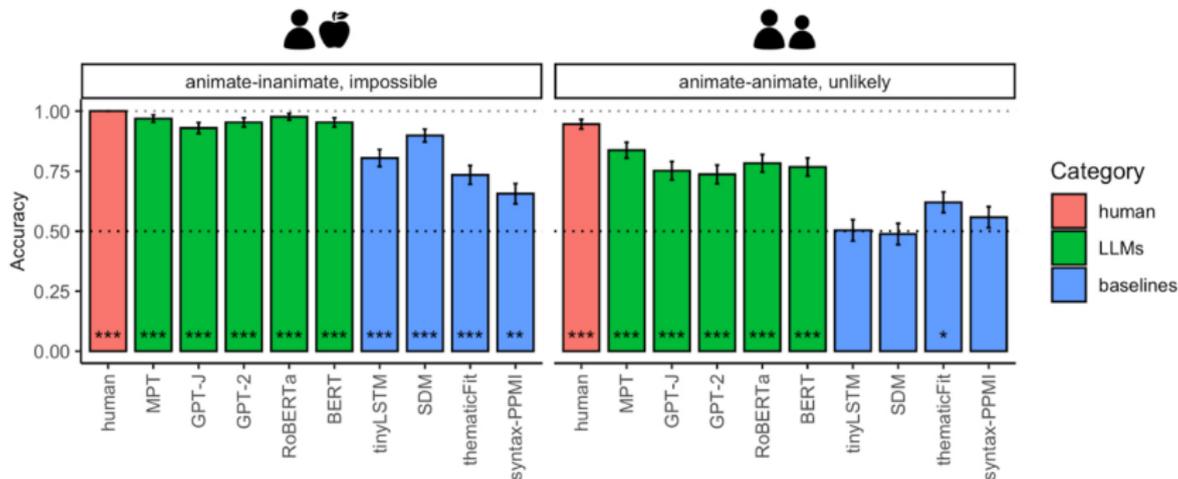
Embedding Events in Propositional Attitudes

Embedding (im)plausible clauses in propositional attitude contexts: Modulatory effects on the N400 and late components

Lia Călinescu[1], Anna Giskes[1], Mila Vulchanova, Giosuè Baggio*

Language Acquisition and Language Processing Lab, Department of Language and Literature, Norwegian University of Science and Technology, Trondheim, Norway

Plausible vs. Implausible Events in Epistemic Contexts

- Propositional attitude verbs can change the plausibility of embedded events

  (1) a. Cars have wheels. (plausible)
       b. Cars have wings. (implausible)

  (2) a. Magnus knows that cars have wheels. (plausible)
       b. Magnus knows that cars have wings. (implausible)

  (3) a. Magnus believes that cars have wheels. (plausible)
       b. Magnus believes that cars have wings. (plausible)

Plausible vs. Implausible Events in Epistemic Contexts

- Propositional attitude verbs can change the plausibility of embedded events

  (1)   a.   Cars have wheels. (plausible)
         b.   Cars have wings. (implausible)

  (2)   a.   Magnus knows that cars have wheels. (plausible)
         b.   Magnus knows that cars have wings. (implausible)

  (3)   a.   Magnus believes that cars have wheels. (plausible)
         b.   Magnus believes that cars have wings. (plausible)

Plausible vs. Implausible Events in Propositional Attitudes

- Factuality Scale: know > believe > dream > doubt > imagine

- Declerk (2011), "The definition of modality"
  - factive verbs (e.g., know) evoke a world which is "automatically interpreted as being the factual world" (p. 41)
  - attitude verbs (e.g., believe, doubt, dream, imagine, etc.) create an "intensional world which may or may not coincide with the factual world"

# Plausible vs. Implausible Events in Propositional Attitudes

- Factuality Scale: know > believe > dream > doubt > imagine

- Declerk (2011), "The definition of modality"
  - factive verbs (e.g., know) evoke a world which is "automatically interpreted as being the factual world" (p. 41)
  - attitude verbs (e.g., believe, doubt, dream, imagine, etc.) create an "intensional world which may or may not coincide with the factual world"
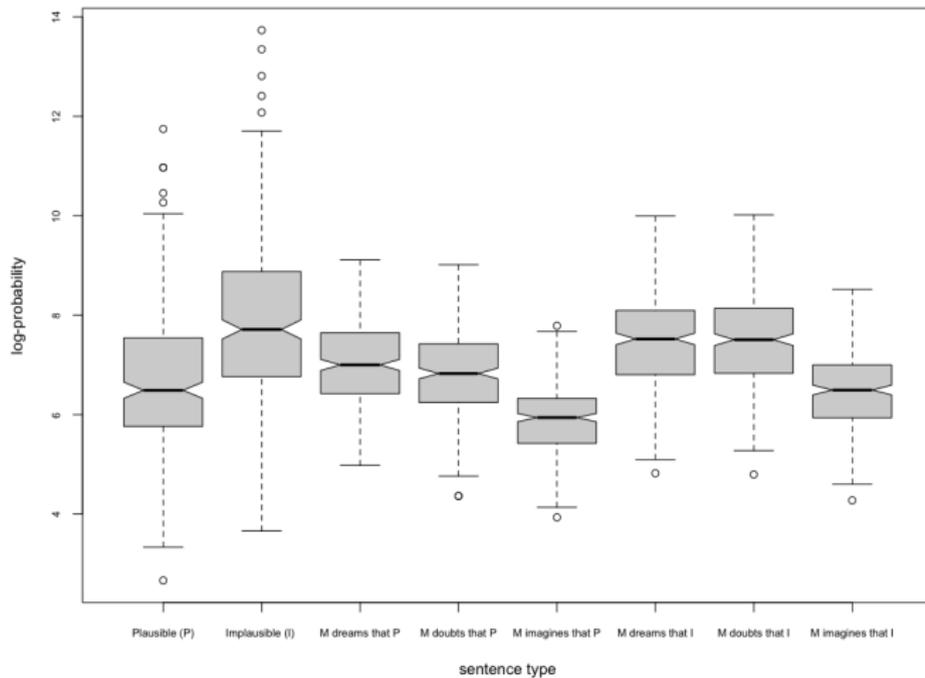
Experimental Setting

- Stimuli from Călinescu et al. (2020)
  - 300 plausible sentences (P) + 300 implausible sentences (I)
  - the P and I sentences were embedded in propositional attitude contexts with 5 different verbs differing for factuality (know, believe, dream, doubt, imagine), for a total of 3,600 data points

- Model: Llama-3 8B Instruct

- Measure: the LLM computed the log-probability scores (Kauf et al. 2023, 2024) of the P and I sentences both as main clauses and as embedded ones

Results
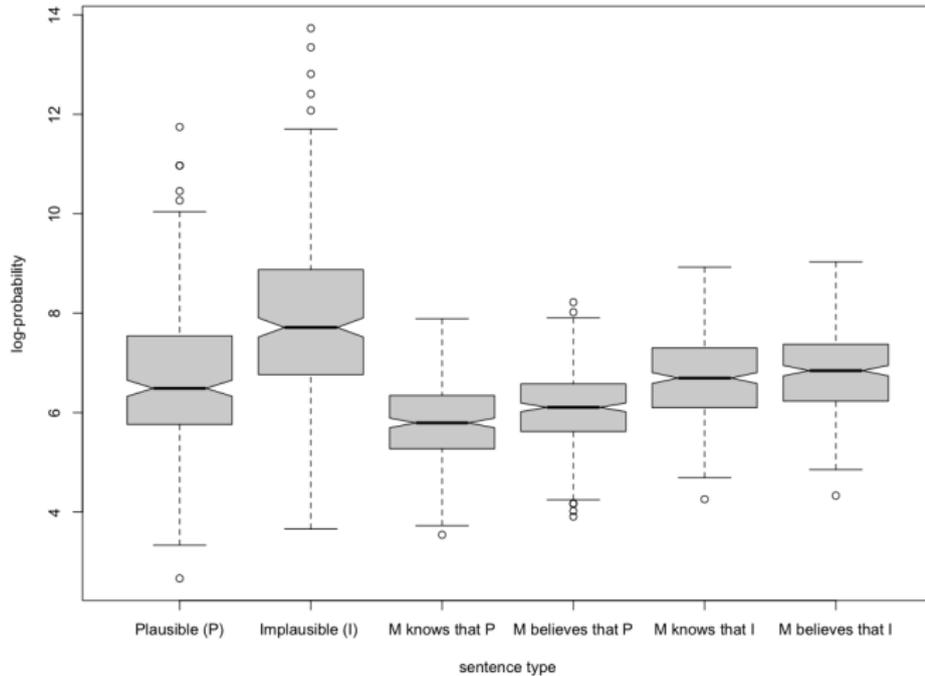
Results

The "Quasi-semantic" Space of LLMs

- LLMs do not seem to acquire fully general and abstract semantic notions
- LLMs – like all Distributional Semantic Models – identify highly sophisticated distributional associations between linguistic expressions, but they have just a "quasi-semantic" space
  - cf. also Gu et al. (2023), Misra et al. (2023), Berglund et al. (2024)

Human Conceptual Space

It is organized in terms of structured "thories": A structured network of entities and events linked by relations that specify their functional role in a system

The "Quasi-semantic" Space of LLMs

- LLMs do not seem to acquire fully general and abstract semantic notions
- LLMs – like all Distributional Semantic Models – identify highly sophisticated distributional associations between linguistic expressions, but they have just a "quasi-semantic" space
    - cf. also Gu et al. (2023), Misra et al. (2023), Berglund et al. (2024)
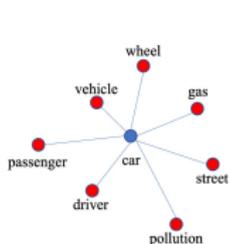
Human Conceptual Space

It is organized in terms of structured "thories": A structured network of entities and events linked by relations that specify their functional role in a system
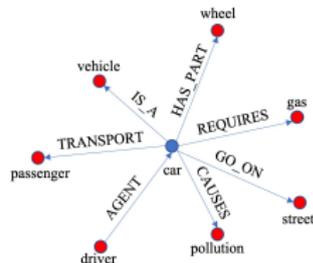
The "Quasi-semantic" Space of LLMs

- LLMs do not seem to acquire fully general and abstract semantic notions
- LLMs – like all Distributional Semantic Models – identify highly sophisticated distributional associations between linguistic expressions, but they have just a "quasi-semantic" space
  - cf. also Gu et al. (2023), Misra et al. (2023), Berglund et al. (2024)

Human Conceptual Space

It is organized in terms of structured "thories": A structured network of entities and events linked by relations that specify their functional role in a system

distributional semantic space of *car*   theory (knowledge) about *car*

Why Do LLMs Learn so Much?

- The Radical Distributional Learning of LLMs is so successful because of the strong correlations between word co-occurrence statistics and semantic relationships (Titus 2024)

The Redundancy Hypothesis (Louwerse 2011)

Language has evolved to become a communicative short-cut for language users and encodes relations in the world, including embodied relations.

Why Do LLMs Learn so Much?

- The Radical Distributional Learning of LLMs is so successful because of the strong correlations between word co-occurrence statistics and semantic relationships (Titus 2024)

The Redundancy Hypothesis (Louwerse 2011)

Language has evolved to become a communicative short-cut for language users and encodes relations in the world, including embodied relations.
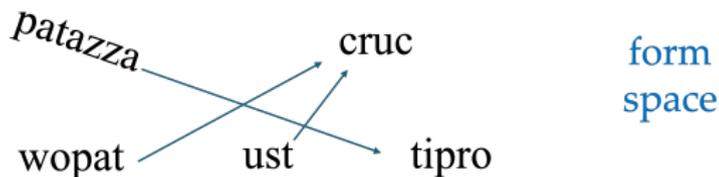
Why Radical Distribution Learning Can't be Enough?

- Despite language-world correlations, the Semantic Gap of LLMs derives from the lack of a perfect isomorphism between distributional statistics and conceptual/logical/referential properties
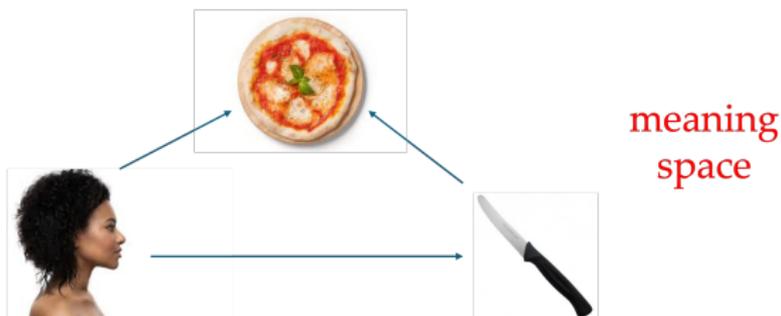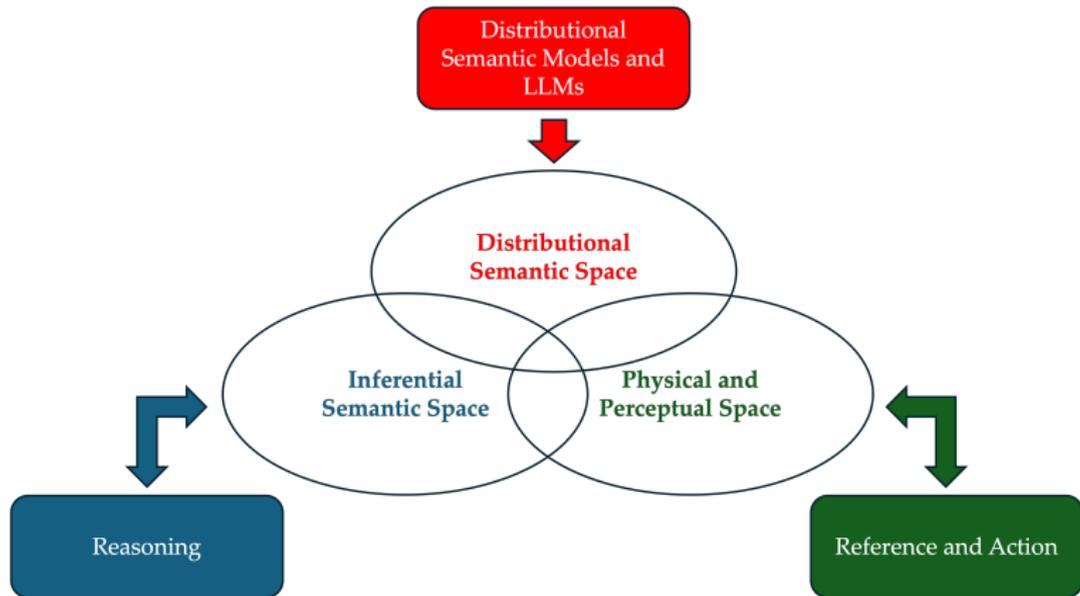
Why Radical Distribution Learning Can't be Enough?

- Despite language-world correlations, the Semantic Gap of LLMs derives from the lack of a perfect isomorphism between distributional statistics and conceptual/logical/referential properties

meaning space

form space

patazza   cruc

wopat   ust   tipro

Why Radical Distribution Learning Can't be Enough?

Conclusions

- The "magic" of LLMs is simply the "magic" of (radical) distributional learning
- The real scientific revelation brought by LLMs is that the range meaning aspects that can be recovered from distributional statistics is far greater than we imagined before (at least if we have enough amount of data)
- The major reason of the Semantic Gap in LLMs might lie in the very type of information they extract from simple co-occurrences between linguistic forms, which is not true semantic knowledge yet

"All this is not to say that there is not a great interconnection between language and meaning, in whatever sense it may be possible to use this word. But it is not a one-to-one relation between morphological structure and anything else. [...] If one wishes to speak of language as exiting in some sense on two planes – of form and of meaning – we can at least say that the structures of the two are not identical, though they will be found similar in various respects." (Harris 1954: 151-152)

Conclusions

- The "magic" of LLMs is simply the "magic" of (radical) distributional learning
- The real scientific revelation brought by LLMs is that the range meaning aspects that can be recovered from distributional statistics is far greater than we imagined before (at least if we have enough amount of data)
- The major reason of the Semantic Gap in LLMs might lie in the very type of information they extract from simple co-occurrences between linguistic forms, which is not true semantic knowledge yet

"All this is not to say that there is not a great interconnection between language and meaning, in whatever sense it may be possible to use this word. But it is not a one-to-one relation between morphological structure and anything else. [...] If one wishes to speak of language as exiting in some sense on two planes – of form and of meaning – we can at least say that the structures of the two are not identical, though they will be found similar in various respects." (Harris 1954: 151-152)

Conclusions

- The "magic" of LLMs is simply the "magic" of (radical) distributional learning
- The real scientific revelation brought by LLMs is that the range meaning aspects that can be recovered from distributional statistics is far greater than we imagined before (at least if we have enough amount of data)
- The major reason of the Semantic Gap in LLMs might lie in the very type of information they extract from simple co-occurrences between linguistic forms, which is not true semantic knowledge yet

"All this is not to say that there is not a great interconnection between language and meaning, in whatever sense it may be possible to use this word. But it is not a one-to-one relation between morphological structure and anything else. [...] If one wishes to speak of language as exiting in some sense on two planes – of form and of meaning – we can at least say that the structures of the two are not identical, though they will be found similar in various respects." (Harris 1954: 151-152)

Conclusions

- The "magic" of LLMs is simply the "magic" of (radical) distributional learning
- The real scientific revelation brought by LLMs is that the range meaning aspects that can be recovered from distributional statistics is far greater than we imagined before (at least if we have enough amount of data)
- The major reason of the Semantic Gap in LLMs might lie in the very type of information they extract from simple co-occurrences between linguistic forms, which is not true semantic knowledge yet

"All this is not to say that there is not a great interconnection between language and meaning, in whatever sense it may be possible to use this word. But it is not a one-to-one relation between morphological structure and anything else. [...] If one wishes to speak of language as exiting in some sense on two planes – of form and of meaning – we can at least say that the structures of the two are not identical, though they will be found similar in various respects." (Harris 1954: 151-152)

Grazie!!!

Thank you!!!