

Perché l'Intelligenza Artificiale crede di fare a meno della teoria linguistica, ma in realtà non potrà farlo

Alessandro Lenci



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

Intelligenza Artificiale (IA) e linguaggio

- Uno degli obiettivi dell'IA è dotare i sistemi artificiali dell'abilità di comprendere e generare il linguaggio come gli esseri umani

Culicover and Jackendoff (2005), *Simpler Syntax*, p. 10

“A speaker of a human language can create and understand an **unlimited number** of different utterances, concerning an unlimited number of different topics.

This entails that a language user with a finite brain must have a **productive system** for constructing new utterances online (in both production and perception) from a **finite basis** stored in memory”

- (1) Ieri sera Gianni ha chiesto a Paolo se volesse andare al cinema con lui. Paolo gli ha risposto: “Domani devo alzarmi presto”. Cosa intendeva dire Paolo a Gianni?

Intelligenza Artificiale (IA) e linguaggio

- Uno degli obiettivi dell'IA è dotare i sistemi artificiali dell'abilità di comprendere e generare il linguaggio come gli esseri umani

Culicover and Jackendoff (2005), *Simpler Syntax*, p. 10

“A speaker of a human language can create and understand an **unlimited number** of different utterances, concerning an unlimited number of different topics.

This entails that a language user with a finite brain must have a **productive system** for constructing new utterances online (in both production and perception) from a **finite basis** stored in memory”

- (1) Ieri sera Gianni ha chiesto a Paolo se volesse andare al cinema con lui. Paolo gli ha risposto: “Domani devo alzarmi presto”. Cosa intendeva dire Paolo a Gianni?

Intelligenza Artificiale (IA) e linguaggio

- Uno degli obiettivi dell'IA è dotare i sistemi artificiali dell'abilità di comprendere e generare il linguaggio come gli esseri umani

Culicover and Jackendoff (2005), *Simpler Syntax*, p. 10

“A speaker of a human language can create and understand an **unlimited number** of different utterances, concerning an unlimited number of different topics.

This entails that a language user with a finite brain must have a **productive system** for constructing new utterances online (in both production and perception) from a **finite basis** stored in memory”

- (1) Ieri sera Gianni ha chiesto a Paolo se volesse andare al cinema con lui. Paolo gli ha risposto: “Domani devo alzarmi presto”. Cosa intendeva dire Paolo a Gianni?

“Il linguaggio è la casa dell’Essere. Nella sua dimora abita l’uomo”

Martin Heidegger, *Lettera sull’“umanismo”*, 1947

L’argomentazione fondante l’IA ... *almeno fino ai Large Language Models (LLM)*

- Il linguaggio è la quintessenza dell’essere umano
- È possibile costruire macchine capaci di generare e comprendere il linguaggio solo implementando le teorie e i modelli elaborati dalle scienze del linguaggio umano (= linguistica)

“Il linguaggio è la casa dell’Essere. Nella sua dimora abita l’uomo”

Martin Heidegger, *Lettera sull’“umanismo”*, 1947

L’argomentazione fondante l’IA ... *almeno fino ai Large Language Models (LLM)*

- Il linguaggio è la quintessenza dell’essere umano
- È possibile costruire macchine capaci di generare e comprendere il linguaggio solo implementando le teorie e i modelli elaborati dalle scienze del linguaggio umano (= linguistica)

“Il linguaggio è la casa dell’Essere. Nella sua dimora abita l’uomo”

Martin Heidegger, *Lettera sull’“umanismo”*, 1947

L’argomentazione fondante l’IA ... *almeno fino ai Large Language Models (LLM)*

- Il linguaggio è la quintessenza dell’essere umano
- È possibile costruire macchine capaci di generare e comprendere il linguaggio solo **implementando le teorie e i modelli elaborati dalle scienze del linguaggio umano** (= linguistica)

IA e scienze del linguaggio, *ante LLM*

- **Scienze del linguaggio**: individuano e formalizzano le conoscenze e i processi che determinano le nostre abilità linguistiche
- **IA**: i modelli linguistici sono forniti ai sistemi artificiali con l'obiettivo di dotarli delle medesime capacità
- Le prime due generazioni di IA:
 - **Modelli a regole** (Good Old-Fashioned AI, GOFAI)
 - **Modelli statistici** (data driven) basati su **Machine Learning (ML)** supervisionato

Assunzione comune

- I sistemi di IA elaborano il linguaggio utilizzando **rappresentazioni esplicite di categorie e strutture** definite dalla linguistica
- La linguistica definisce il **metalinguaggio** con cui vengono rappresentati i processi delle macchine

IA e scienze del linguaggio, *ante LLM*

- **Scienze del linguaggio**: individuano e formalizzano le conoscenze e i processi che determinano le nostre abilità linguistiche
- **IA**: i modelli linguistici sono forniti ai sistemi artificiali con l'obiettivo di dotarli delle medesime capacità
- Le prime due generazioni di IA:
 - **Modelli a regole** (Good Old-Fashioned AI, GOFAI)
 - **Modelli statistici** (data driven) basati su **Machine Learning (ML)** supervisionato

Assunzione comune

- I sistemi di IA elaborano il linguaggio utilizzando **rappresentazioni esplicite di categorie e strutture** definite dalla linguistica
- La linguistica definisce il **metalinguaggio** con cui vengono rappresentati i processi delle macchine

IA e scienze del linguaggio, *ante LLM*

- **Scienze del linguaggio**: individuano e formalizzano le conoscenze e i processi che determinano le nostre abilità linguistiche
- **IA**: i modelli linguistici sono forniti ai sistemi artificiali con l'obiettivo di dotarli delle medesime capacità
- Le prime due generazioni di IA:
 - **Modelli a regole** (Good Old-Fashioned AI, GOFAI)
 - **Modelli statistici** (data driven) basati su **Machine Learning (ML)** supervisionato

Assunzione comune

- I sistemi di IA elaborano il linguaggio utilizzando **rappresentazioni esplicite di categorie e strutture** definite dalla linguistica
- La linguistica definisce il **metalinguaggio** con cui vengono rappresentati i processi delle macchine

IA e scienze del linguaggio, *ante LLM*

- **Scienze del linguaggio**: individuano e formalizzano le conoscenze e i processi che determinano le nostre abilità linguistiche
- **IA**: i modelli linguistici sono forniti ai sistemi artificiali con l'obiettivo di dotarli delle medesime capacità
- Le prime due generazioni di IA:
 - **Modelli a regole** (Good Old-Fashioned AI, GOFAI)
 - **Modelli statistici** (data driven) basati su **Machine Learning (ML)** supervisionato

Assunzione comune

- I sistemi di IA elaborano il linguaggio utilizzando **rappresentazioni esplicite di categorie e strutture** definite dalla linguistica
- La linguistica definisce il **metalinguaggio** con cui vengono rappresentati i processi delle macchine

IA e scienze del linguaggio, *ante LLM*

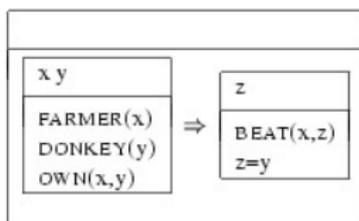
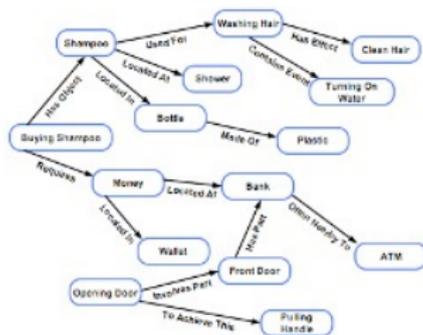
- **Scienze del linguaggio**: individuano e formalizzano le conoscenze e i processi che determinano le nostre abilità linguistiche
- **IA**: i modelli linguistici sono forniti ai sistemi artificiali con l'obiettivo di dotarli delle medesime capacità
- Le prime due generazioni di IA:
 - **Modelli a regole** (Good Old-Fashioned AI, GOFAI)
 - **Modelli statistici** (data driven) basati su **Machine Learning (ML)** supervisionato

Assunzione comune

- I sistemi di IA elaborano il linguaggio utilizzando **rappresentazioni esplicite di categorie e strutture** definite dalla linguistica
- La linguistica definisce il **metalinguaggio** con cui vengono rappresentati i processi delle macchine

Rappresentazioni linguistiche

il caso della semantica



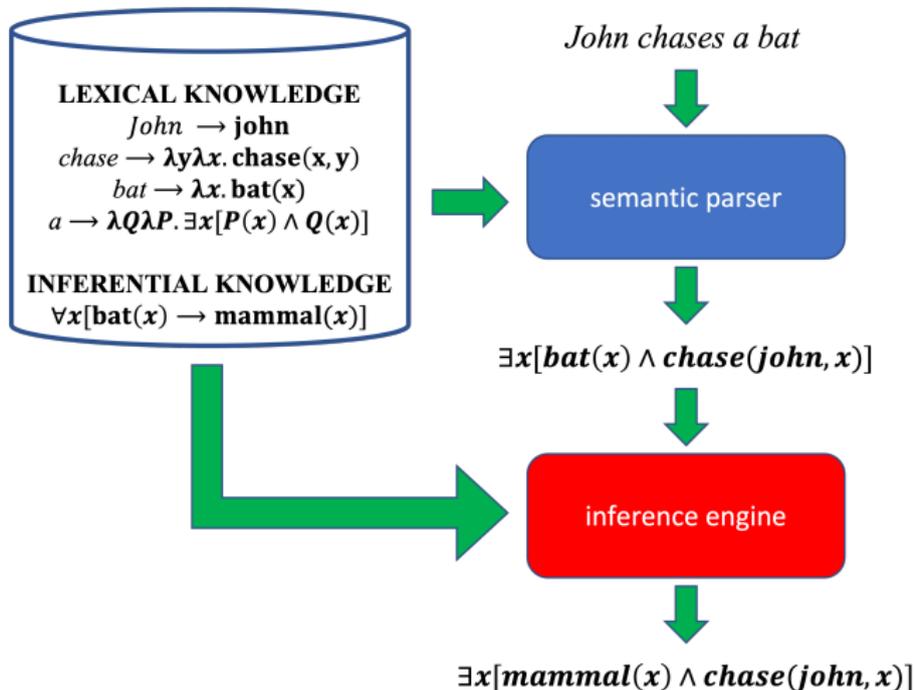
$$\left[\begin{array}{l}
 \text{kill} \\
 \text{EVENTSTR} = \begin{bmatrix} E_1 = e_1:\text{process} \\ E_2 = e_2:\text{state} \\ \text{RESTR} = <\alpha \\ \text{HEAD} = e_1 \end{bmatrix} \\
 \text{ARGSTR} = \begin{bmatrix} \text{ARG1} = \boxed{1} \begin{bmatrix} \text{ind} \\ \text{FORMAL} = \text{physobj} \end{bmatrix} \\ \text{ARG2} = \boxed{2} \begin{bmatrix} \text{animate_ind} \\ \text{FORMAL} = \text{physobj} \end{bmatrix} \end{bmatrix} \\
 \text{QUALIA} = \begin{bmatrix} \text{cause-lcp} \\ \text{FORMAL} = \text{dead}(e_2, \boxed{2}) \\ \text{AGENTIVE} = \text{kill_act}(e_1, \boxed{1}, \boxed{2}) \end{bmatrix}
 \end{array} \right]$$

ROLES

- AGENT [+ANIMATE | +ORGANIZATION]
- THEME
- SOURCE [+CONCRETE]
- BENEFICIARY [+ANIMATE | +ORGANIZATION]
- ASSET [-LOCATION & -REGION]

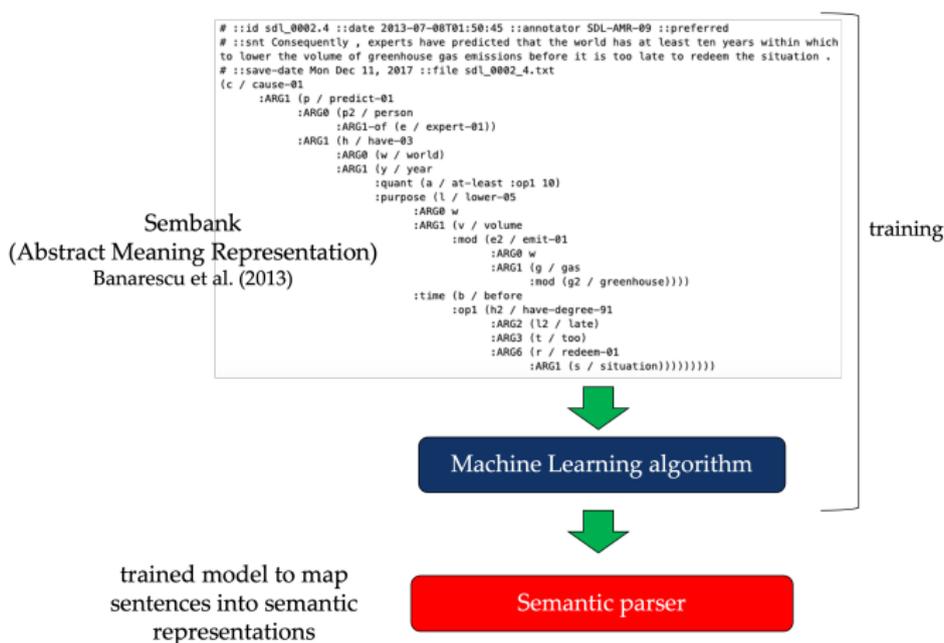
[EVENT GO ([THING Ray], [PATH TO ([PLACE IN ([THING the room])]])])]

IA Ante LLM - GOFAI a regole



IA Ante LLM - ML supervisionato

- Le categorie e strutture linguistiche sono fornite ai sistemi come **dati annotati** su cui sono addestrati gli algoritmi di ML supervisionato



I vantaggi dell'IA *Ante LLM*

Interpretabilità – sappiamo quale informazione è codificata dai modelli, e questa può essere facilmente interpretata dagli umani

[S [NP Il cane]₁ [VP entrò₂ [PP nella₃ [NP stanza]₄]]].

[_{EVENT} GO ([_{OBJECT} DOG]₁, [_{PATH TO} ([_{PLACE} IN ([_{OBJECT} ROOM]₄))]₃)]₂]

Spiegabilità – il comportamento dei modelli può essere spiegato direttamente in termine delle strutture esplicite che essi manipolano

- *John chases a bat* \Rightarrow *John chases an animal*
- *kill* $\rightarrow \lambda x \lambda y. [\text{kill}(x, y)] \Leftrightarrow \lambda x \lambda y. [\text{CAUSE}(x, \text{BECOME}(\text{DEAD}(y)))]$

I vantaggi dell'IA *Ante LLM*

Interpretabilità – sappiamo quale informazione è codificata dai modelli, e questa può essere facilmente interpretata dagli umani

[S [NP Il cane]₁ [VP entrò₂ [PP nella₃ [NP stanza]₄]]].

[EVENT GO ([OBJECT DOG]₁, [PATH TO ([PLACE IN ([OBJECT ROOM]₄))]₃)]₂

Spiegabilità – il comportamento dei modelli può essere **spiegato** direttamente in termine delle strutture esplicite che essi manipolano

- *John chases a bat* \Rightarrow *John chases an animal*
- *kill* $\rightarrow \lambda x \lambda y. [\mathbf{kill}(x, y)] \Leftrightarrow \lambda x \lambda y. [\mathbf{CAUSE}(x, \mathbf{BECOME}(\mathbf{DEAD}(y)))]$

I limiti dell'IA *Ante LLM*

Learnability (per GOFAI a regole) – manca un modello computazionale di come le strutture linguistiche vengono apprese

- cf. **knowledge acquisition bottleneck** in IA

Dati annotati – I modelli di ML supervisionato richiedono molti dati annotati

- cf. **data annotation bottleneck** in IA

Generalità – manca un modello computazionale generale delle diverse capacità linguistiche

- cf. paradigma “one task, one model”

I limiti dell'IA *Ante LLM*

Learnability (per GOFAI a regole) – manca un modello computazionale di come le strutture linguistiche vengono apprese

- cf. **knowledge acquisition bottleneck** in IA

Dati annotati – I modelli di ML supervisionato richiedono molti dati annotati

- cf. **data annotation bottleneck** in IA

Generalità – manca un modello computazionale generale delle diverse capacità linguistiche

- cf. paradigma “one task, one model”

I limiti dell'IA *Ante LLM*

Learnability (per GOFAI a regole) – manca un modello computazionale di come le strutture linguistiche vengono apprese

- cf. **knowledge acquisition bottleneck** in IA

Dati annotati – I modelli di ML supervisionato richiedono molti dati annotati

- cf. **data annotation bottleneck** in IA

Generalità – manca un modello computazionale generale delle diverse capacità linguistiche

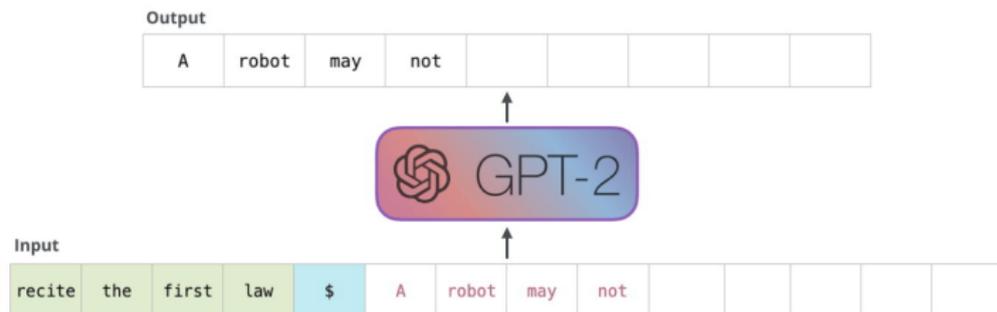
- cf. paradigma **“one task, one model”**

LLM (Foundation Models, Pretrained Models)

tratti comuni

● Pretraining

- reti neurali (Transformer) che acquisiscono grandi quantità di conoscenze sulle **strutture linguistiche** e sul **mondo** durante l'addestramento su **dati testuali non annotati** con un **self-supervised string prediction task**



- la conoscenza è codificata nei **vettori** che corrispondono ai livelli interni della rete

● Task adaptation

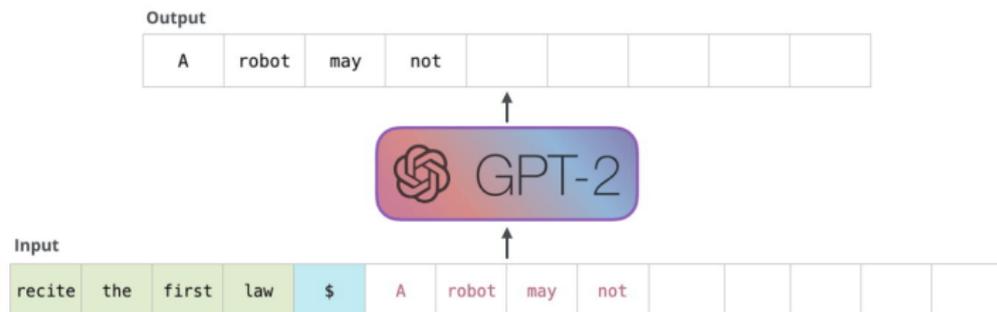
- utilizza la conoscenza acquisita durante il pre-training per risolvere compiti linguistici con **fine-tuning** o **in-context learning (prompting)**

LLM (Foundation Models, Pretrained Models)

tratti comuni

● Pretraining

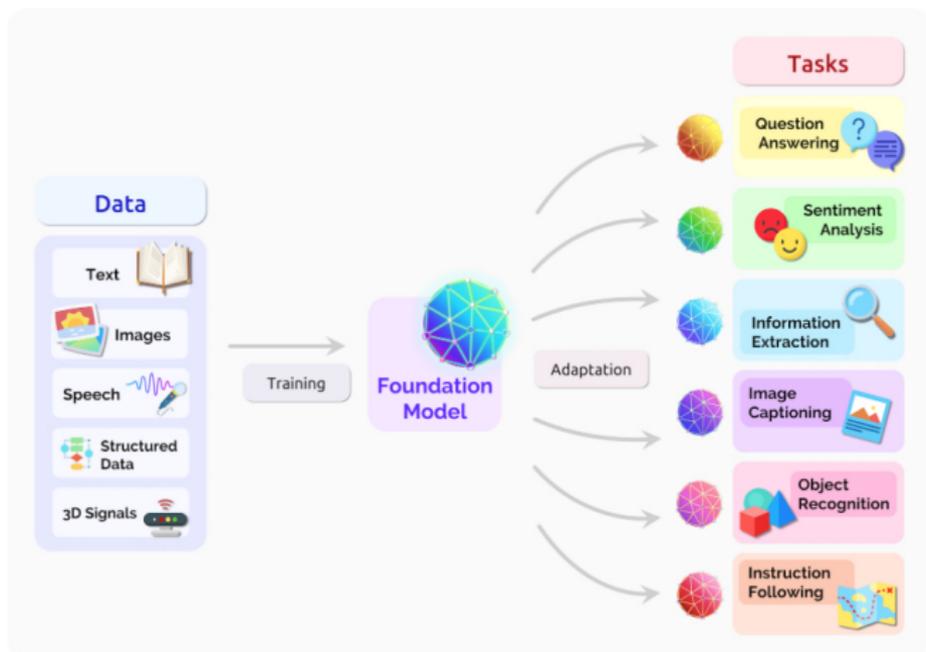
- reti neurali (Transformer) che acquisiscono grandi quantità di conoscenze sulle **strutture linguistiche** e sul **mondo** durante l'addestramento su **dati testuali non annotati** con un **self-supervised string prediction task**



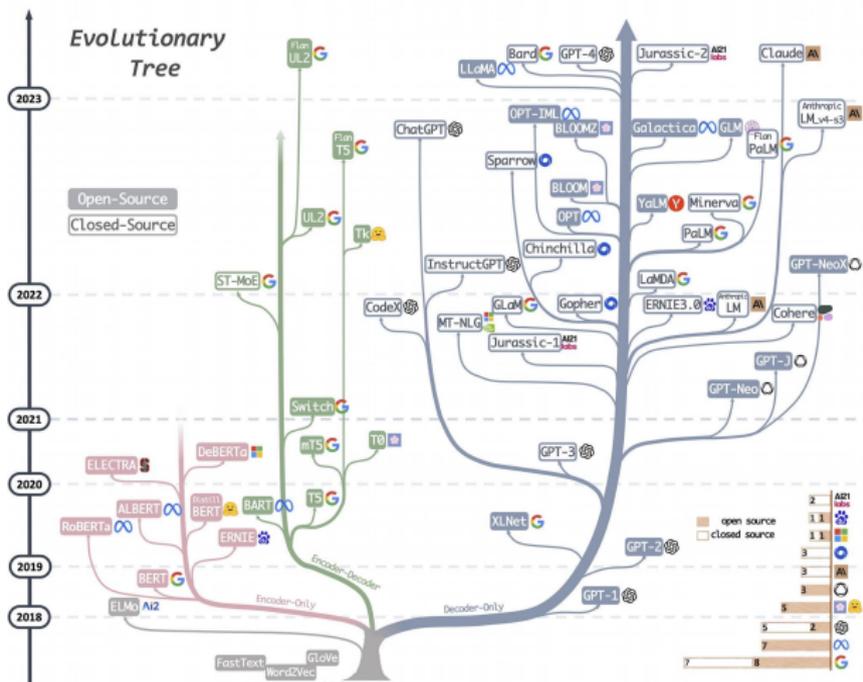
- la conoscenza è codificata nei **vettori** che corrispondono ai livelli interni della rete
- **Task adaptation**
- utilizza la conoscenza acquisita durante il pre-training per risolvere compiti linguistici con **fine-tuning** o **in-context learning (prompting)**

LLM (Foundation Models, Pretrained Models)

Bommasani et al. (2021). "On the opportunities and risks of foundation models". *ArXiv*: 2108.07258



L'albero degli LLM e le sue radici



LLM come Distributional Semantic Models (DSMs)

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

Distributional Semantics

Il contenuto delle espressioni linguistiche è rappresentato con vettori (**embeddings**) creati a partire da **dati distribuzionali** estratti da corpora

3.2 0.3 4.5 0.1 0.2 7.8 0.3 2.9 1.5 0.9 4.7 4.8

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 ...

Scienze del linguaggio e IA *post LLM*

- Gli LLM manifestano capacità linguistiche superiori alle generazioni precedenti, **senza essere stati specificatamente addestrati a riconoscere e a manipolare categorie e strutture linguistiche**
- Le conoscenze non sono fornite ai LLM né sotto forma di regole, né di dati annotati, ma sono **‘abilità emergenti’** (Wei et al. 2022), poiché nascono come risultato del compito di predizione con cui le reti vengono addestrate
 - ma anche con il contributo **determinante** di raffinamenti a posteriori come instruction tuning, reinforcement learning, ecc.

Scienze del linguaggio e IA *post LLM*

- Gli LLM manifestano capacità linguistiche superiori alle generazioni precedenti, **senza essere stati specificatamente addestrati a riconoscere e a manipolare categorie e strutture linguistiche**
- Le conoscenze non sono fornite ai LLM né sotto forma di regole, né di dati annotati, ma sono **‘abilità emergenti’** (Wei et al. 2022), poiché nascono come risultato del compito di predizione con cui le reti vengono addestrate
 - ma anche con il contributo **determinante** di raffinamenti a posteriori come instruction tuning, reinforcement learning, ecc.

Scienze del linguaggio e IA *post LLM*

- (2) a. *prompt*: Ieri sera Gianni ha chiesto a Paolo se volesse andare al cinema con lui. Paolo gli ha risposto: “Domani devo alzarmi presto”. Cosa intendeva dire Paolo a Gianni?
- b. *ChatGPT*: Paolo ha risposto a Gianni dicendo “Domani devo alzarmi presto”, il che implica che Paolo non può andare al cinema con Gianni la sera stessa. Questo perché deve svegliarsi presto il giorno seguente e probabilmente vuole riposare e andare a letto presto per essere riposato
- (3) a. *prompt*: Qual è il soggetto di “volesse andare” nella frase “Ieri sera Gianni ha chiesto a Paolo se volesse andare al cinema con lui.”?
- b. *ChatGPT*: Il soggetto di “volesse andare” nella frase “Ieri sera Gianni ha chiesto a Paolo se volesse andare al cinema con lui.” è “Paolo”. La frase può essere parafrasata come “Ieri sera Gianni ha chiesto a Paolo se Paolo volesse andare al cinema con lui.”

IA linguistica senza scienze del linguaggio?

Gli LLM sono le prime “macchine parlanti” costruite **senza un contributo diretto delle teorie della linguistica sull'organizzazione e il funzionamento del linguaggio umano**

Scienze del linguaggio e IA *post LLM*

- (2) a. *prompt*: Ieri sera Gianni ha chiesto a Paolo se volesse andare al cinema con lui. Paolo gli ha risposto: “Domani devo alzarmi presto”. Cosa intendeva dire Paolo a Gianni?
- b. *ChatGPT*: Paolo ha risposto a Gianni dicendo “Domani devo alzarmi presto”, il che implica che Paolo non può andare al cinema con Gianni la sera stessa. Questo perché deve svegliarsi presto il giorno seguente e probabilmente vuole riposare e andare a letto presto per essere riposato
- (3) a. *prompt*: Qual è il soggetto di “volesse andare” nella frase “Ieri sera Gianni ha chiesto a Paolo se volesse andare al cinema con lui.”?
- b. *ChatGPT*: Il soggetto di “volesse andare” nella frase “Ieri sera Gianni ha chiesto a Paolo se volesse andare al cinema con lui.” è “Paolo”. La frase può essere parafrasata come “Ieri sera Gianni ha chiesto a Paolo se Paolo volesse andare al cinema con lui.”

IA linguistica senza scienze del linguaggio?

Gli LLM sono le prime “macchine parlanti” costruite **senza un contributo diretto delle teorie della linguistica sull'organizzazione e il funzionamento del linguaggio umano**

Scienze del linguaggio e IA *post LLM*

- (2) a. *prompt*: Ieri sera Gianni ha chiesto a Paolo se volesse andare al cinema con lui. Paolo gli ha risposto: “Domani devo alzarmi presto”. Cosa intendeva dire Paolo a Gianni?
- b. *ChatGPT*: Paolo ha risposto a Gianni dicendo “Domani devo alzarmi presto”, il che implica che Paolo non può andare al cinema con Gianni la sera stessa. Questo perché deve svegliarsi presto il giorno seguente e probabilmente vuole riposare e andare a letto presto per essere riposato
- (3) a. *prompt*: Qual è il soggetto di “volesse andare” nella frase “Ieri sera Gianni ha chiesto a Paolo se volesse andare al cinema con lui.”?
- b. *ChatGPT*: Il soggetto di “volesse andare” nella frase “Ieri sera Gianni ha chiesto a Paolo se volesse andare al cinema con lui.” è “Paolo”. La frase può essere parafrasata come “Ieri sera Gianni ha chiesto a Paolo se Paolo volesse andare al cinema con lui.”

IA linguistica senza scienze del linguaggio?

Gli LLM sono le prime “macchine parlanti” costruite **senza un contributo diretto delle teorie della linguistica sull'organizzazione e il funzionamento del linguaggio umano**

“Us and Them” (Pink Floyd, *The dark Side of the Moon*)

- In cosa gli LLM sono **simili a noi**:
 - centralità dei meccanismi predittivi nell’elaborazione del linguaggio
 - uso di rappresentazioni vettoriali continue
 - importanza dell’informazione distribuzionale (cf. linguistica strutturalista)

- In cosa gli LLM sono **diversi da noi**:
 - sono addestrati su quantità “**sovraumane**” di dati linguistici
 - i dati per addestrare ChatGPT corrispondono a **36.540 anni uomo** (Lan et al. 2024)
 - ricavano tutta la loro conoscenza attraverso un processo di **estrazione statistica meramente passiva**
 - i bambini imparano a parlare **interagendo** con altri esseri umani
 - la **probabilità condizionata** è l’unico criterio che determina le risposte del sistema

Un nuovo tipo di entità

“It could thus be argued that in recent years, the field of AI has created machines with new modes of understanding, most likely new species in a larger zoo of related concepts.” (Mitchell e Krakauer 2023)

“Us and Them” (Pink Floyd, *The dark Side of the Moon*)

- In cosa gli LLM sono **simili a noi**:
 - centralità dei meccanismi predittivi nell’elaborazione del linguaggio
 - uso di rappresentazioni vettoriali continue
 - importanza dell’informazione distribuzionale (cf. linguistica strutturalista)
- In cosa gli LLM sono **diversi da noi**:
 - sono addestrati su quantità “**sovraumane**” di dati linguistici
 - i dati per addestrare ChatGPT corrispondono a **36.540 anni uomo** (Lan et al. 2024)
 - ricavano tutta la loro conoscenza attraverso un processo di **estrazione statistica meramente passiva**
 - i bambini imparano a parlare **interagendo** con altri esseri umani
 - la **probabilità condizionata** è l’unico criterio che determina le risposte del sistema

Un nuovo tipo di entità

“It could thus be argued that in recent years, the field of AI has created machines with new modes of understanding, most likely new species in a larger zoo of related concepts.” (Mitchell e Krakauer 2023)

“Us and Them” (Pink Floyd, *The dark Side of the Moon*)

- In cosa gli LLM sono **simili a noi**:
 - centralità dei meccanismi predittivi nell’elaborazione del linguaggio
 - uso di rappresentazioni vettoriali continue
 - importanza dell’informazione distribuzionale (cf. linguistica strutturalista)
- In cosa gli LLM sono **diversi da noi**:
 - sono addestrati su quantità “**sovraumane**” di dati linguistici
 - i dati per addestrare ChatGPT corrispondono a **36.540 anni uomo** (Lan et al. 2024)
 - ricavano tutta la loro conoscenza attraverso un processo di **estrazione statistica meramente passiva**
 - i bambini imparano a parlare **interagendo** con altri esseri umani
 - la **probabilità condizionata** è l’unico criterio che determina le risposte del sistema

Un nuovo tipo di entità

“It could thus be argued that in recent years, the field of AI has created machines with new modes of understanding, most likely new species in a larger zoo of related concepts.” (Mitchell e Krakauer 2023)

“Us and Them” (Pink Floyd, *The dark Side of the Moon*)

- In cosa gli LLM sono **simili a noi**:
 - centralità dei meccanismi predittivi nell’elaborazione del linguaggio
 - uso di rappresentazioni vettoriali continue
 - importanza dell’informazione distribuzionale (cf. linguistica strutturalista)
- In cosa gli LLM sono **diversi da noi**:
 - sono addestrati su quantità “**sovraumane**” di dati linguistici
 - i dati per addestrare ChatGPT corrispondono a **36.540 anni uomo** (Lan et al. 2024)
 - ricavano tutta la loro conoscenza attraverso un processo di **estrazione statistica meramente passiva**
 - i bambini imparano a parlare **interagendo** con altri esseri umani
 - la **probabilità condizionata** è l’unico criterio che determina le risposte del sistema

Un nuovo tipo di entità

“It could thus be argued that in recent years, the field of AI has created machines with new modes of understanding, most likely new species in a larger zoo of related concepts.” (Mitchell e Krakauer 2023)

“Us and Them” (Pink Floyd, *The dark Side of the Moon*)

- In cosa gli LLM sono **simili a noi**:
 - centralità dei meccanismi predittivi nell’elaborazione del linguaggio
 - uso di rappresentazioni vettoriali continue
 - importanza dell’informazione distribuzionale (cf. linguistica strutturalista)

- In cosa gli LLM sono **diversi da noi**:
 - sono addestrati su quantità “**sovraumane**” di dati linguistici
 - i dati per addestrare ChatGPT corrispondono a **36.540 anni uomo** (Lan et al. 2024)
 - ricavano tutta la loro conoscenza attraverso un processo di **estrazione statistica meramente passiva**
 - i bambini imparano a parlare **interagendo** con altri esseri umani
 - la **probabilità condizionata** è l’unico criterio che determina le risposte del sistema

Un nuovo tipo di entità

“It could thus be argued that in recent years, the field of AI has created machines with new modes of understanding, most likely new species in a larger zoo of related concepts.” (Mitchell e Krakauer 2023)

“Us and Them” (Pink Floyd, *The dark Side of the Moon*)

- In cosa gli LLM sono **simili a noi**:
 - centralità dei meccanismi predittivi nell’elaborazione del linguaggio
 - uso di rappresentazioni vettoriali continue
 - importanza dell’informazione distribuzionale (cf. linguistica strutturalista)
- In cosa gli LLM sono **diversi da noi**:
 - sono addestrati su quantità **“sovraumane”** di dati linguistici
 - i dati per addestrare ChatGPT corrispondono a **36.540 anni uomo** (Lan et al. 2024)
 - ricavano tutta la loro conoscenza attraverso un processo di **estrazione statistica meramente passiva**
 - i bambini imparano a parlare **interagendo** con altri esseri umani
 - la **probabilità condizionata** è l’unico criterio che determina le risposte del sistema

Un nuovo tipo di entità

“It could thus be argued that in recent years, the field of AI has created machines with new modes of understanding, most likely new species in a larger zoo of related concepts.” (Mitchell e Krakauer 2023)

“Us and Them” (Pink Floyd, *The dark Side of the Moon*)

- In cosa gli LLM sono **simili a noi**:
 - centralità dei meccanismi predittivi nell’elaborazione del linguaggio
 - uso di rappresentazioni vettoriali continue
 - importanza dell’informazione distribuzionale (cf. linguistica strutturalista)
- In cosa gli LLM sono **diversi da noi**:
 - sono addestrati su quantità **“sovraumane”** di dati linguistici
 - i dati per addestrare ChatGPT corrispondono a **36.540 anni uomo** (Lan et al. 2024)
 - ricavano tutta la loro conoscenza attraverso un processo di **estrazione statistica meramente passiva**
 - i bambini imparano a parlare **interagendo** con altri esseri umani
 - la **probabilità condizionata** è l’unico criterio che determina le risposte del sistema

Un nuovo tipo di entità

“It could thus be argued that in recent years, the field of AI has created machines with new modes of understanding, most likely new species in a larger zoo of related concepts.” (Mitchell e Krakauer 2023)

“Us and Them” (Pink Floyd, *The dark Side of the Moon*)

- In cosa gli LLM sono **simili a noi**:
 - centralità dei meccanismi predittivi nell’elaborazione del linguaggio
 - uso di rappresentazioni vettoriali continue
 - importanza dell’informazione distribuzionale (cf. linguistica strutturalista)
- In cosa gli LLM sono **diversi da noi**:
 - sono addestrati su quantità **“sovraumane”** di dati linguistici
 - i dati per addestrare ChatGPT corrispondono a **36.540 anni uomo** (Lan et al. 2024)
 - ricavano tutta la loro conoscenza attraverso un processo di **estrazione statistica meramente passiva**
 - i bambini imparano a parlare **interagendo** con altri esseri umani
 - la **probabilità condizionata** è l’unico criterio che determina le risposte del sistema

Un nuovo tipo di entità

“It could thus be argued that in recent years, the field of AI has created machines with new modes of understanding, most likely new species in a larger zoo of related concepts.” (Mitchell e Krakauer 2023)

“Us and Them” (Pink Floyd, *The dark Side of the Moon*)

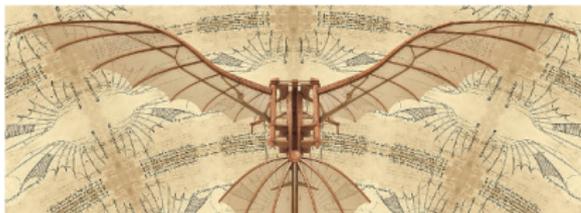
- In cosa gli LLM sono **simili a noi**:
 - centralità dei meccanismi predittivi nell’elaborazione del linguaggio
 - uso di rappresentazioni vettoriali continue
 - importanza dell’informazione distribuzionale (cf. linguistica strutturalista)
- In cosa gli LLM sono **diversi da noi**:
 - sono addestrati su quantità **“sovraumane”** di dati linguistici
 - i dati per addestrare ChatGPT corrispondono a **36.540 anni uomo** (Lan et al. 2024)
 - ricavano tutta la loro conoscenza attraverso un processo di **estrazione statistica meramente passiva**
 - i bambini imparano a parlare **interagendo** con altri esseri umani
 - la **probabilità condizionata** è l’unico criterio che determina le risposte del sistema

Un nuovo tipo di entità

“It could thus be argued that in recent years, the field of AI has created machines with new modes of understanding, most likely new species in a larger zoo of related concepts.” (Mitchell e Krakauer 2023)

Realizzare funzioni della natura senza imitarla

il caso delle macchine volanti



Perché gli LLM sono una “rivoluzione copernicana” nel rapporto tra IA e linguistica?

- Gli LLM hanno (**apparentemente!**) tolto l'uomo dal centro del linguaggio
 - le abilità di comunicazione linguistica sembrano adesso manifestarsi anche in macchine che, seppure ispirate a meccanismi presenti negli esseri umani, però le realizzano in maniera radicalmente differente da ciò che significa imparare e usare il linguaggio per l'uomo
- Gli LLM hanno (**apparentemente!**) tolto la linguistica dal centro della ricerca dell'IA sulle “macchine parlanti”

Linguistica e IA *post LLM*

C'è un posto per la linguistica nel nuovo universo creato dagli LLM?

Perché gli LLM sono una “rivoluzione copernicana” nel rapporto tra IA e linguistica?

- Gli LLM hanno (**apparentemente!**) tolto l’uomo dal centro del linguaggio
 - le abilità di comunicazione linguistica sembrano adesso manifestarsi anche in macchine che, seppure ispirate a meccanismi presenti negli esseri umani, però le realizzano in maniera radicalmente differente da ciò che significa imparare e usare il linguaggio per l’uomo
- Gli LLM hanno (**apparentemente!**) tolto la linguistica dal centro della ricerca dell’IA sulle “macchine parlanti”

Linguistica e IA *post LLM*

C’è un posto per la linguistica nel nuovo universo creato dagli LLM?

Perché gli LLM sono una “rivoluzione copernicana” nel rapporto tra IA e linguistica?

- Gli LLM hanno (**apparentemente!**) tolto l’uomo dal centro del linguaggio
 - le abilità di comunicazione linguistica sembrano adesso manifestarsi anche in macchine che, seppure ispirate a meccanismi presenti negli esseri umani, però le realizzano in maniera radicalmente differente da ciò che significa imparare e usare il linguaggio per l’uomo
- Gli LLM hanno (**apparentemente!**) tolto la linguistica dal centro della ricerca dell’IA sulle “macchine parlanti”

Linguistica e IA *post LLM*

C’è un posto per la linguistica nel nuovo universo creato dagli LLM?

Linguistica e IA *post LLM*

- Gli LLM sono ottimi **imitatori** dei testi umani ma sono spesso pessimi a livello dei ragionamenti (Mahowald et al. 2024)
- Gli LLM hanno **capacità di generalizzazione** diverse da quelle umane (Yiu et al. 2023)
 - sono bravi nel gestire le strutture che hanno maggiore probabilità di aver osservato nella fase di addestramento, ma hanno una minore capacità di generalizzazione rispetto a eventi più rari (McCoy et al. 2023)

Il ruolo della linguistica nell'IA *post LLM*

1. La linguistica per capire se quanto queste macchine parlano realmente come noi

Linguistica e IA *post LLM*

- Gli LLM sono ottimi **imitatori** dei testi umani ma sono spesso pessimi a livello dei ragionamenti (Mahowald et al. 2024)
- Gli LLM hanno **capacità di generalizzazione** diverse da quelle umane (Yiu et al. 2023)
 - sono bravi nel gestire le strutture che hanno maggiore probabilità di aver osservato nella fase di addestramento, ma hanno una minore capacità di generalizzazione rispetto a eventi più rari (McCoy et al. 2023)

Il ruolo della linguistica nell'IA *post LLM*

1. La linguistica per capire se quanto queste macchine parlano realmente come noi

Linguistica e IA *post LLM*

- Gli LLM sono ottimi **imitatori** dei testi umani ma sono spesso pessimi a livello dei ragionamenti (Mahowald et al. 2024)
- Gli LLM hanno **capacità di generalizzazione** diverse da quelle umane (Yiu et al. 2023)
 - sono bravi nel gestire le strutture che hanno maggiore probabilità di aver osservato nella fase di addestramento, ma hanno una minore capacità di generalizzazione rispetto a eventi più rari (McCoy et al. 2023)

Il ruolo della linguistica nell'IA *post LLM*

1. La linguistica per capire se quanto queste macchine parlano realmente come noi

Linguistica e IA *post LLM*

- Manca una chiara **teoria esplicativa** del perché gli LLM hanno successo o sbagliano
- Non sappiamo **quali informazioni** sono codificate nelle rappresentazioni interne degli LLM
 - gli LLM come sistemi “black box”
- Manca ancora una **teoria causale** di come gli stati interni della macchina determinino il suo comportamento

Il ruolo della linguistica nell'IA *post LLM*

2. La linguistica per creare modelli esplicativi delle conoscenze e dei comportamenti degli LLM

Linguistica e IA *post LLM*

- Manca una chiara **teoria esplicativa** del perché gli LLM hanno successo o sbagliano
- Non sappiamo **quali informazioni** sono codificate nelle rappresentazioni interne degli LLM
 - gli LLM come sistemi “black box”
- Manca ancora una **teoria causale** di come gli stati interni della macchina determinino il suo comportamento

Il ruolo della linguistica nell'IA *post LLM*

2. La linguistica per creare modelli esplicativi delle conoscenze e dei comportamenti degli LLM

Linguistica e IA *post LLM*

- Manca una chiara **teoria esplicativa** del perché gli LLM hanno successo o sbagliano
- Non sappiamo **quali informazioni** sono codificate nelle rappresentazioni interne degli LLM
 - gli LLM come sistemi “black box”
- Manca ancora una **teoria causale** di come gli stati interni della macchina determinino il suo comportamento

Il ruolo della linguistica nell'IA *post LLM*

2. La linguistica per creare modelli esplicativi delle conoscenze e dei comportamenti degli LLM

Linguistica e IA *post LLM*

- Manca una chiara **teoria esplicativa** del perché gli LLM hanno successo o sbagliano
- Non sappiamo **quali informazioni** sono codificate nelle rappresentazioni interne degli LLM
 - gli LLM come sistemi “black box”
- Manca ancora una **teoria causale** di come gli stati interni della macchina determinino il suo comportamento

Il ruolo della linguistica nell'IA *post LLM*

2. La linguistica per creare modelli esplicativi delle conoscenze e dei comportamenti degli LLM

Linguistica e IA *post LLM*

- Il linguaggio umano è il risultato di un processo di adattamento evolutivo in cui una funzionalità comunicativa altamente sofisticata si è sviluppata **ottimizzando l'uso di limitate risorse cognitive e fisiologiche**
- Gli LLM raggiungono competenze linguistiche **quasi-umane**, ma con risorse **sovraumane**
 - volumi insostenibili di dati linguistici e risorse computazionali

Il ruolo della linguistica nell'IA *post LLM*

3. La linguistica per individuare i **bias induttivi** e gli **a priori cognitivi** che consentono agli esseri umani di acquisire il linguaggio ottimizzando l'uso di dati e risorse cognitive

Linguistica e IA *post LLM*

- Il linguaggio umano è il risultato di un processo di adattamento evolutivo in cui una funzionalità comunicativa altamente sofisticata si è sviluppata **ottimizzando l'uso di limitate risorse cognitive e fisiologiche**
- Gli LLM raggiungono competenze linguistiche **quasi-umane**, ma con risorse **sovraumane**
 - volumi insostenibili di dati linguistici e risorse computazionali

Il ruolo della linguistica nell'IA *post LLM*

3. La linguistica per individuare i **bias induttivi** e gli **a priori cognitivi** che consentono agli esseri umani di acquisire il linguaggio ottimizzando l'uso di dati e risorse cognitive

Linguistica e IA *post LLM*

- Il linguaggio umano è il risultato di un processo di adattamento evolutivo in cui una funzionalità comunicativa altamente sofisticata si è sviluppata **ottimizzando l'uso di limitate risorse cognitive e fisiologiche**
- Gli LLM raggiungono competenze linguistiche **quasi-umane**, ma con risorse **sovraumane**
 - volumi insostenibili di dati linguistici e risorse computazionali

Il ruolo della linguistica nell'IA *post LLM*

3. La linguistica per individuare i **bias induttivi** e gli **a priori cognitivi** che consentono agli esseri umani di acquisire il linguaggio ottimizzando l'uso di dati e risorse cognitive

Linguistica e IA *post LLM*

- Il linguaggio umano è il risultato di un processo di adattamento evolutivo in cui una funzionalità comunicativa altamente sofisticata si è sviluppata **ottimizzando l'uso di limitate risorse cognitive e fisiologiche**
- Gli LLM raggiungono competenze linguistiche **quasi-umane**, ma con risorse **sovraumane**
 - volumi insostenibili di dati linguistici e risorse computazionali

Il ruolo della linguistica nell'IA *post LLM*

3. La linguistica per individuare i **bias induttivi** e gli **a priori cognitivi** che consentono agli esseri umani di acquisire il linguaggio ottimizzando l'uso di dati e risorse cognitive

Mechanistic Interpretability

- Individuare un legame causale tra specifiche componenti delle reti neurali, feature che codificano, e i comportamenti dei modelli
 - **feature** = elementi di informazione che sono “human-interpretable”
- Il problema della **polisemanticità**: molti neuroni rispondono a molteplici feature dell’input che sono del tutto irrelate tra di loro
 - cf. **entanglement** e **sovrapposizione** nella meccanica quantistica
- È necessario individuare un mapping tra le rappresentazioni continue ed “entangled” dei modelli neurali e rappresentazioni con feature discrete
 - cf. Random Sparse Autoencoder

Linguistic Alignment

- Allineare gli LLM con le feature e le strutture linguistiche del linguaggio umano e le appropriate relazioni causali
- L’alignment linguistico deve essere guidato dai modelli teorici e dai dati sperimentali delle scienze del linguaggio

Mechanistic Interpretability

- Individuare un legame causale tra specifiche componenti delle reti neurali, feature che codificano, e i comportamenti dei modelli
 - **feature** = elementi di informazione che sono “human-interpretable”
- Il problema della **polisemanticità**: molti neuroni rispondono a molteplici feature dell’input che sono del tutto irrelate tra di loro
 - cf. **entanglement** e **sovrapposizione** nella meccanica quantistica
- È necessario individuare un mapping tra le rappresentazioni continue ed “entangled” dei modelli neurali e rappresentazioni con feature discrete
 - cf. Random Sparse Autoencoder

Linguistic Alignment

- Allineare gli LLM con le feature e le strutture linguistiche del linguaggio umano e le appropriate relazioni causali
- L’alignment linguistico deve essere guidato dai modelli teorici e dai dati sperimentali delle scienze del linguaggio

Mechanistic Interpretability

- Individuare un legame causale tra specifiche componenti delle reti neurali, feature che codificano, e i comportamenti dei modelli
 - **feature** = elementi di informazione che sono “human-interpretable”
- Il problema della **polisemanticità**: molti neuroni rispondono a molteplici feature dell’input che sono del tutto irrelate tra di loro
 - cf. **entanglement** e **sovrapposizione** nella meccanica quantistica
- È necessario individuare un mapping tra le rappresentazioni continue ed “entangled” dei modelli neurali e rappresentazioni con feature discrete
 - cf. Random Sparse Autoencoder

Linguistic Alignment

- Allineare gli LLM con le feature e le strutture linguistiche del linguaggio umano e le appropriate relazioni causali
- L’alignment linguistico deve essere guidato dai modelli teorici e dai dati sperimentali delle scienze del linguaggio

Mechanistic Interpretability

- Individuare un legame causale tra specifiche componenti delle reti neurali, feature che codificano, e i comportamenti dei modelli
 - **feature** = elementi di informazione che sono “human-interpretable”
- Il problema della **polisemanticità**: molti neuroni rispondono a molteplici feature dell’input che sono del tutto irrelate tra di loro
 - cf. **entanglement** e **sovrapposizione** nella meccanica quantistica
- È necessario individuare un mapping tra le rappresentazioni continue ed “entangled” dei modelli neurali e rappresentazioni con feature discrete
 - cf. Random Sparse Autoencoder

Linguistic Alignment

- Allineare gli LLM con le feature e le strutture linguistiche del linguaggio umano e le appropriate relazioni causali
- L’alignment linguistico deve essere guidato dai modelli teorici e dai dati sperimentali delle scienze del linguaggio

Mechanistic Interpretability

- Individuare un legame causale tra specifiche componenti delle reti neurali, feature che codificano, e i comportamenti dei modelli
 - **feature** = elementi di informazione che sono “human-interpretable”
- Il problema della **polisemanticità**: molti neuroni rispondono a molteplici feature dell’input che sono del tutto irrelate tra di loro
 - cf. **entanglement** e **sovrapposizione** nella meccanica quantistica
- È necessario individuare un mapping tra le rappresentazioni continue ed “entangled” dei modelli neurali e rappresentazioni con feature discrete
 - cf. Random Sparse Autoencoder

Linguistic Alignment

- Allineare gli LLM con le feature e le strutture linguistiche del linguaggio umano e le appropriate relazioni causali
- L’alignment linguistico deve essere guidato dai modelli teorici e dai dati sperimentali delle scienze del linguaggio

Baby LM

- Sviluppare LM che siano più cognitivamente realistici per dimensione e quantità di input di addestramento
- **PRIN 2022 - CLEVER**: sviluppo di Baby LM per l'italiano addestrati su dati linguistici “ecologici” in quantità incrementale
 - **I livello (età prescolare)** - training data: ca. 60M di token di trascrizione di parlato e di film per bambini, libri per l'infanzia, ecc.
 - **II livello (ca. 10 anni)** - training data: ca. 100M di token inclusivi di trascrizioni di libri, manuali scolastici, trascrizione di programmi TV e film, ecc.
- **Obiettivi**:
 - uso di metodi di **curriculum learning** basati su evidenze relative al percorso acquisizionale del linguaggio umano
 - studio di **learning bias cognitivi** da fornire ai modelli neurali

Baby LM

- Sviluppare LM che siano più cognitivamente realistici per dimensione e quantità di input di addestramento
- **PRIN 2022 - CLEVER**: sviluppo di Baby LM per l'italiano addestrati su dati linguistici “ecologici” in quantità incrementale
 - **I livello (età prescolare)** - training data: ca. 60M di token di trascrizione di parlato e di film per bambini, libri per l'infanzia, ecc.
 - **II livello (ca. 10 anni)** - training data: ca. 100M di token inclusivi di trascrizioni di libri, manuali scolastici, trascrizione di programmi TV e film, ecc.
- Obiettivi:
 - uso di metodi di **curriculum learning** basati su evidenze relative al percorso acquisizionale del linguaggio umano
 - studio di **learning bias cognitivi** da fornire ai modelli neurali

Baby LM

- Sviluppare LM che siano più cognitivamente realistici per dimensione e quantità di input di addestramento
- **PRIN 2022 - CLEVER**: sviluppo di Baby LM per l'italiano addestrati su dati linguistici “ecologici” in quantità incrementale
 - **I livello (età prescolare)** - training data: ca. 60M di token di trascrizione di parlato e di film per bambini, libri per l'infanzia, ecc.
 - **II livello (ca. 10 anni)** - training data: ca. 100M di token inclusivi di trascrizioni di libri, manuali scolastici, trascrizione di programmi TV e film, ecc.
- **Obiettivi**:
 - uso di metodi di **curriculum learning** basati su evidenze relative al percorso acquisizionale del linguaggio umano
 - studio di **learning bias cognitivi** da fornire ai modelli neurali

Linguaggio e pensiero

- Mahowald et al. (2024) “Dissociating language and thought in large language models: A cognitive perspective. *Trends in Cognitive Sciences*, distinguono tra:
 - **competenza linguistica formale**, come conoscenza di regole e schemi linguistici
 - **competenza funzionale**
 - formal reasoning (ragionamento logico e novel problem solving), conoscenza del mondo, situation modeling (capacità di costruire rappresentazioni incrementali di situazioni, stories), social reasoning (Theory of Mind)
- Gli LLMs hanno una competenza formale superficiale e una ricca conoscenza di fatti e “fattoidi”, ma mancano di molteplici competenze funzionali nell’uso del linguaggio, nel ragionamento e nella comunicazione

Costruire macchine parlanti o macchine onniscienti?

- gli LLM imparano simultaneamente e dalla medesima fonte le **strutture linguistiche** e **tutto quello che sanno del mondo**
- gli esseri umani sviluppano le competenze linguistico-comunicative prima, con meno risorse, e indipendentemente dalle conoscenze enciclopediche del mondo

Linguaggio e pensiero

- Mahowald et al. (2024) “Dissociating language and thought in large language models: A cognitive perspective. *Trends in Cognitive Sciences*, distinguono tra:
 - **competenza linguistica formale**, come conoscenza di regole e schemi linguistici
 - **competenza funzionale**
 - formal reasoning (ragionamento logico e novel problem solving), **conoscenza del mondo**, **situation modeling** (capacità di costruire rappresentazioni incrementali di situazioni, stories), **social reasoning** (Theory of Mind)
- Gli LLMs hanno una competenza formale superficiale e una ricca conoscenza di **fatti e “fattoidi”**, ma mancano di molteplici competenze funzionali nell’uso del linguaggio, nel ragionamento e nella comunicazione

Costruire macchine parlanti o macchine onniscienti?

- gli LLM imparano simultaneamente e dalla medesima fonte le **strutture linguistiche** e **tutto quello che sanno del mondo**
- gli esseri umani sviluppano le competenze linguistico-comunicative prima, con meno risorse, e indipendentemente dalle conoscenze enciclopediche del mondo

Linguaggio e pensiero

- Mahowald et al. (2024) “Dissociating language and thought in large language models: A cognitive perspective. *Trends in Cognitive Sciences*, distinguono tra:
 - **competenza linguistica formale**, come conoscenza di regole e schemi linguistici
 - **competenza funzionale**
 - formal reasoning (ragionamento logico e novel problem solving), **conoscenza del mondo**, **situation modeling** (capacità di costruire rappresentazioni incrementali di situazioni, stories), **social reasoning** (Theory of Mind)
- Gli LLMs hanno una competenza formale superficiale e una ricca conoscenza di **fatti e “fattoidi”**, ma mancano di molteplici competenze funzionali nell’uso del linguaggio, nel ragionamento e nella comunicazione

Costruire macchine parlanti o macchine onniscienti?

- gli LLM imparano simultaneamente e dalla medesima fonte le **strutture linguistiche** e **tutto quello che sanno del mondo**
- gli esseri umani sviluppano le competenze linguistico-comunicative prima, con meno risorse, e indipendentemente dalle conoscenze enciclopediche del mondo

Linguaggio e pensiero

- Mahowald et al. (2024) “Dissociating language and thought in large language models: A cognitive perspective. *Trends in Cognitive Sciences*, distinguono tra:
 - **competenza linguistica formale**, come conoscenza di regole e schemi linguistici
 - **competenza funzionale**
 - formal reasoning (ragionamento logico e novel problem solving), **conoscenza del mondo**, **situation modeling** (capacità di costruire rappresentazioni incrementali di situazioni, stories), **social reasoning** (Theory of Mind)
- Gli LLMs hanno una competenza formale superficiale e una ricca conoscenza di **fatti e “fattoidi”**, ma mancano di molteplici competenze funzionali nell’uso del linguaggio, nel ragionamento e nella comunicazione

Costruire macchine parlanti o macchine onniscienti?

- gli LLM imparano simultaneamente e dalla medesima fonte le **strutture linguistiche** e **tutto quello che sanno del mondo**
- gli esseri umani sviluppano le competenze linguistico-comunicative prima, con meno risorse, e indipendentemente dalle conoscenze enciclopediche del mondo

Linguaggio e pensiero

- Mahowald et al. (2024) “Dissociating language and thought in large language models: A cognitive perspective. *Trends in Cognitive Sciences*, distinguono tra:
 - **competenza linguistica formale**, come conoscenza di regole e schemi linguistici
 - **competenza funzionale**
 - formal reasoning (ragionamento logico e novel problem solving), **conoscenza del mondo**, **situation modeling** (capacità di costruire rappresentazioni incrementali di situazioni, stories), **social reasoning** (Theory of Mind)
- Gli LLMs hanno una competenza formale superficiale e una ricca conoscenza di **fatti e “fattoidi”**, ma mancano di molteplici competenze funzionali nell’uso del linguaggio, nel ragionamento e nella comunicazione

Costruire macchine parlanti o macchine onniscienti?

- gli LLM imparano simultaneamente e dalla medesima fonte le **strutture linguistiche** e **tutto quello che sanno del mondo**
- gli esseri umani sviluppano le competenze linguistico-comunicative prima, con meno risorse, e indipendentemente dalle conoscenze enciclopediche del mondo

“Quand’è che abbiamo smesso di capire il mondo?”

B. Labatut

- Abbiamo bisogno della linguistica per capire la nuova generazione di “macchine parlanti”, ma questo sarà possibile solo con **nuove teorie del linguaggio umano**
 - i successi e gli errori degli LLM possono portarci ad abbandonare molti degli attuali presupposti sul linguaggio
- Abbiamo bisogno di una teoria del linguaggio umano che sappia essa stessa “parlare” il linguaggio delle macchine: **vettori, matrici e probabilità**
 - è lo stesso linguaggio del sostrato neuronale che sostiene la capacità del linguaggio
- Abbiamo bisogno di una teoria del linguaggio umano che sia più vicina alla **meccanica quantistica** che al cartesianesimo della fisica classica

“Quand’è che abbiamo smesso di capire il mondo?”

B. Labatut

- Abbiamo bisogno della linguistica per capire la nuova generazione di “macchine parlanti”, ma questo sarà possibile solo con **nuove teorie del linguaggio umano**
 - i successi e gli errori degli LLM possono portarci ad abbandonare molti degli attuali presupposti sul linguaggio
- Abbiamo bisogno di una teoria del linguaggio umano che sappia essa stessa “parlare” il linguaggio delle macchine: **vettori, matrici e probabilità**
 - è lo stesso linguaggio del sostrato neuronale che sostiene la capacità del linguaggio
- Abbiamo bisogno di una teoria del linguaggio umano che sia più vicina alla **meccanica quantistica** che al cartesianesimo della fisica classica

“Quand’è che abbiamo smesso di capire il mondo?”

B. Labatut

- Abbiamo bisogno della linguistica per capire la nuova generazione di “macchine parlanti”, ma questo sarà possibile solo con **nuove teorie del linguaggio umano**
 - i successi e gli errori degli LLM possono portarci ad abbandonare molti degli attuali presupposti sul linguaggio
- Abbiamo bisogno di una teoria del linguaggio umano che sappia essa stessa “parlare” il linguaggio delle macchine: **vettori, matrici e probabilità**
 - è lo stesso linguaggio del sostrato neuronale che sostiene la capacità del linguaggio
- Abbiamo bisogno di una teoria del linguaggio umano che sia più vicina alla **meccanica quantistica** che al cartesianesimo della fisica classica

Grazie!!!