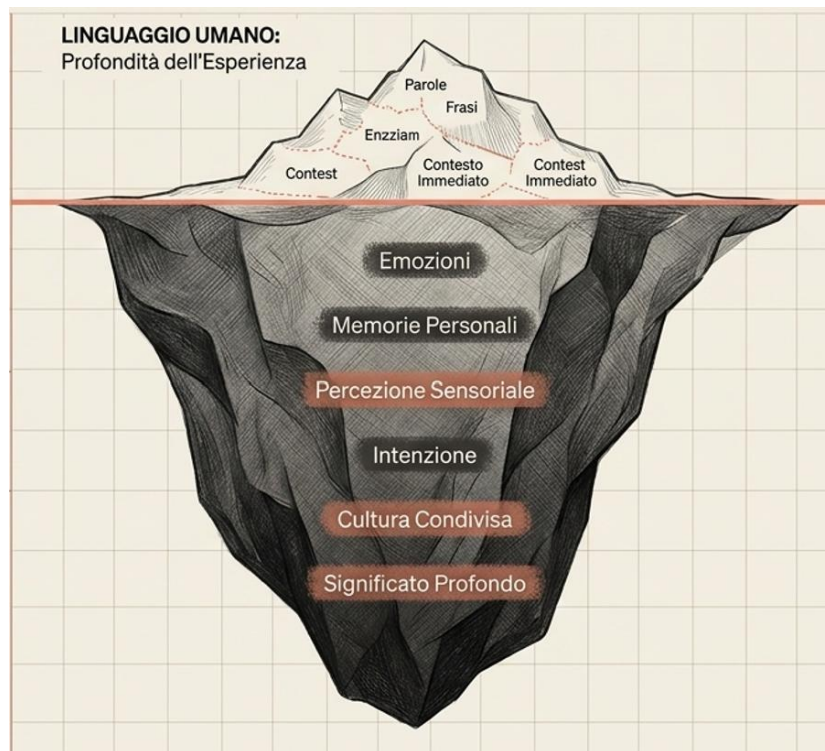


La grammatica non è più quella di una volta...

Cosa ci insegna l'intelligenza artificiale



Francesco Cutugno

Università Federico II di Napoli

Como – Villa del Grumello 21 maggio 2026

Grammatica

Vocabolario on line

DAL VOCABOLARIO

grammatica

grammatica (ant. **gramàtica**) s. f. [dal lat. *grammatīca*, gr. γραμματική (τέχνη), dall'agg. γραμματικός v. grammatico]. – **1. a.** Il complesso delle norme che costituiscono il particolare modo di essere di una lingua (o di un dialetto), cioè il suo sistema fonemico, morfologico, sintattico, considerato nella sua totalità, come si riflette di volta in volta nelle singole espressioni: *la g. italiana, tedesca, francese; l'evoluzione della g. inglese*. **b.** Rappresentazione sistematica di una lingua e dei suoi elementi costitutivi, articolata tradizionalmente in fonologia, morfologia, sintassi, lessicologia ed etimologia; anche, la disciplina stessa (considerata nel medioevo una delle «arti del trivio») che ha per oggetto tale rappresentazione: *insegnare, studiare la g.; maestro di grammatica*.

Il problema della astrattezza delle regole rispetto alla concretezza dell'informazione linguistica

Regole Astratte (Grammatica Tradizionale)



Modelli teorici fissi, spesso decontestualizzati dall'uso reale. Cercano di imporre una logica universale che non sempre rispecchia la fluidità della lingua parlata e scritta.

Informazione Linguistica Concreta (Uso Reale)



Dati linguistici vivi, ricchi di sfumature e dipendenti dal contesto. Emergono da miliardi di esempi pratici, rivelando schemi che le regole astratte faticano a catturare.

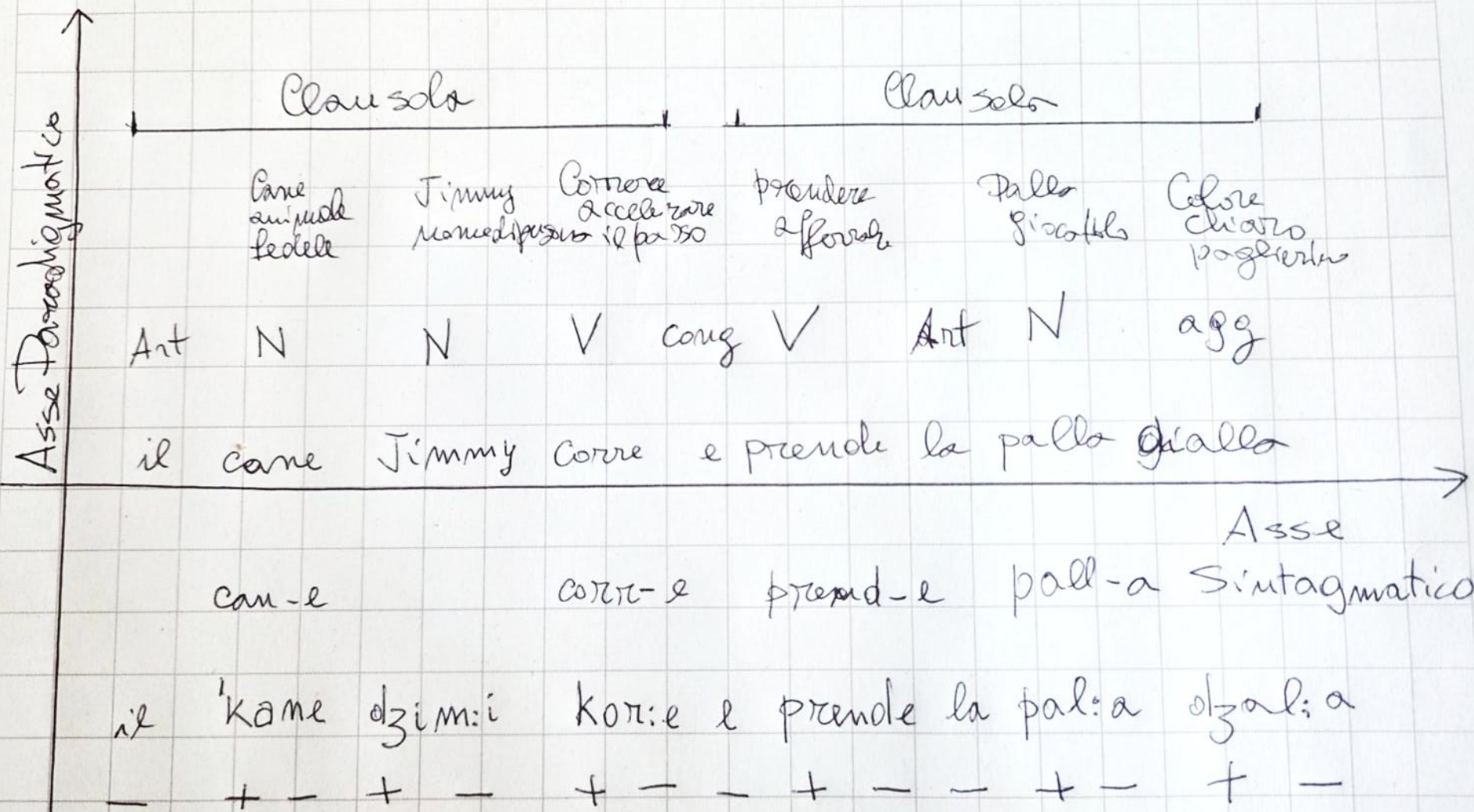
Esercizi di Grammatica Tradizionale: L'Approccio Analitico

Analisi Logica

Scomposizione della frase semplice nei suoi costituenti fondamentali: soggetto, predicato e complementi. Obiettivo è identificare la funzione sintattica di ciascun elemento all'interno della proposizione.

Analisi del Periodo

Scomposizione della frase complessa (periodo) nelle diverse proposizioni che la compongono. Obiettivo è individuare la proposizione principale e le diverse proposizioni subordinate, comprendendone i rapporti logici.





I neonati apprendono le lingue e la conoscenza del mondo grazie a:

- Multimodalità
- Multigranularità

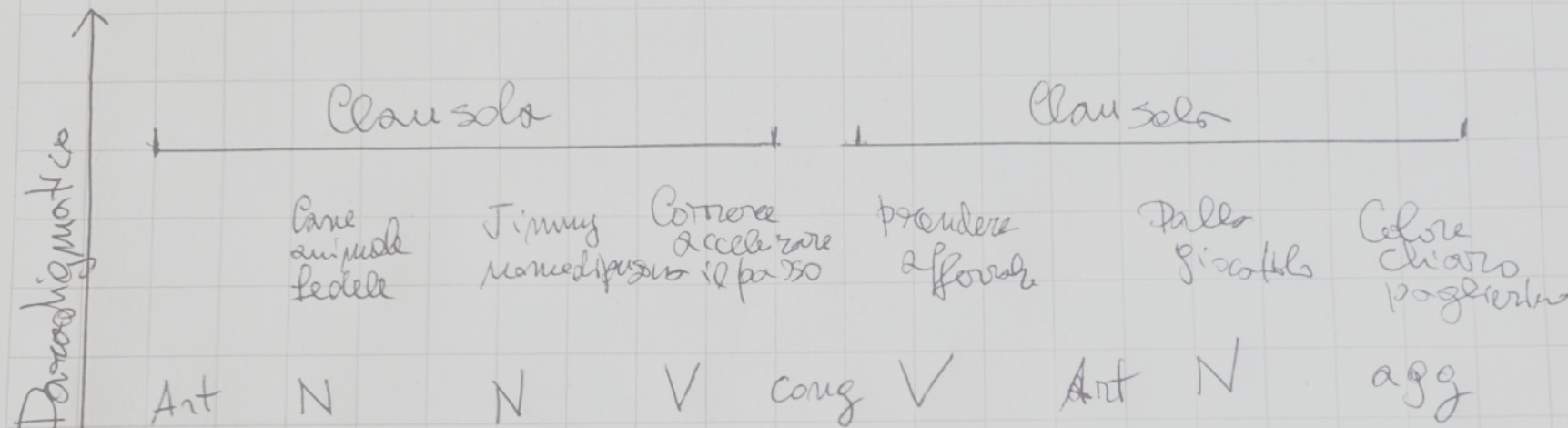
Sono esposti a informazione puramente sintagmatica su molti «canali» contemporaneamente e con diverse dimensioni delle unità minime di analisi

Proprietà paradigmatiche

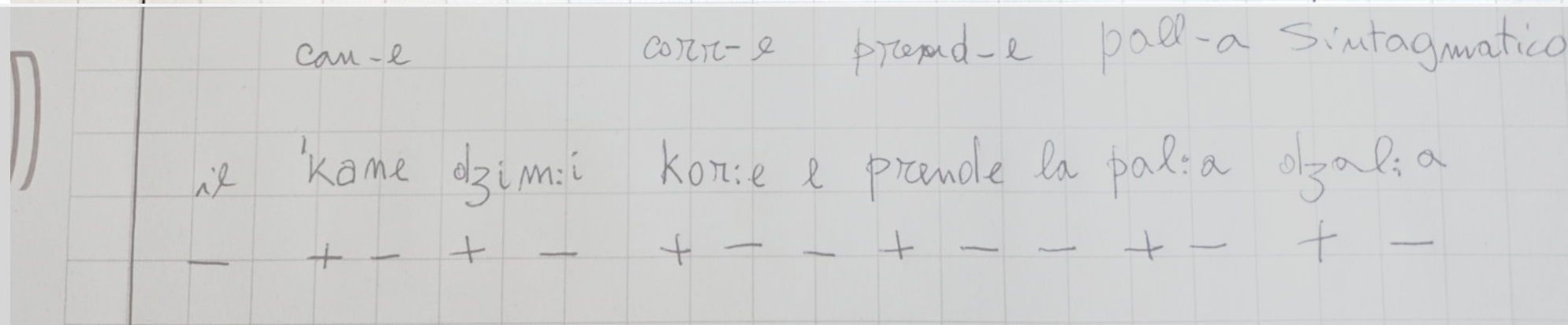
- Costruzione di una rappresentazione «ontologica» della conoscenza

Individuano unità minime di «senso» e apprendono la loro organizzazione in sequenze

Sulle unità definite organizzano «verticalmente» il piano paradigmatico dei significati e delle relazioni fra questi



Asse
il cane Jimmy corre e prende la palla gialla
Asse



Studies in
NATURAL
LANGUAGE
PROCESSING

DISTRIBUTIONAL SEMANTICS



Alessandro Lenci and Magnus Sahlgren

1. every language has a structure of its own, and there are no universal linguistic categories;
2. the study of language must be primarily synchronic;
3. linguistics must be autonomous with respect to other disciplines, especially psychology.

Distributional Hypothesis

The semantic similarity between two lexemes is a function of the similarity of their linguistic contexts.

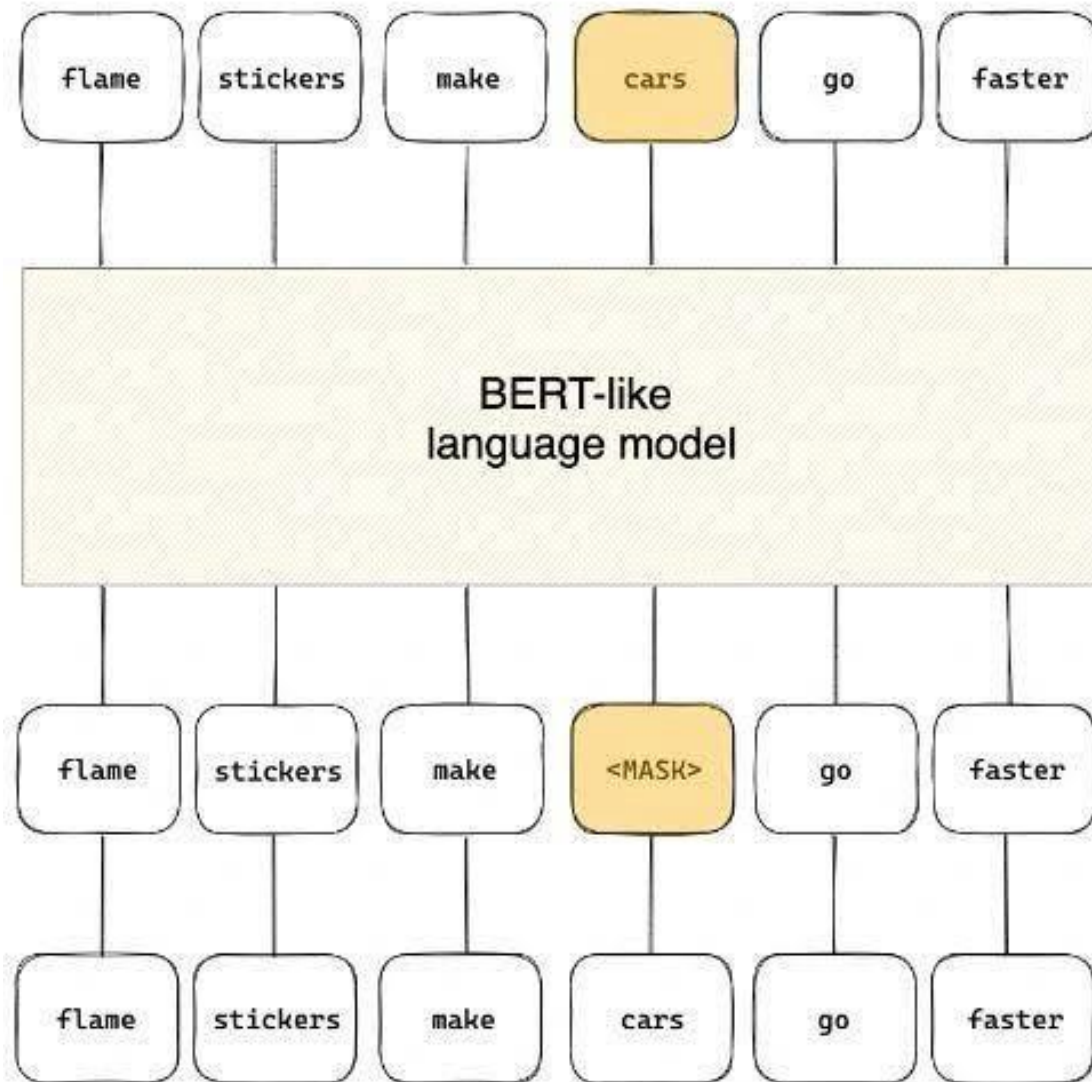
Linguistic elements are identified by their **distributions**, defined as the sum of the contexts in which they occur.

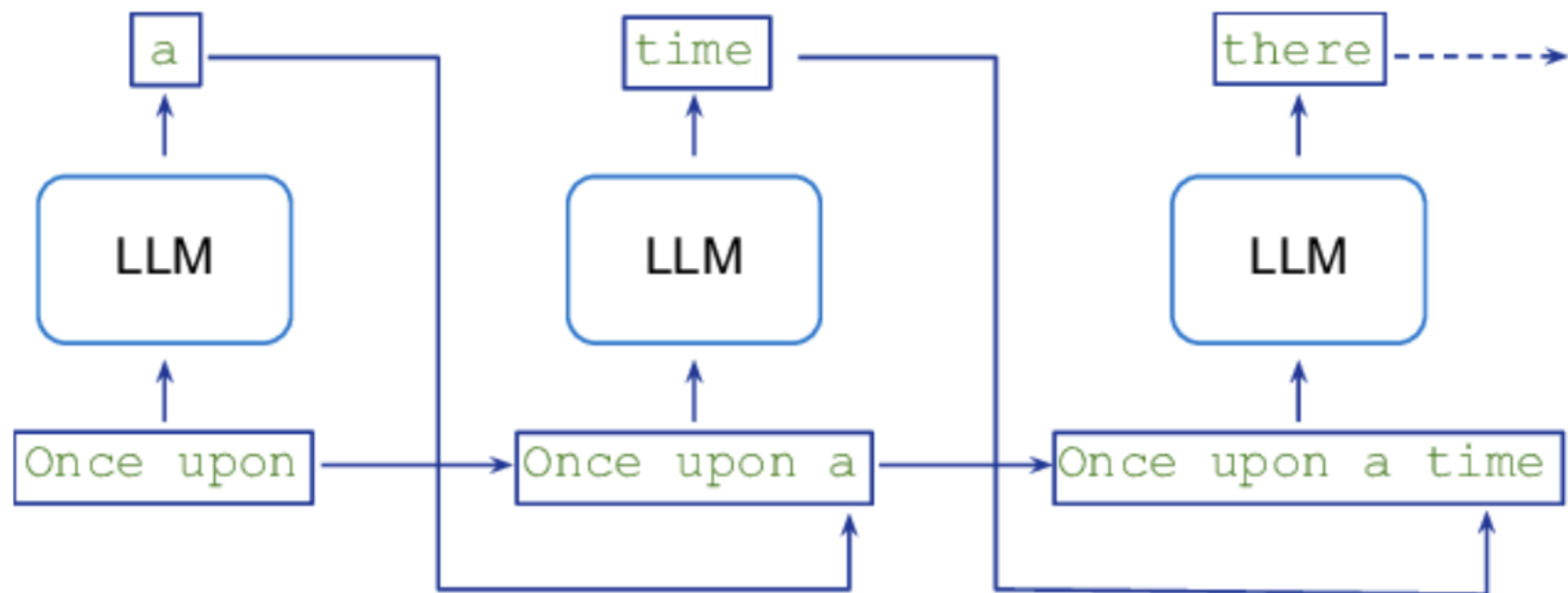


Durante l'addestramento X e Y sono conosciute, W è incognita

In esercizio Y è incognita, W e X sono conosciute

Masked Language Modeling MLM





Self - attention

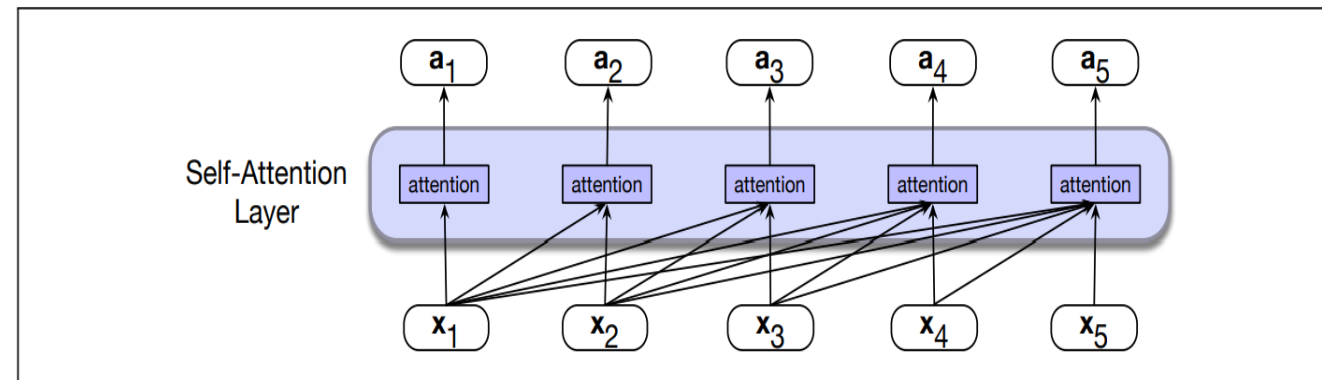
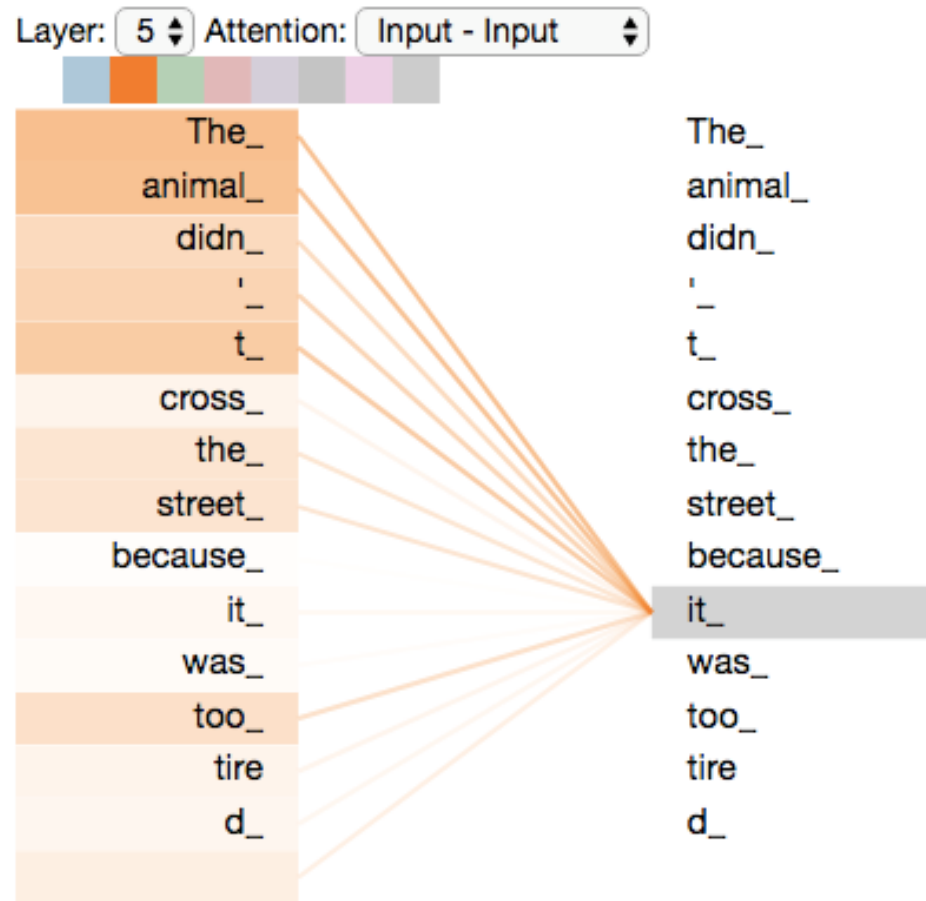
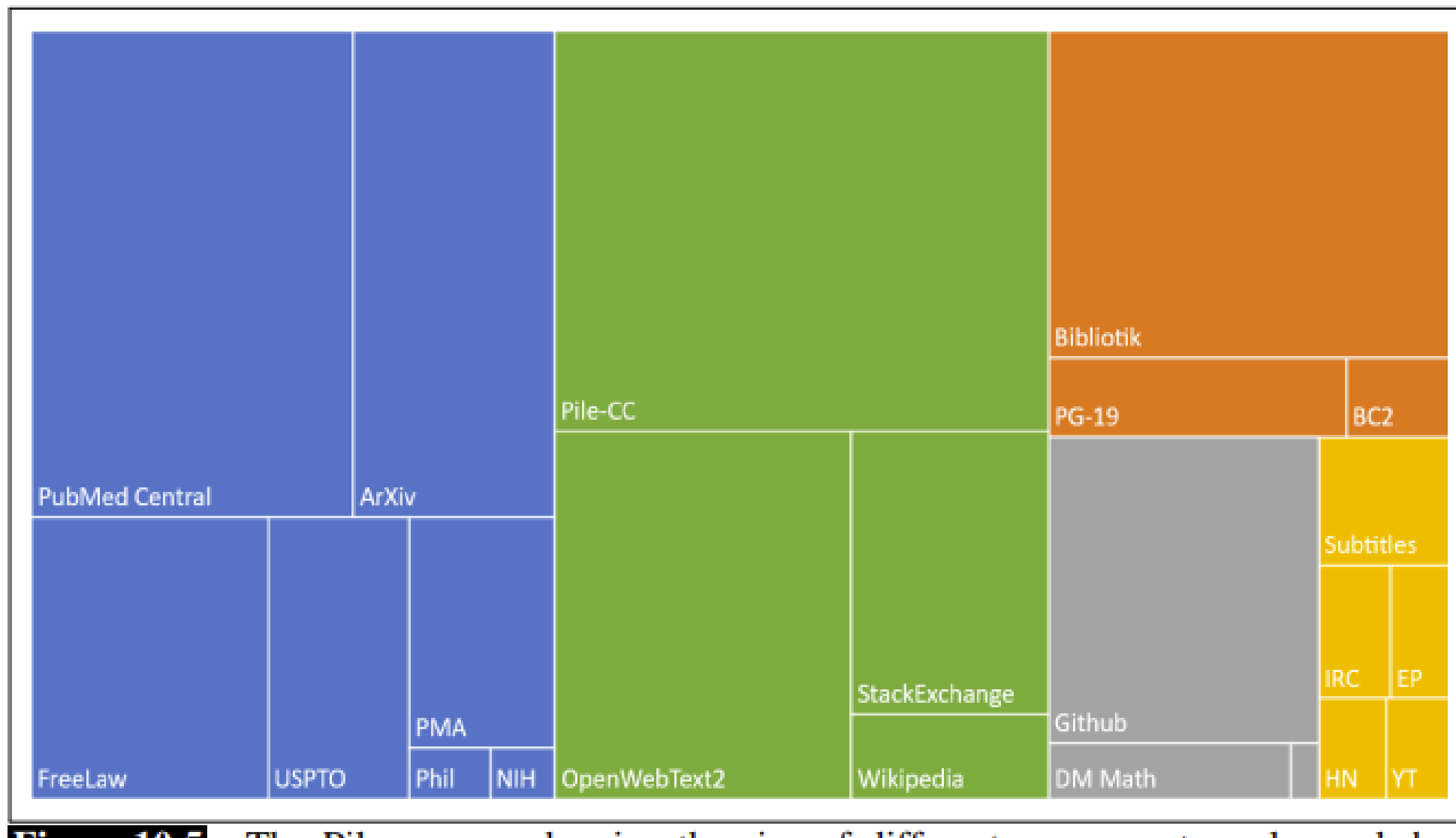


Figure 8.4 Information flow in causal self-attention. When processing each input x_i , the model attends to all the inputs up to, and including x_i .

The Pile (825GB)



1. Tempo di lettura per 300 miliardi di parole:

- $300.000.000.000 \text{ parole} \div 150 \text{ parole/minuto} \approx 2.000.000.000 \text{ minuti}$

2. Conversione in anni:

- $2.000.000.000 \text{ minuti} \div (60 \text{ minuti} \times 24 \text{ ore} \times 365 \text{ giorni}) \approx \text{circa } 3.800 \text{ anni}$

Risultato:

Ci vorrebbero circa 3.800 anni di lettura continua, ad alta voce, senza pause, per leggere tutto il testo che ha contribuito ad addestrare GPT-3.

Nota: Questa è una stima molto approssimativa e serve solo a dare un'idea della vastità del corpus. La quantità di dati è enormemente grande e rappresenta un insieme di testi molto variegato e complesso.

L'intelligenza generativa è un sofisticato sistema di autocompletamento statistico.

Mascheramento (Masking)

Il gatto si siede sul [MASK].

Una parola viene nascosta per sfidare il modello a prevederla.

Valutazione del Contesto Linguistico

Il modello analizza le parole circostanti ("Il gatto si siede sul") per comprendere la relazione grammaticale e semantica.

Analisi Statistica

Il modello non analizza la logica, ma calcola la probabilità di combinazione delle parole nei contesti reali d'uso.

Conoscenza del Mondo

Utilizza dati di addestramento per sapere dove i gatti si siedono solitamente.

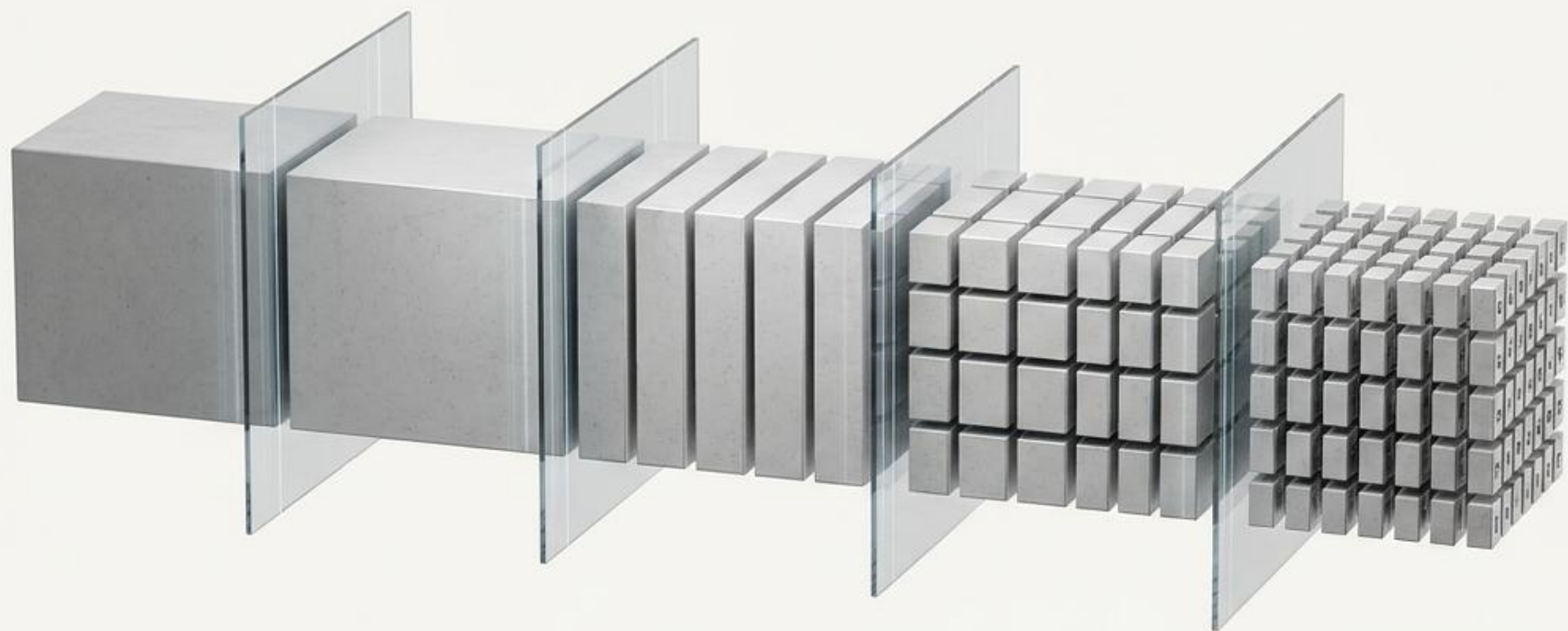
Previsione

La funzione principale è prevedere la parola più probabile per riempire il [MASK], basandosi sull'analisi del contesto e su milioni di esempi precedenti.

...tappeto [85%]

...divano [10%]

...tetto [5%]



PAROLE

SUBWORD

SILLABE

CARATTERI

La lente distribuzionale

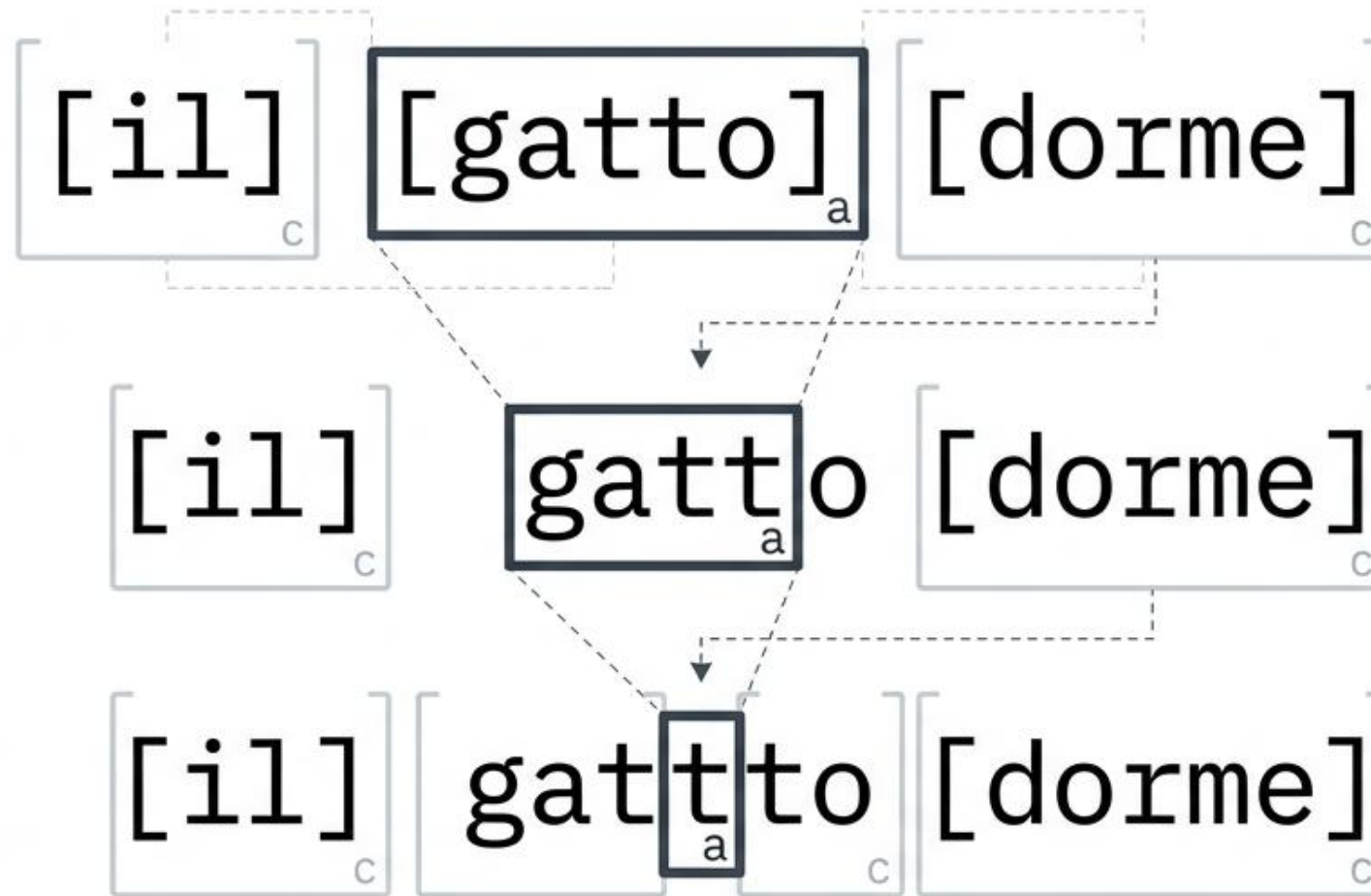
Oltre i confini della parola: un'analisi comparativa delle architetture NLP a granularità variabile



Forma	Parola	Subword tipo BPE	Sillabe	Caratteri
gatto	gatto	gatt + o	gat + to	g a t t o
gatti	gatti	gatt + i	gat + ti	g a t t i
gattone	gattone	gatt + one	gat + to + ne	g a t t o n e



La semantica appartiene al contesto, non alla parola isolata



Nel senso classico della linguistica distribuzionale, un elemento è definito dalla totalità dei suoi ambienti di occorrenza. Se formalizzata in sequenze e probabilità, la costruzione matematica è granularity-agnostic. Le proprietà distribuzionali dipendono dalla definizione di contesto, non dallo statuto lessicale dell'unità.

Le proprietà semantiche contestuali valgono per tutti i livelli di granularità dell'informazione linguistica, dalle parole alle subword ai singoli caratteri

Più astratto



PAROLE
(word level)

Il **banco** è vicino alla finestra.

Unità: parola

Significato contestuale

"banco" = tavolo da scuola



SUBWORD
(subword level)

Il **ban** **co** è vicino alla finestra.

Unità: subword (es. BPE)

Significato contestuale

"ban" + "co" → insieme evoca il concetto di tavolo da scuola



CARATTERI
(character level)

Il **b** **a** **n** **co** è vicino alla finestra.

Unità: carattere

Significato contestuale

La sequenza di caratteri "b-a-n-c-o" attiva il concetto di tavolo da scuola



LIVELLO PIÙ FINE
(byte / char level)

Il **banco** è vicino alla finestra.

82 61 6E 63 6F

Unità: byte / codice (es. UTF-8)

Significato contestuale

Anche a livello di codice, la sequenza corrispondente contribuisce al significato

01100010
01100001
01101110
01100011
01101111

Più fine



Il significato non è legato al livello, ma emerge dal contesto.

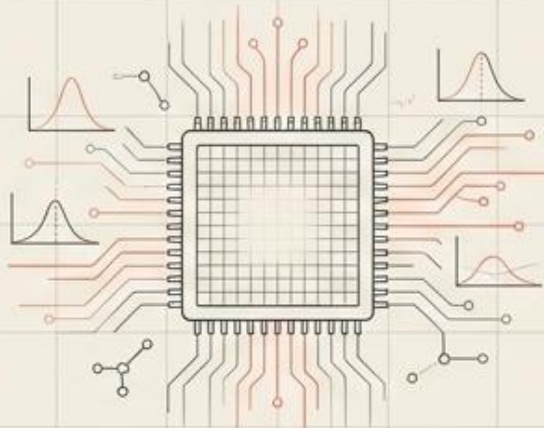
Le rappresentazioni a tutti i livelli (parole, subword, caratteri, byte) veicolano e attivano proprietà semantiche contestuali.

La correttezza formale nasce dall'elaborazione di probabilità su scala massiva.



1. Ingestione Dati (Il Materiale Grezzo)

Addestramento su ingenti quantità di dati linguistici reali. Il sistema assorbe milioni di testi, conversazioni e documenti.



2. Potenza di Calcolo (Il Motore)

Impiego di un'infrastruttura di calcolo rilevante per elaborare tutta l'informazione necessaria. Trasforma il testo in vettori statistici.



3. Schemi Emergenti (Il Risultato)

Il sistema non memorizza frasi, ma apprende regolarità e schemi probabilistici che simulano la padronanza della lingua.

Il problema della astrattezza delle regole rispetto alla concretezza dell'informazione linguistica

Regole Astratte (Grammatica Tradizionale)



Modelli teorici fissi, spesso decontestualizzati dall'uso reale. Cercano di imporre una logica universale che non sempre rispecchia la fluidità della lingua parlata e scritta.

Informazione Linguistica Concreta (Uso Reale)



Dati linguistici vivi, ricchi di sfumature e dipendenti dal contesto. Emergono da miliardi di esempi pratici, rivelando schemi che le regole astratte faticano a catturare.

- Dal parlato parlato allo scritto scritto su asse continuo (Nencioni 1976)
- Voghera (2022) multimodalità

Caratteristiche	Parlato	Scritto	Scritto digitale non continuo	Scritto dialogico
Canale	Fonico-uditivo Gestuale-visivo	Grafico-visivo	Grafico-visivo Fonico-uditivo Audiovisivo	Grafico-visivo Fonico-uditivo Audiovisivo
Produzione e ricezione	In presenza	In assenza	In assenza	In assenza
	In tempo reale	Differita	Differita	Differita
	Continua	Discontinua	Discontinua	Discontinua
	Non pianificata	Pianificata	Pianificata	± Pianificata
Rapporto tra produttore e ricevente	Dialogo	Monologo	Monologo	Dialogo/Monologo
Codici	Lingua, gesti	Lingua	Lingua, calcoli, foto, filmati, musica...	Lingua, foto, filmati, musica...
Testo verbale	Discontinuo in turni	Continuo	Discontinuo senza turnazione	Discontinuo in turni

Un possibile circolo vizioso...

- osserviamo gli usi linguistici;
- da questi inferiamo una grammatica;
- la grammatica guida ciò che consideriamo “corretto”;
- ma i dati che osserviamo sono già filtrati da pratiche sociali, interpretazioni e convenzioni;
- quindi la grammatica non è mai completamente indipendente dall’uso che pretende di descrivere.
- E i modelli statistici moderni (LLM inclusi) accentuano il problema: apprendono regolarità sintagmatiche dai testi senza necessariamente “sapere” cosa sia una regola grammaticale nel senso tradizionale.

La soluzione normativa: postulare una grammatica ideale

È la via classica delle accademie linguistiche e della grammatica prescrittiva. Si decide che:

- alcuni usi sono “corretti”, altri no;
- le regole valgono indipendentemente dalla variabilità empirica.

Il vantaggio:

- stabilità;
- insegnabilità;
- Interoperabilità sociale.

Il limite:

- arbitrarietà inevitabile;
- congelamento di un uso storico particolare;
- esclusione delle variazioni reali.

Questa soluzione rompe il circolo imponendo un punto fermo esterno, ma lo fa per autorità, non per necessità logica.

La soluzione strutturalista: la grammatica è una regolarità emergente

- la lingua non è un insieme di regole eterne;
- è un sistema di differenze stabilizzato socialmente.

Qui la grammatica non “precede” l’uso ma emerge statisticamente dall’uso collettivo.

In questo quadro non c’è davvero un circolo vizioso: c’è un sistema dinamico auto-organizzato.

La grammatica è una compressione delle regolarità osservate.

Gli LLM funzionano sorprendentemente bene proprio perché sfruttano questa proprietà emergente.

Ovviamente questo non vuol dire che non abbiamo bisogno di una grammatica, solo che va ripensata con regole meno rigide.

La soluzione pragmatica/contemporanea: il significato nasce dall'uso coordinato

- ❖ Una grammatica non è vera o falsa;
- ❖ Le regole linguistiche sopravvivono se:
 - riducono l'ambiguità;
 - stabilizzano aspettative comunicative.
- ❖ In questa prospettiva:
 - “riconoscere la grammatica” esplicitamente non è necessario;
 - basta partecipare correttamente al gioco linguistico.

E infatti gli LLM mostrano qualcosa di importante:
si può produrre linguaggio altamente grammaticale senza possedere regole simboliche esplicite nel senso tradizionale.

Questo non prova che la grammatica non esista.

La Resistenza al Cambiamento Tecnologico: Echi Storici



La Critica alla Scrittura (Platone, Fedro)

Nel Fedro, Platone esprime il timore che la scrittura indebolisca la memoria e crei un'apparenza di sapienza senza vera comprensione, una 'droga' per la memoria.



L'Invenzione della Stampa

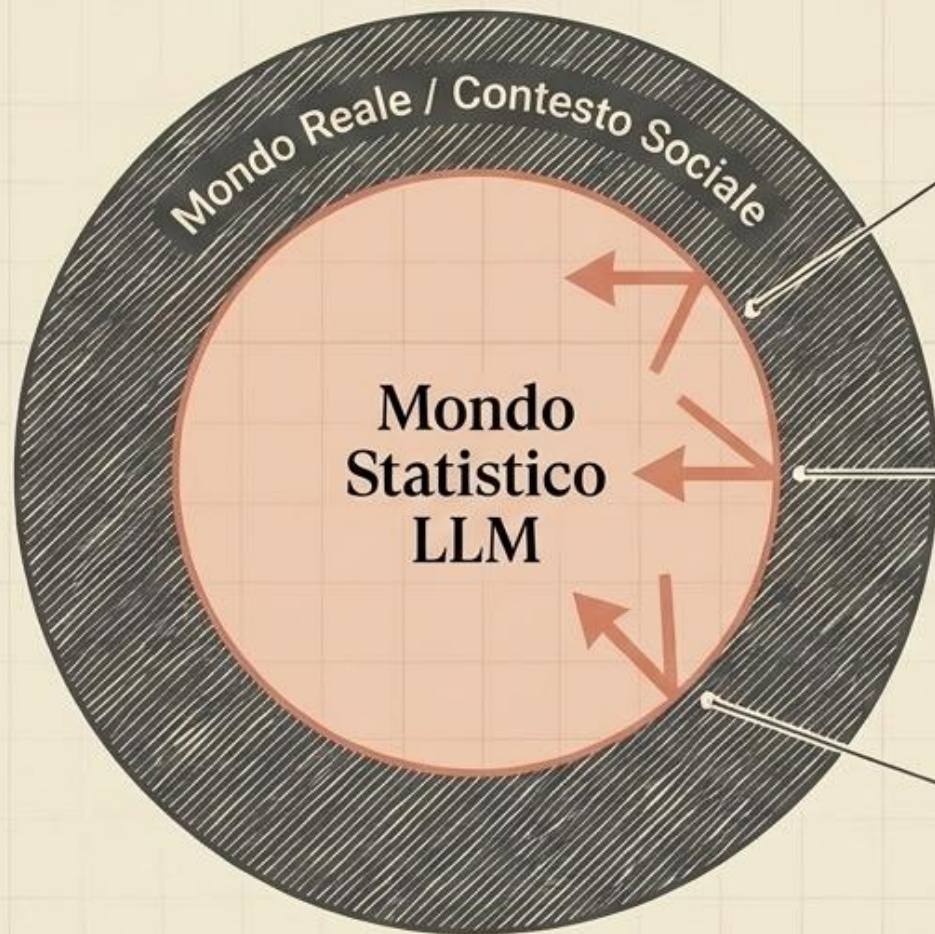
La diffusione della stampa generò timori di un sovraccarico di informazioni, la propagazione di idee considerate pericolose e la perdita del controllo sulla conoscenza.



La Calcolatrice e la Matematica

L'introduzione delle calcolatrici nelle scuole ha suscitato un intenso dibattito tra gli educatori riguardo all'impatto sulle competenze aritmetiche di base e sul calcolo mentale.

L'assenza di vera comprensione genera allucinazioni e riproduce pregiudizi sistemici.



Le Allucinazioni

Invenzione di fatti e concetti presentati con un'assoluta e impeccabile sicurezza grammaticale, privi di aderenza alla realtà.

Il Problema dei Bias

Riproduzione acritica e amplificazione dei pregiudizi culturali e sociali già presenti nei dati di addestramento originali.

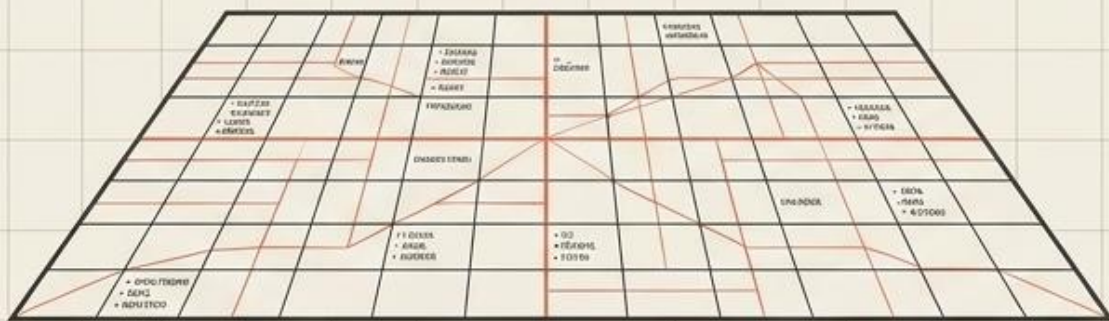
Il Vuoto Cognitivo

Totale assenza di una reale cognizione del mondo fisico e del contesto sociale in cui avviene la comunicazione umana.

www.halupedia.com :-D

L'intenzionalità e l'esperienza vissuta rimangono un'esclusiva del linguaggio umano.

LINGUAGGIO MACCHINA: Superficie Statistica



I limiti dell'Intelligenza Artificiale aiutano a chiarire, per contrasto, cosa distingue veramente il nostro linguaggio.

Dimensione Pragmatica

Noi comunichiamo radicati in un contesto sociale reale, adattando il messaggio a chi ci ascolta.

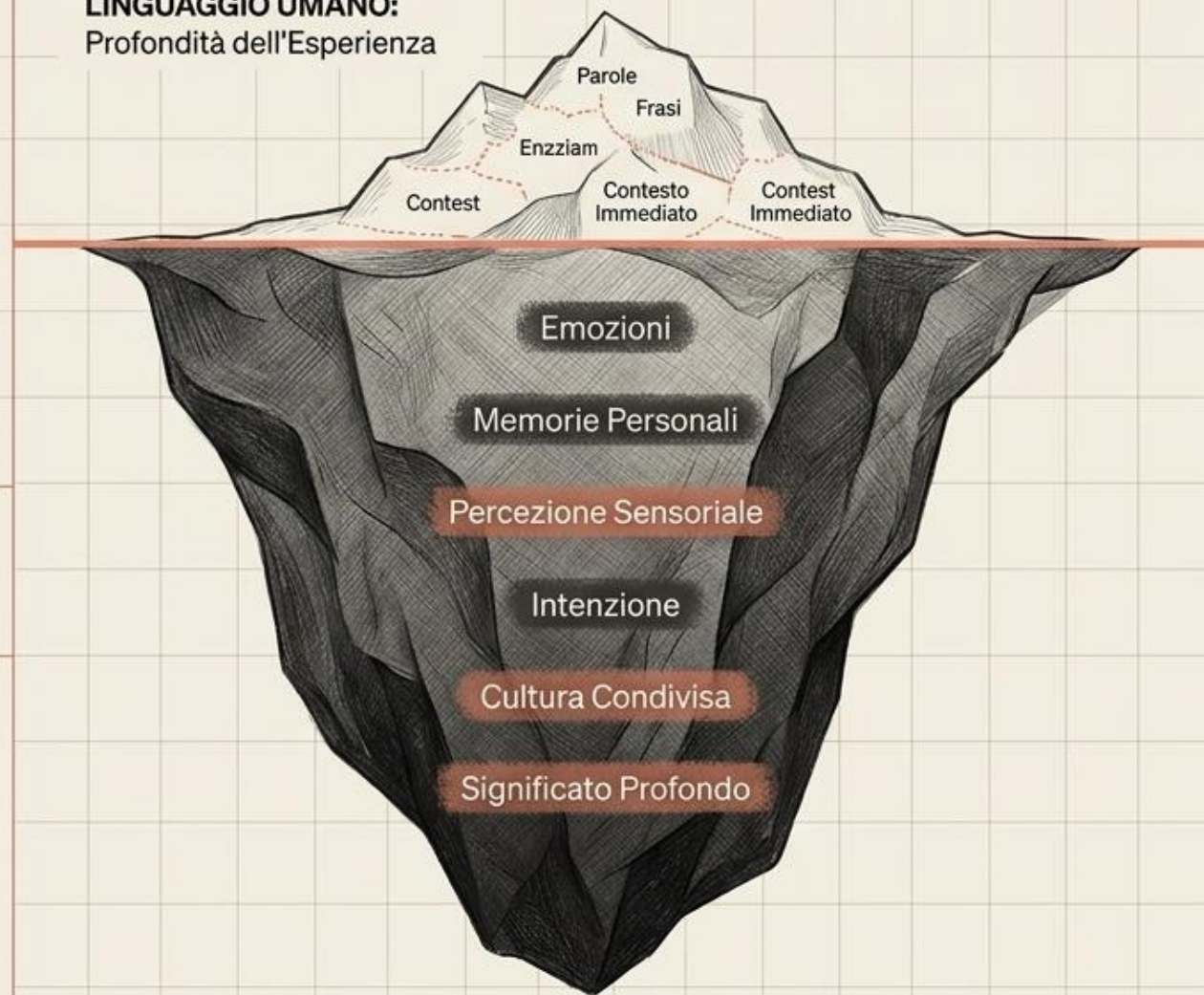
L'Esperienza

Le nostre parole sono legate indissolubilmente alla nostra percezione del mondo fisico e alle nostre emozioni.

L'Intenzionalità

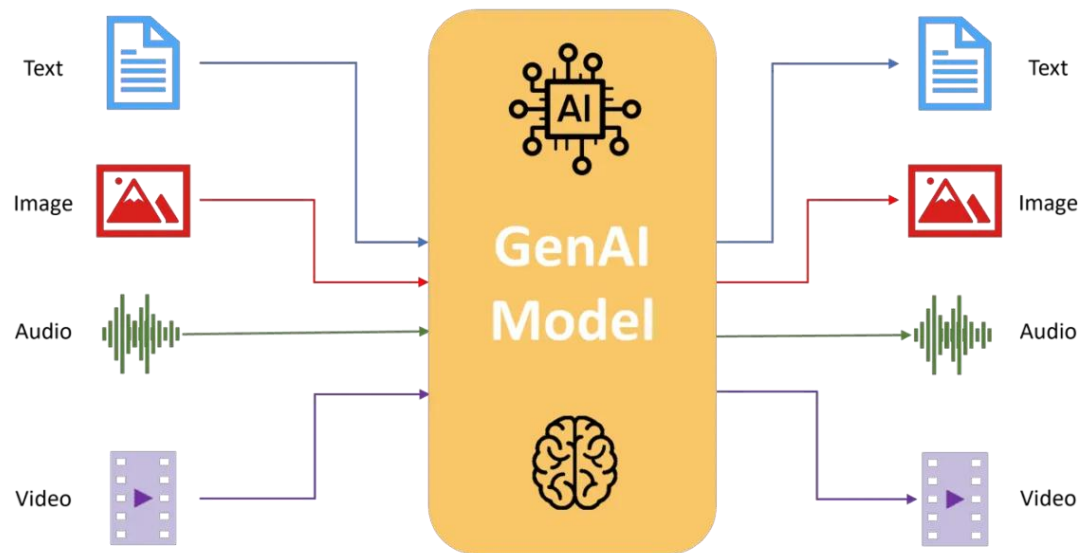
Non prevediamo la prossima parola; parliamo mossi da un fine, da un'intenzione comunicativa profonda che la statistica non può replicare.

LINGUAGGIO UMANO: Profondità dell'Esperienza



Quando gli LLM riescono a guardare il mondo:





How Multimodal AI Works

Understanding Modalities: Text, Image, Audio, Video, and Sensor Data

The Architecture Behind Multimodal AI Models



Training Multimodal AI: Challenges and Techniques

How Multimodal AI Enhances Contextual Understanding

The Role of Large Language Models (LLMs) in Multimodal AI

Cosa non cambierà

Importanza della pragmatica

La pragmatica interpreta il significato nel contesto e richiede emozioni, intenzioni e competenze culturali uniche agli umani.



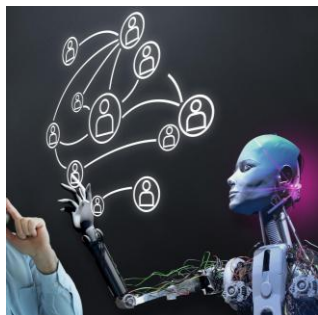
Lingua come fenomeno sociale

La lingua costruisce relazioni, identità e comunità, aspetti che nessun modello statistico può replicare efficacemente.



Centralità dell'esperienza umana

L'esperienza umana resta al centro della comunicazione; l'IA è uno strumento ma non un sostituto dell'intelligenza naturale.



Un esempio divertente...

- <https://www.instagram.com/tardanifederico/reel/DSvEPSnDF5s/>

Epistemicità: I Modelli Linguistici “Pensano”?

La Posizione di Quattrociocchi



Nei LLM esistono rappresentazioni interne stabili e manipolabili associate a pattern statistici latenti, ma non una vera “rappresentazione del mondo”. Il passaggio da strutture interne correlate a comprensione è una sovrainterpretazione: manca l'accesso referenziale, il *grounding epistemico* e la capacità autonoma di validazione.



La riduzione degli LLM a semplici predittori probabilistici è contestata dalla ricerca recente. La definizione stessa di “comprensione” è complessa: si pensi alla fisica teorica, che prescinde dall'esperienza empirica immediata. Inoltre, la capacità dell'IA di dimostrare teoremi va oltre la mera generazione di testo convincente. Il dibattito tra esperti è aperto e richiede prudenza intellettuale, non posizioni semplicistiche.

L'Incertezza Fondamentale

Non abbiamo una chiara idea dei processi mentali che regolano l'apprendimento, la produzione e la percezione del linguaggio umano.



Il termine «cognitivo» è superato, come lo sono i termini «cibernetica» e «intelligenza artificiale». Lo studio separato di percezione, deduzione, apprendimento, e così via – cioè della cognizione –, non si consolida in una unità disciplinare ben definita e dalla crescita bene indirizzata. Risulta quindi opportuno aprire l'indagine – come di fatto sta accadendo – ad altri, più problematici aspetti della mente: emozione, processi inconsci, intenzionalità, attitudini, dolore. In una parola: coscienza.

Trautteur, Giuseppe. "Undici tesi sulla scienza cognitiva." *Adelphiana* 1 (2002): 71-96.

Macchine e Coscienza: Il Dibattito

È possibile?

Come e Quando?

E Se...?



Il dibattito fondamentale:
le macchine possono
sviluppare una vera
esperienza soggettiva?

Funzionalismo vs. Biologia

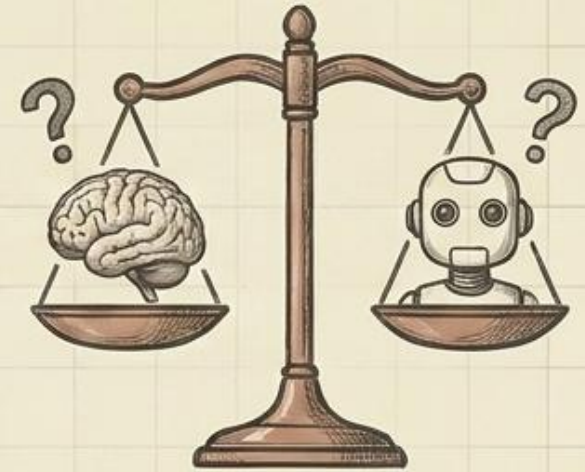


Simulazione funzionale o
necessità di un substrato biologico?



**Il "Problema Difficile"
& Test di Turing**

Superare la mera imitazione
del comportamento intelligente.



L'incertezza rimane. La
coscienza potrebbe essere
un fenomeno emergente o
un'esclusiva biologica?