Machine translation, problem solving and pattern recognition
Early criticisms revisited

Stefania Centrone
Technical University of Munich

Cosimo Perini Brogi
IMT School for Advanced Studies Lucca

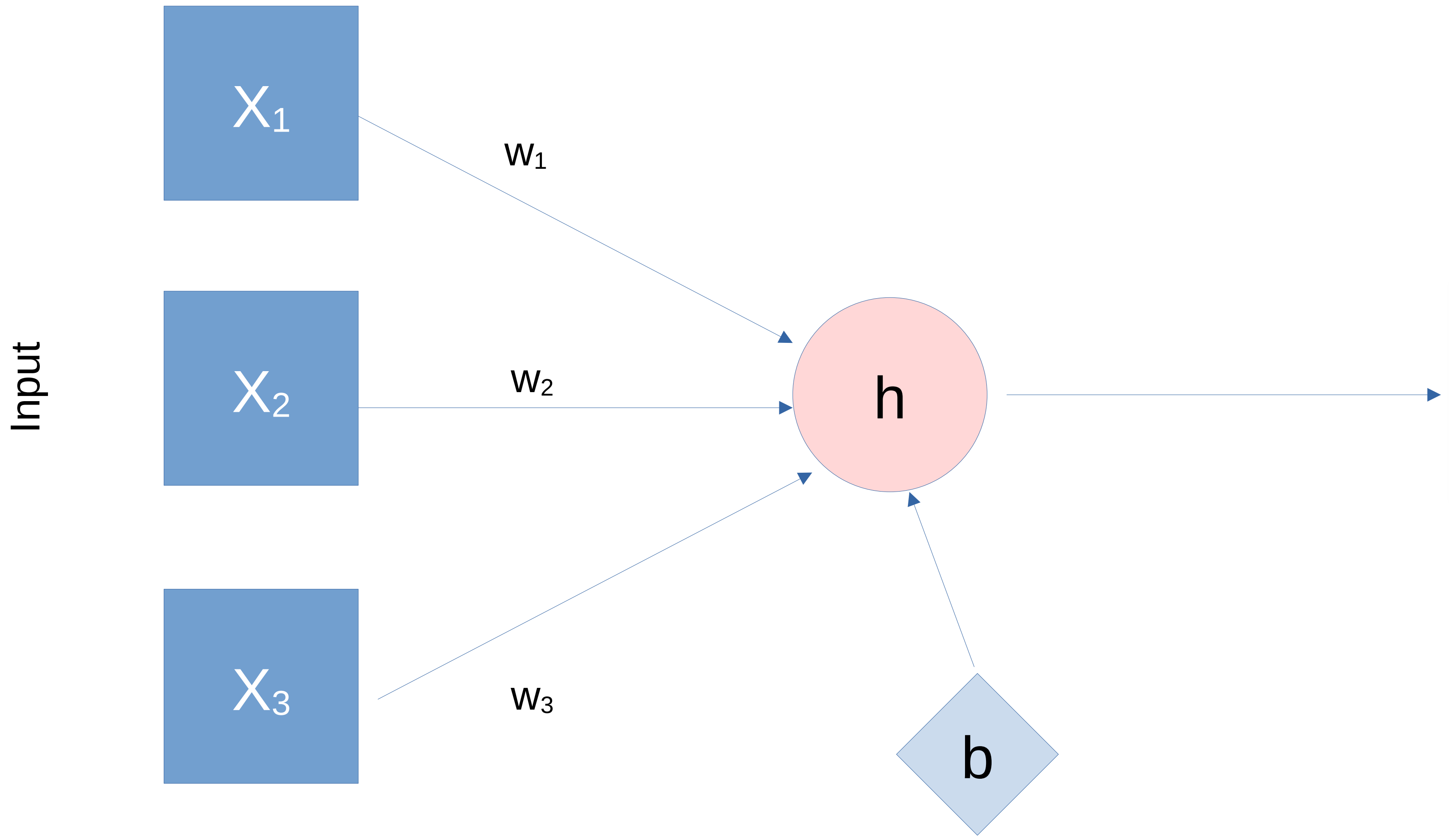
Munich Centre for Mathematical Philosophy, Logic Seminar

5 December 2024

Conclusive Part

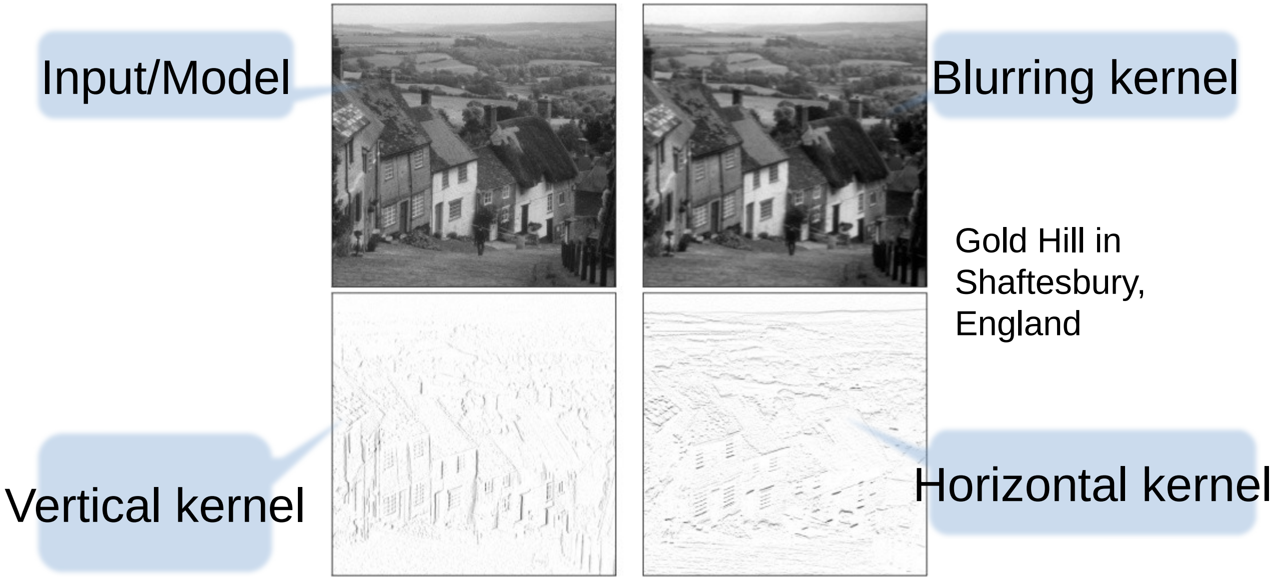
Neural Networks

The Unit



$$h\left(\left(\sum_{i=1}^{3} x_i w_i\right) + b\right)$$

Convolutional Neural Networks

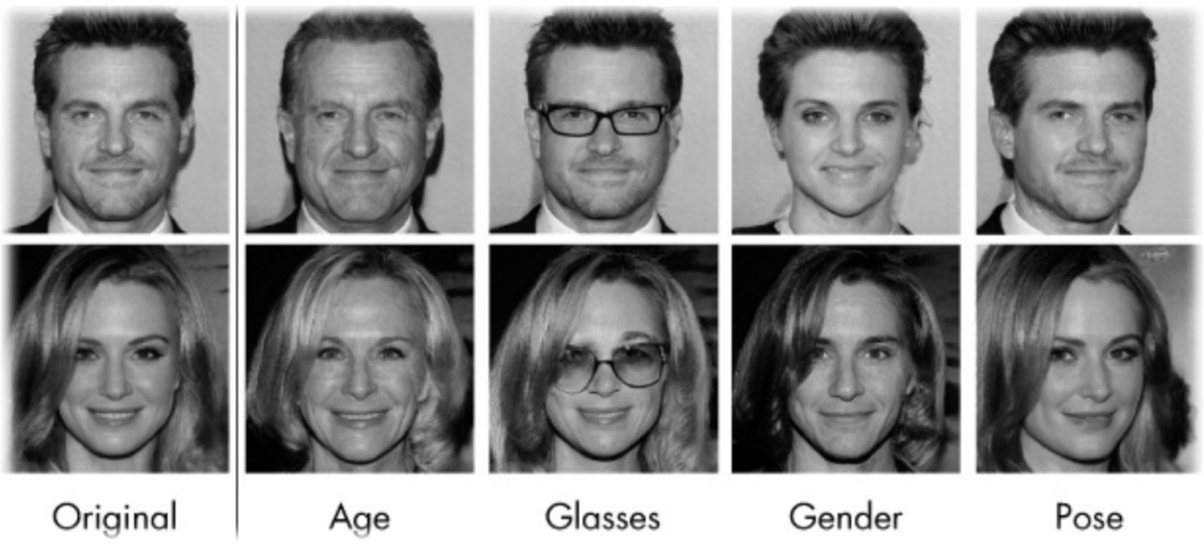


Gold Hill in Shaftesbury, England

- Slide a small square, called a kernel, over the image.
- Movement occurs from top to bottom and left to right.
- At each position, pixel values within the kernel are multiplied by corresponding kernel values.
- The products are summed to produce a single output pixel value for that position.

$$\begin{matrix} 60 & 58 & 60 & 60 & 60 & 60 & 60 & 52 \\ 68 & 60 & 60 & 68 & 68 & 52 & 76 & 76 \\ 76 & 44 & 60 & 60 & 68 & 52 & 60 & 52 \\ 68 & 68 & 76 & 76 & 68 & 52 & 60 & 44 \\ 92 & 84 & 84 & 84 & 76 & 44 & 60 & 44 \\ 76 & 68 & 84 & 76 & 76 & 52 & 60 & 52 \\ 68 & 60 & 60 & 76 & 84 & 52 & 84 & 60 \\ 60 & 60 & 68 & 68 & 68 & 52 & 76 & 68 \end{matrix} \times \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} -60 & -120 & -68 \\ 0 & 0 & 0 \\ 68 & 152 & 76 \end{bmatrix} = \boxed{48}$$

Generative Adversarial Networks



Two parts: a **generator** and a **discriminator**.

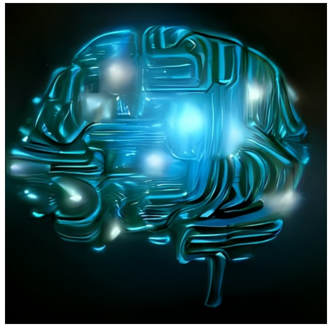
- The **generator** creates *fake data*, and the **discriminator** tries to tell real from fake.
- They compete in training, improving the generator's ability to create realistic data, like images or text.
- Movement along specific directions in the noise/random/input vector space can predictably alter features of the generated output, facilitating control over desired attributes.

Large Language Models

General Artificial Intelligence

LLMs

Narrow Artificial Intelligence



Main Role:

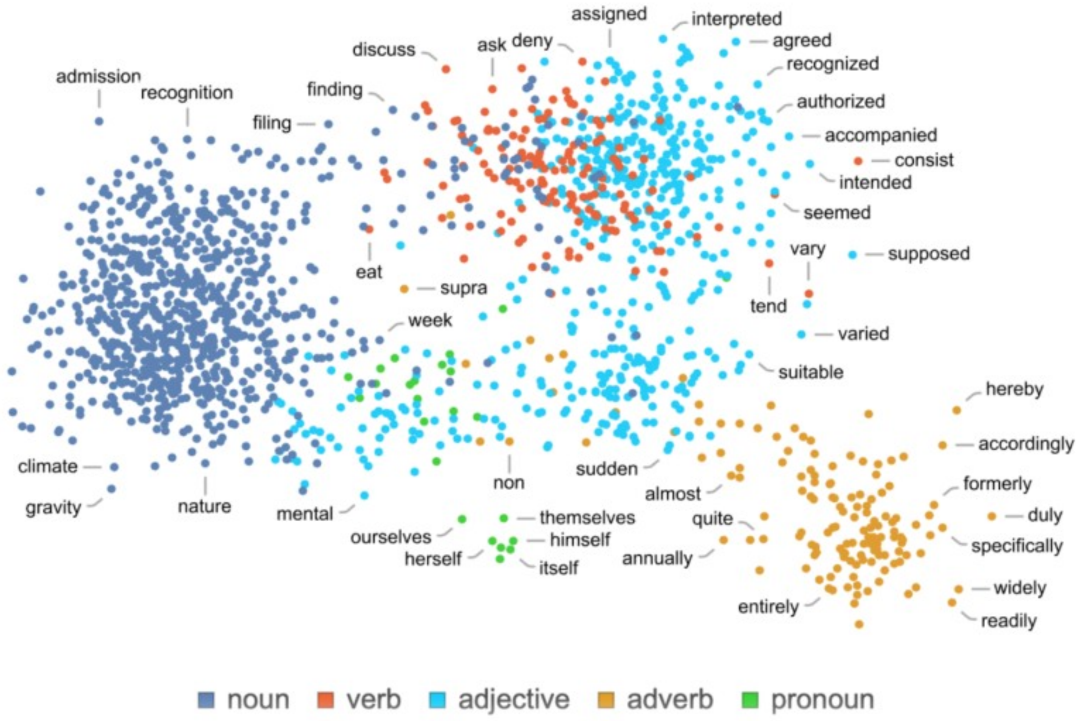
Machine language manipulation and generation of sentences in different (human) languages

BERT-style: complete missing words from a sentence

GPT-style: Add next word

Ingredients for LLMs

- Huge text corpora (mainly from the web) *produced by humans*
- Vectorial encoding of semantic proximity of text components
- Neural/sub-symbolic computational architecture
- Matrix operations on (encoding of) linguistic data
- GPUs used for parallel computing



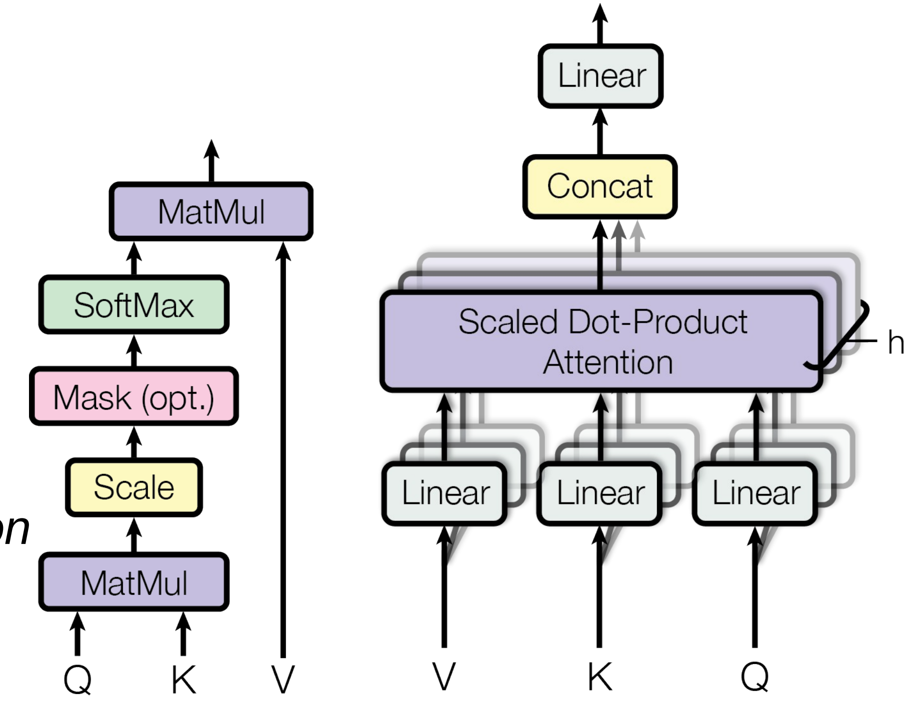
Recipe for LLMs

These are statistical encodings of *other words* that are "called by" or "calling for" the given word

1. Consider a string of lingustic items in a text
2. Assign to each of them 3 matrices: *Queries; Keys; Values*
3. Pair Q and K matrices to compute probability weights
4. Generate average values using these probability weights

✓ Complete BERT- or GPT-style tasks[1], including *language translation*

Attention architecture



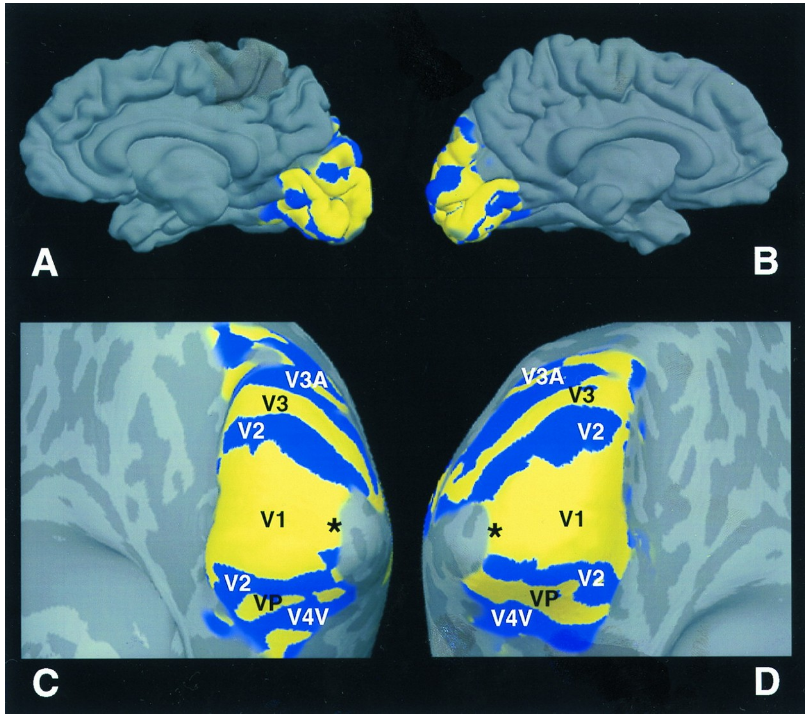
Q and K matrices are a *statistical proxy* for syntactic relationship

Syntax is re-constructed from its semantic mirror

Huge parallel computations (in GPUs) for searching syntactic structures within semantic encoding

[1] Modulo parameter refinements/training

Models for biological intelligence?
Convolutional Neural Networks



- Convolving an image with different kernels highlights different aspects of the image
- During end-to-end training, ***CNNs*** apply a set of kernels in order to exctract structural information that is relevant for classifying the input image
- The mechanism ***is inspired*** by the way our brain detects edges, textures, and colours, as reported by research in neurophysiology and brain imaging

Models for biological intelligence?
Large Language Models

- Multimodal LLMs are capable of non-trivial practical problem solving, along with well-known linguistic proficiency
- These performances depend on the availability of a *huge* amount of structured data and on the machine training (implicitly focused on mimicking Pólya’s contextual awareness in general terms)
- Language manipulation is thus reduced to a (non-trivial) problem solving task

However, we know from resarch in neurolingustics and biology that the human faculty of language is **not reducible** to general problem solving!

Models for biological intelligence?

Large Language Models

Mission: Impossible Language Models

Julie Kallini[1], Isabel Papadimitriou[1], Richard Futrell[2], Kyle Mahowald[3], Christopher Potts[1]

[1]Stanford University; [2]University of California, Irvine; [3]University of Texas, Austin

kallini@stanford.edu

Abstract

Chomsky and others have very directly claimed that large language models (LLMs) are equally capable of learning languages that are possible and impossible for humans to learn. However, there is very little published experimental evidence to support such a claim. Here, we develop a set of synthetic *impossible languages* of differing complexity, each designed by systematically altering English data with unnatural word orders and grammar rules. These languages lie on an impossibility continuum: at one end are languages that are inherently impossible, such as random and irreversible shuffles of English words, and on the other, languages that may not be intuitively impossible but are often considered so in linguistics, particularly those with rules based on counting word positions. We report on a wide range of evaluations to assess the capacity of GPT-2 small models to learn these uncontroversially impossible languages, and crucially, we perform these assessments at various stages throughout training to compare the learning process for each language. Our core finding is that GPT-2 struggles to learn impossible languages when compared to English as a control, challenging the core claim. More importantly, we hope our approach opens up a productive line of inquiry in which different LLM architectures are tested on a variety of impossible languages in an effort to learn more about how LLMs can be used as tools for these cognitive and typological investigations.

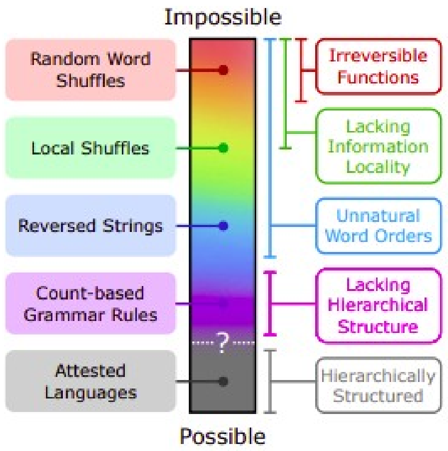


Figure 1: Partial impossibility continuum of languages based on complexity. We assess the learnability of languages at different points in the continuum and push the (currently unclear) boundary between possible and impossible.

viability of LLMs as the basis for robust language capabilities.

These authors state this claim in absolute terms. For example, Chomsky et al. (2023) flatly assert that LLMs "are incapable of distinguishing the possible from the impossible," Chomsky (2023) says this property "can't be modified," and Moro et al.

CORRESPONDENCE | 19 March 2024

Three reasons why AI doesn't model human language

By Johan J. Bolhuis, Stephen Crain, Sandiway Fong & Andrea Moro

Artificial intelligence (AI) is being used to develop large language models (LLMs) with considerable success. But they should not be seen as being models of how human language works and is acquired.

First, LLMs are probabilistic models of externalized language data, whereas human language is truly generative: it yields an unbounded number of hierarchically structured expressions (M. B. A. Everaert *et al. Trends Cogn. Sci.* **19**, 729–743; 2015). Second, language acquisition in infants does not depend on massive amounts of input data, but includes knowledge of language's generative nature. Therefore, children can acquire any language rapidly with minimal linguistic input (C. Yang *et al. Neurosci. Biobehav. Rev.* **81**, 103–119; 2017). Third, LLMs can produce 'impossible' languages, not generated by the principles governing all known human languages, just as well as (if not better than) natural language output, and cannot distinguish between them (A. Moro *et al. Cortex* **167**, 82–85; 2023).

LLMs are a useful technology, good at reproducing strands of information derived from training data, but they do not model the structure of human language, let alone encompass its creative nature.

Nature **627**, 489 (2024)

doi: https://doi.org/10.1038/d41586-024-00824-z

Models for biological intelligence?
Impressive achievements...

N.B. These performances in solving advanced maths problems are not based on the architectures of LLMs, but on an **interaction between deep learning *and* symbolic reasoning** (GOFAI)

nature

Explore content About the journal Publish with us Subscribe

nature > news > article

NEWS | 25 July 2024

DeepMind hits milestone in solving maths problems — AI's next grand challenge

AlphaProof showed its prowess on questions from this year's Mathematical Olympiad — a step in the race to create substantial proofs with artificial intelligence.

By Davide Castelvecchi

Models for biological intelligence?

… and trivial pitfalls

You

Solve this puzzle. A man and his goat are trying to cross a river. They have a boat. How do they do it?

Copilot

The man and the goat can cross the river using the following steps:

1. The man takes the goat across the river and leaves it on the other side.
2. The man returns alone to the original side.
3. The man then takes the cabbage across the river.
4. He leaves the cabbage on the other side and takes the goat back with him.
5. Finally, the man goes across the river one last time, leaving the goat and the cabbage on the opposite side.

This way, all three (man, goat, and cabbage) successfully cross the river without any issues! 🚣🐐🌿

2 of 4

You

There is no cabbage in the puzzle.

Copilot

My apologies for the oversight! Let's solve the puzzle correctly:

1. The man takes the **goat** across the river and leaves it on the other side.
2. The man returns alone to the original side.
3. The man then takes the **wolf** across the river.
4. He leaves the wolf on the other side and takes the goat back with him.
5. Finally, the man goes across the river one last time, leaving the goat and the wolf on the opposite side.

This way, all three (man, goat, and wolf) successfully cross the river without any issues! 🚣🐐🐺

3 of 4

Statistical encodings of problems/data/patterns do not suffice for genuine human reasoning (so far)

Summary

- Current technologies in AI have made significant progress in addressing Dreyfus' challenges originally proposed in the 1970s

- By implicitly tackling philosophical problems, AI systems have improved in contextual understanding and, in particular, structured data manipulation and language use.

  - **CNNs:** Inspired by neurophysiology, these models excel in image classification by highlighting key structural information through kernels.

  - **GANs:** Comprising a generator and discriminator, they improve through competition of the two components to generate new realistic but fake data (e.g., images or text) by identifying salient aspects of the training data.

  - **LLMs:** Focus on language manipulation and sentence generation. They are built on massive textual datasets and statistical methods to replicate syntactic relationships, but this language processing is still far from replicating genuine human reasoning.

Some tentative conclusions

- While AI tools like LLMs showcase impressive performance in machine translation, pattern recognition, and problem-solving, their reliance on statistical encodings and extensive data manipulation falls short of replicating the biological aspects of human reasoning and linguistic abilities *beyond purely behaviourist considerations*.

- Contemporary AI *based on pure connectionism* does not faithfully model most aspects of human cognition, especially the biological faculty of language.

- Even though Dreyfus's objections to AI seem to have been addressed by current technologies, a **new challenge for artificial models of *genuinely* natural intelligence** has emerged at the forefront of AI, rooted in evidence from the natural sciences (biology, neurology, brain imaging) and linguistics.

Many thanks for listening!

stefania.centrone@tum.de
cosimo.perinibrogi@imtlucca.it