

Language Is the Manifestation of Thought – Why Language Models Can Speak Without Thinking

“Artificial intelligence poses a double challenge to the human mind. Not only will the most advanced models soon be able to think better than people, but AI has consequences for humanity which are so uncertain, so potentially vast and approaching at such a rapid pace that even the best brains flinch”

The Economist, July 25, 2026

“Where are they, the great next philosophers, the equivalents of Kant or Wittgenstein or even Aristotle? I think we’re going to need that to help navigate society to that next step, because I think A.G.I. and artificial superintelligence are going to change humanity and the human condition.”

Demis Hassabis, DeepMind co-founder, Big Technology Podcast, January 23, 2025

Anyone who wants to know whether artificial intelligence thinks must first know what thinking actually is. That is precisely what is currently missing—in media reporting, in large parts of science, and among the makers of language models themselves. AI is credited with intelligence, understanding, learning, decision-making, or even empathy without first clarifying what these terms actually designate. A system that appears linguistically assured is credited with understanding; one that solves problems is deemed intelligent; and one that says “I” already seems to possess consciousness. Instead of determining the matter conceptually, people are dazzled by the impressive performance and infer an equally impressive machine mind.

Hegel is particularly fruitful for criticizing these misjudgments because he does not define intelligence as a measurable capacity for performance, but as the process by which a subject makes the world its own content through intuition, representation, language, and conceptual thinking.¹ We do not even need to turn Hegel entirely from his head onto his feet; it is enough to strip away his idealist metaphysics: it is not mind that rediscovers itself in the world, but a bodily, social, language-capable being that forms concepts in order to know and change a world that exists independently of it.

In what follows, the argument at first seems to lead away from AI: from a few very basic determinations of language, intelligence, and thinking to bees and primates, infants and children raised in linguistic isolation, deaf people and sign languages, and on to the Inuit, Homer, and the question of whether different languages favor different forms of thinking. Comparing these different forms of cognitive and linguistic ability reveals what perception, recollection, communication, and the use of signs can already accomplish—and where the transition to language and conceptual thinking lies. Against this background, an analysis of what a Large Language Model is and does shows that it condenses the linguistic results of human experience of the world into statistical

¹ Especially in the *Encyclopaedia of the Philosophical Sciences* (1817).

relations devoid of concepts. It can speak with a deceptive appearance of understanding, but possesses language only as a calculable form: it simulates sensation, cognition, understanding, and decision-making by statistically extrapolating new utterances from the linguistic sediment of human perceptions, actions, and experiences—without ever having traversed for itself the path from the world to the concept.

This also answers the question of AI's dangerousness. The danger, however, does not lie in an LLM that one day develops consciousness, sets purposes of its own, and rises up against human beings. As a mere language model, no amount of training or scaling can make it produce consciousness: it remains a system of probabilistic sign-processing without a relation to the world of its own, a standpoint of its own, or self-determined purposes. It can nevertheless become exceedingly dangerous through human use—as an obedient instrument of state, military, industrial, and criminal interests. The problem is not a revolt of the machines, but the speed, reach, and efficiency with which they can serve existing apparatuses of power, violence, and control and carry out their purposes.

This argument stands in clear opposition to widespread positions in philosophy, psychology, cognitive science, and AI research and therefore invites fundamental objections. At the end of the text, the most important objections will first be presented in their strongest form and then answered.

What Is Consciousness?

Consciousness designates a form of relation to the world in which contents of the environment are not merely registered, discriminated, and processed in ways that affect behavior, but **appear to a sentient subject** as objects distinct from it. Consciousness therefore presupposes sensation but goes beyond it: not merely a state is experienced; something is perceived as an object. On current evidence, such a form of relation to the world has been demonstrated only in living beings.

A system's mere ability to take in environmental information, distinguish it, and process it in behaviorally effective ways is not sufficient. Sensory registration is not sensation, and object recognition is not object-directed consciousness. A technical system, too, can process stimuli and classify objects without this showing that anything at all appears to it as an object.²

A conscious organism therefore does not merely experience isolated stimuli or internal states; it sees a tree, for example, tracks prey, or recognizes a member of its own species. Consciousness begins where something appears to a subject as an object distinct from it. It therefore **presupposes neither language nor conceptual thinking**; animals, too, can inhabit an environment structured into objects, recognize individual objects, and orient their behavior toward them.³

² A camera is causally altered by light; an image classifier can then distinguish objects. Both demonstrate information processing, but neither yet demonstrates subjective experience or an object-directed relation to the world.

³ An animal can, for example, experience fear, recognize what occasioned it, and respond accordingly. A human being can additionally know and say that they are afraid, ask about the cause of the fear, and judge whether it is appropriate. What is specifically human, therefore, is not consciousness as such, but its linguistically mediated and conceptually reflected form.

A further level is reached when the subject is not only related to objects but at the same time apprehends itself as the one that is related to them. Consciousness of an object thereby becomes at the same time consciousness of one's own relation to that object: **self-consciousness**. Whereas in consciousness a world stands over against the subject, the self-conscious subject at the same time knows itself as the one to whom that world stands opposed.

Consciousness is grounded in sensation but is not identical with it. Sensation occurs when a stimulus is not merely processed but is connected with subjective experience. Consciousness goes beyond this: what is experienced appears to the subject as something distinct from itself.

From Perception to the Concept

Organisms—from the ant to the mouse to the human infant—initially gain access to their environment through the senses. **Perception** is based on the reception and neural processing of external stimuli, such as visual, auditory, olfactory, or tactile signals. This is not a mere mapping of the external world, however, but a selective, organism-specific processing of sensory information. In this fundamental respect, widely different animal species and human beings have something in common. Technical systems (computers), too, can register and process stimuli; whether this should already count as perception depends on whether the term is understood merely functionally or also in the sense of subjective experience.

Perception comprises sensation and intuition: sensations supply the qualitative contents; intuition apprehends them as determinations of present objects.

At the beginning of the intellectual appropriation of the world stands sensory and affective **sensation**.⁴ Organisms respond to colors, sounds, touch, warmth, pain, pleasure, and displeasure. In developed perception, however, such sensations usually do not occur in isolation. Sensation constitutes the qualitative content of perception.⁵

Whereas in sensation no object as such has yet been posited, in **intuition** the content of sensation is apprehended spatially and temporally as a determination of a present object. The nervous system apprehends it immediately and in its concrete sensory appearance. Intuition is therefore directed toward the individual in the here and now and constitutes a first form in which an organism gains access to its environment. In infants and most animal species, intuition (like sensation) can become directly effective in action. It may then trigger reflex-like, instinctive, learned, or situation-bound responses without the perceived object already being grasped as a general concept and explicitly reflected upon. Thus, a bee's perception of a blossom can trigger complex searching, orienting, and

⁴ In what follows, Hegel's functional determinations are not presented in their systematic order, but rearranged as follows: sensation → intuition → representation → thinking.

⁵ External sensations such as colors, sounds, or smells are ordinarily not perceived in isolation and temporally prior to the objects to which they belong. We do not first see "red" and then assemble an apple out of it; we immediately perceive a red apple. The distinction between sensation and intuition therefore does not designate a temporal sequence, but different moments of the same perceptual process. Sensation is its qualitative content—red, warm, or loud, for example—whereas intuition apprehends this content spatially and temporally as a determination of a present object. Only in conceptual reflection can the sensed red be distinguished from the intuited red object. In Hegel's own words: "Sensation is the form of the mind's dull weaving in its individuality, devoid of consciousness and understanding." (*Encyclopaedia*, § 400).

foraging behavior without the bee conceptually determining the blossom, in the human sense, as a “blossom,” a “plant,” or a “food source.”

In **representation**, the contents gained in intuition are released from their immediate tie to the here and now. A previously perceived object can be recalled and made inwardly present even when it is no longer present here and now. This capacity to reactivate earlier perceptual contents can be called „**reproductive imagination**“. In this way, for example, one can form a representation of a tree seen at an earlier time.

Representations can also be linked with one another on the basis of shared features, recurring connections, or individual experience (“**associative imagination**”⁶). Such associative relations may, for example, connect the representation of a tree with that of an apple. Unlike the concept, representation remains bound to concrete experiences and subjective perspectives. It is often intuitive or image-like, but need not necessarily occur in the form of an inner image.⁷ It also remains comparatively indeterminate because it does not yet explicitly order its contents by general features and fix them conceptually.

Animals with more limited cognitive capacities, too, possess mental representations in which earlier experiences are stored and from which expectations about future events arise; in this broad sense these could be called representations. In computers, by contrast, the term would be appropriate only in a functional and figurative sense: they can retrieve stored representations and link them with one another without this corresponding to subjective experience or an inner representation.

In **recollection**, the subject releases the content of intuition from its immediate tie to spatial and temporal presence. What was intuited becomes a representation and thus an internalized content of the subject, which the subject can reproduce even in the renewed absence of the object.

Sensation, perception, recollection, and representation are not yet intelligence in the narrower sense. Rather, they form its cognitive prerequisites. Animals, too, can perceive, remember what they have experienced, recognize objects, and form expectations without thereby appropriating the world conceptually. The decisive transition occurs only where representations are fixed by generally usable signs, mediated linguistically with one another, and ultimately determined as concepts.

What Is Intelligence?

Intelligence designates the activity through which a language-capable and self-conscious subject conceptually appropriates the world that is initially given to it as external. It presupposes

⁶ The occasion for an association is an intuition or representation. Which features come to the fore and which further representations are thereby called up initially depend on the subject’s attention, experience, and arbitrariness. This is the deficiency of associative thinking: the connection does not necessarily arise from the essential content of the thing, but often from contingent, external, or merely subjectively emphasized similarities. Associative imagination nevertheless forms an important preliminary stage of conceptual thinking. By noticing relations among representations, the subject discovers identities, similarities, and recurring features. In this way, general representations arise that point beyond the individual object. They become durably fixed and generally available by being given a name (see the discussion of memory and language below).

⁷ Human representations need not be pictorial. People with aphantasia can remember, process, and conceptually think through contents without experiencing them as conscious inner images. They thus vividly demonstrate that thinking does not depend on calling up visual representations.

perception, representation, and recollection, but goes beyond them: the individual object is no longer merely perceived or made present in representation, but is brought under general determinations through language, related to other objects and concepts, and finally known in accordance with its concept.

Memory is an essential prerequisite here for linguistic signs to be stably connected with particular meanings and used repeatedly. In this respect, intelligence operates as a sign-forming activity: it connects representations with sounds, gestures, or written signs and thereby makes them durably available and communicable.

As such signs are retained in memory, recognized, and used according to shared rules, **language** emerges as a supra-individual system. It confronts the individual speaker as a pre-existing order of signs, meanings, and grammatical structures that must be learned in social use.

Intelligence is the intellectual appropriation of a content by a subject that knows itself. It is the activity through which a subject takes up a content that initially appears to be externally given, internalizes it, and conceptually makes it its own. This occurs through intuition, representation, and thinking. The object is apprehended as an individual, present content (“cognized”); the content is released from the immediate here and now, recalled, and made present in representation (“represented”); and it is fixed by signs and names (“language”). The object can then be known in thought no longer merely through intuition or representation, but in its universality and according to its concept.⁸

The rational content of **Hegel’s concept of intelligence**⁹ remains intact once it is stripped of its idealism. The subject (in Hegel, “the theoretical mind”) does not merely register the world, but appropriates it by internalizing perceptions, ordering them, and determining them conceptually. In the concept, however, human beings do not, as Hegel argues, rediscover their own reason as the hidden structure of reality. Concepts are not forms of thought dwelling in things themselves, but naturally, historically, and socially developed forms of intellectual appropriation by which human beings apprehend a material world that exists independently of them. Hegel’s concept of intelligence must therefore be freed from his idealist metaphysics: it is not mind that recognizes itself in the world, but a material, social, language-capable being that forms concepts in order to orient itself in the world, understand it, and change it.¹⁰

⁸ In ordinary usage and comparative cognition research, cognitively highly developed animals are also described as intelligent. What is usually meant are capacities such as learning, problem-solving, flexible behavioral adaptation, memory, planning, or tool use. Whether animals count as intelligent therefore depends substantially on the concept of intelligence being used. If intelligence is understood functionally as a capacity for successful adaptation and problem-solving, many animal species are undoubtedly intelligent. If, however, intelligence is understood—as it is here—as the intellectual appropriation of the world through linguistically mediated conceptual thinking by a subject that knows itself, then animals possess numerous prerequisites of intelligence but would not, in this sense, be described as intelligent. This does not deny their complex cognitive achievements; it merely distinguishes them from conceptual intelligence.

⁹ G. W. F. Hegel, *Enzyklopädie der philosophischen Wissenschaften im Grundrisse III. Die Philosophie des Geistes*, in *Werke in zwanzig Bänden*, vol. 10, Frankfurt am Main, 1970, especially pp. 229–287 and §§ 440–468.

¹⁰ For Hegel, the concept belongs not only to the cognizing subject but also to the objective structure of reality. When mind conceptually apprehends the world, then, according to Hegel, it does not merely impose its own categories on it. It recognizes in the world a rationality that is at the same time the structure of its own thinking. This is why

What Is Language?

In **memory**, linguistic signs become durably associated with representations. A thing receives a name—“tree” or “state,” for example—and thereby becomes linguistically determinable. Someone who hears the word “tree” does not ordinarily direct attention to the sequence of sounds or letters T-R-E-E as such; rather, the designated object or its conceptual content is immediately made present. The word thus functions as the bearer of the concept. In it, the thing is released from its contingent individual sensory form and preserved in an ideal form available to thought. Word and object are not the same in reality; in the act of thinking, however, sign and meaning coincide so closely that the object becomes intellectually present in the name.

Language is therefore not merely an external means by which thoughts that already exist independently of it are subsequently expressed. Rather, it is a constitutive activity of conceptual thinking: the word constitutes the objective **existence of thought**, through which thought can be determined, retained, and communicated. Language is in this sense the necessary form of appearance in which conceptual thinking acquires reality and intersubjective validity.¹¹

The assumption of a fully developed inner “language of mind”—a *Mentalese*—whose contents are subsequently merely translated into a public language is untenable (for the reasons discussed below). Language emerges and takes shape in the use of linguistic signs within social practices. Meanings are not formed in isolated consciousness, but in shared activities of perceiving, acting, naming, distinguishing, teaching, arguing, and correcting. On this basis, concepts develop as generalized and reflected determinations of meaning. Concepts are formed linguistically and socially, but they do not acquire their objective validity from social agreement. They acquire it insofar as they correctly apprehend the necessary and essential determinations of their object. Social practice mediates cognition of the thing; it does not produce the thing’s objective determinations.

Human beings relate to the world cognitively by determining it in **thought**. **Genuine understanding** abstracts from the initially disordered-seeming manifold of nature and social reality. It asks not only how a thing immediately presents itself, but what it is in its essence: which determinations are necessary and essential to it and which remain merely contingent, external, or

Hegel says that, in cognizing the object, „mind is ultimately at home with itself“. This is idealist metaphysics. Science approaches its purpose of correctly apprehending and explaining natural and social phenomena not because mind recognizes itself in the world. It does so because false theories—at least in the natural sciences—eventually fail because of internal contradictions or deficiencies in their explanation of their object. That a genuine understanding of the subject matters is nevertheless by no means guaranteed becomes especially clear in the humanities and other “soft” sciences, including parts of medicine: Genuine understanding fails here not only because of substantive complexity, methodological limits, or insufficient evidence, but often also because of extraneous interests, methodological errors, biases, and ideological presuppositions on the part of researchers.

¹¹ Language, then, is by no means merely a means of thinking. Karl Kraus put it this way in *Die Fackel* (288) in 1909: “Language is the mother, not the maidservant, of thought.” Even then, Kraus was opposing the common idea that a thought first exists in finished form and is subsequently merely “put into words.” Rather, the thought proper comes into being only by finding its linguistic form.

changeable. To grasp the **concept**¹² of a thing therefore means to know its essential determinations and their inner connection.¹³

What Is Thinking?

In thinking, we generalize individual experiences and what we encounter in the world into concepts, make our own thinking an object of reflection, and judge statements, actions, or reasons. Conceptual thinking is therefore not only abstractive but also self-reflective, and it distinguishes between true and false, appropriate and inappropriate, justified and unjustified.¹⁴

Methodologically, conceptual thinking proceeds essentially in the forms of **judgment** and **inference**. In a judgment, an object is posited as subject and further determined by a predicate: “The rose is red” or “The state has a military.” Subject and predicate are initially distinguished from one another and are then connected by the copula “is” or “has.”

A judgment states **that** a particular property belongs to an object. Taken by itself, however, it does not yet provide a sufficient justification for **why** this connection holds. In that respect, it initially remains an assertion. Only in an inference is the relation between subject and predicate justified by a mediating determination. The inference shows on the basis of what general connection the relevant predicate belongs to the particular or individual object. It thus mediates the universal, the particular, and the individual.

A simplified example is this: “This apple is sweet” is initially a judgment. Its justification can take the form of an inference: ripe apples typically contain sugar and therefore taste sweet; this apple is ripe; consequently, it is to be expected that it will taste sweet. The mediating determination here is the ripeness of the apple and the sugar content associated with it.

A physical connection can be derived in the same way: “This piece of metal expands when heated” is initially a judgment. Its justification lies in the fact that the thermal motion of atoms increases as temperature rises and that the bonding relations in a solid are generally asymmetric. This increases

¹² Note the metaphor in the German *begreifen* (“to comprehend”): literally, to get hold of something with the hand.

¹³ The idealist Hegel argues that thinking is objective and that the concept establishes identity with the object (*Encyclopaedia*, §§ 213, 468). This does not mean merely that a subjective thought adapts itself as closely as possible to an external object, so that one can grasp in thought what the object really is. For Hegel, the concept designates the inner, necessary structure of the thing itself—its essence and the logic of its development. True thinking therefore does not impose its determinations on the object from outside, but follows the object’s own movement. Because, for Hegel, actuality is itself rational or conceptually structured, reason ultimately recognizes its own determinations in the object. In this sense, actuality becomes conscious of itself in cognition. Hegel’s famous sentence from the Preface to the *Philosophy of Right* points in the same direction: “What is rational is actual; and what is actual is rational.” For Hegel, actuality is therefore not conceptless material external to mind. It itself possesses a rational structure that is not first produced in thinking but conceptually grasped there. In science, reason therefore rediscovers the conceptual order that, according to Hegel, is already at work in the world.

¹⁴ All of this, of course, often fails—including, and especially, in science. From deception through error to the supreme discipline of ideology: only for the idealist Hegel does having “the thing in the name” already amount to having the concept. The all-too-common tendency to explain phenomena simply by citing the conditions under which they occur belongs here as well: When all the conditions are present, what is to be explained merely exists. But one still does not know what it is merely by knowing its conditions—or, worse, by making those conditions its explanation. The same applies to explaining a phenomenon by its “faculty” or “inner force”: a tautology, a duplication into manifestation (the thing) and force (for example, “wars arise from aggression”).

the mean distance between the atoms. Since the object in question consists of such a metal lattice, its volume increases when heated.

In inference, the connection that was initially only asserted in the judgment is thus made intelligible through a mediating ground. Conceptual thinking does not stop at establishing that something is so; it seeks to show the necessary, or at least regular, connection from which this determination follows.¹⁵

When explaining a thing—and thus determining its concept—we ensure that the ground logically precedes what is grounded. An explanation shows on the basis of what presuppositions, conditions, or connections something is as it is: because a particular ground obtains, a concrete determination follows from it. This is the essence of **derivation**. It does not merely state a fact, but shows how what is grounded follows from its ground. A conceptual derivation thus makes intelligible the inner connection between the ground and what follows from it.

Do Animals Think?

If thinking is understood—as it is here—as linguistically mediated conceptual thinking, then thinking in this narrower sense cannot be attributed to animals. This does not in any way deny their diverse cognitive and affective capacities. Many animals perceive in highly differentiated ways, remember past events, form expectations, learn from experience, and find themselves in moods or emotional states. These preconceptual forms of relation to the world and to self are precisely the appropriate starting point for asking what constitutes the specific transition to thinking.

Moods are diffuse, as yet unreflected forms of how a living being finds itself. A living being is in a certain affective condition without explicitly distinguishing that condition from itself as its own state or determining it conceptually. **Feelings**¹⁶, too, are initially immediate and subjective: they are experienced but not necessarily cognized, put into language, or understood in their causes. In this sense, they constitute an elementary form of self-relation that can also be assumed in many animals.

Conceptual thinking begins where a subject can make its own state an object. It may, for example, feel anger, then recognize and name it: “I am angry,” and finally investigate why the anger arose and whether it is appropriate to its occasion. The feeling is not thereby eliminated but **sublated**: it persists as an experienced state, while at the same time being held at a distance, determined conceptually, and situated within a broader context.

¹⁵ The exposition and examples followed the classical form of syllogistic inference, in which a predicate is ascribed to an individual through a universal and a particular determination. The basic idea, however, also applies to other forms of inference: deductive inferences justify a statement through its necessary derivation from premises, inductive inferences through the generalization of observed individual cases, and probabilistic inferences through determining its degree of probability on the basis of available evidence. Their degree of validity differs: whereas a valid deductive inference with true premises yields a necessary conclusion, inductive and probabilistic inferences can support their results only more or less strongly.

¹⁶ Hegel acknowledges that ordinary usage does not consistently distinguish sensation from feeling. He nevertheless suggests a difference (*Encyclopaedia*, § 402): sensation places greater emphasis on the passive finding of an immediate determinacy; feeling places greater emphasis on the fact that the determinacy is related to the self and experienced as its own state. Thus one is more likely to speak of a “sensation of pain,” but of a “feeling of fear,” “feeling of shame,” or “feeling of grief.” In feeling, the subject experiences itself more comprehensively as being in a particular condition; even this form, however, remains below the level of conceptual thinking.

Animals, too, can respond to their environment flexibly, learn from experience, and act purposefully. Their behavior rests on innate drives, affective dispositions, perceptions, recollections, and learned patterns of behavior. It is not, however, determined by linguistically mediated reflection on their own motives, by explicitly formulated reasons, or by general principles of action. Animals experience affective states and act out of them, but they do not in the same way make these states the object of a conceptual relation to themselves.

This at the same time marks a limit to their **freedom**. Freedom does not consist in acting without natural drives. Human beings, too, remain needy, sentient, drive-governed living beings. Rather, the will becomes **free** insofar as a person can make immediate needs, inclinations, and purposes objects of thought, judge them, and consciously take a position toward them.¹⁷ A person can affirm them, defer them, weigh them against one another, or orient action toward general reasons and self-set purposes.

Thinking realizes this freedom in will and action. So long as a thought remains a mere representation, possibility, or theoretical insight, it does not yet possess objective reality. Only in conscious decision and action does the subject objectify itself and change the world in accordance with a represented purpose. An animal may feel hunger, seek food, and learn how to find it. A human being, by contrast, can think “food” as a general concept, reflect on the production and distribution of food, or consciously decide to fast for religious, political, or health reasons.

The boundary between animal and human cognition by no means coincides with the boundary between **communication** and noncommunication. Animals exchange information even without language—for example through posture, odors, colors, gestures, or sounds. A well-known example is the honeybee waggle dance, by which information about the direction and distance of a food source is conveyed. Such a sign system can be astonishingly complex without thereby already constituting a freely combinable language.

In some animal species, communication is more flexible and more strongly learned. Sperm whales use characteristic sound sequences by which social groups, and possibly individual animals as well, can be distinguished. The songs of many bird species are based partly on innate dispositions and partly on learning and imitation. Parrots can reproduce human sound sequences and use individual words appropriately in particular situations. Such achievements show that animal communication cannot be reduced to rigid reflexes. Yet their signs generally remain bound to limited situations, needs, and contexts of action.

The gestures of nonhuman primates are especially versatile. Chimpanzees can learn artificial sign systems, retain numerous signs, and in some cases form new combinations. They can refer to objects, places, persons, and actions and express simple wants or assignments. This does not yet

¹⁷ For Hegel, genuine freedom consists in willing what is rational. The free will is the will that does not allow itself to be determined merely by external circumstances or immediate inclinations, but rationally determines itself and realizes this self-determination in a rational social order. Freedom therefore acquires its objective form for Hegel in right, morality, and finally in the institutions of ethical life—especially the family, civil society, and the state. He does not regard these as mere restrictions on individual freedom, but as the forms in which the rational will first acquires objective and enduring actuality. In this respect, Hegel is both a theorist and an apologist of the modern state, whose institutional form was only then in the historical process of being established.

show, however, that they use signs, as humans do, as elements of a grammatically structured system of general concepts, or make the meaning of those signs itself an object of reflection.

Nor is the capacity for categorization unique to human beings. Cephalopods, corvids, dolphins, and primates can treat different objects as instances of recurring classes, distinguish edible from inedible objects, select suitable tools, and transfer learned distinctions to new situations. Some animals can prepare actions, anticipate future events, recognize themselves in a mirror, or, to a limited extent, anticipate the behavior of others from their perspective.¹⁸

Such achievements show that perception, memory, learning, categorization, planning, and social orientation can already be highly developed prior to language. They must not, however, be equated with conceptual thinking. The decisive transition does not consist simply in grasping similarities or responding to signs. It consists in fixing a general determination by means of a publicly shared linguistic sign, detaching it from the individual situation, systematically connecting it with other concepts, and explicitly reflecting on its meaning. Animals can use categories practically; human beings can think them as concepts.

Language Deprivation and Sign Language

Reports about people who grew up largely isolated from others in early childhood—so-called “feral children” or “**Kaspar Hauser cases**”—are also illuminating for the significance of language to thinking. If they were unable to acquire a language during the first years of life,¹⁹ they later found it extremely difficult to learn one despite intensive support. In most cases, their communication remained limited to simple signs, gestures, and a small number of words. In many respects, their linguistic and intellectual abilities thus resembled those of primates that had learned sign systems.

For completeness, one should also mention people who are deaf from birth and acquire a **sign language** early in life. Sign languages are fully fledged natural languages. They can express meanings, grammatical structures, and abstract thoughts just as complex as those expressed by spoken languages.²⁰

¹⁸ “Theory of mind.”

¹⁹ Babies usually begin babbling at around six to eight months. At roughly ten to twelve months they use their first words with recognizable meaning. Between the eighteenth and twentieth month there is often a marked developmental leap: vocabulary grows rapidly and children begin combining two words. By around four years of age, most children have already mastered the basic grammatical structures of their language.

²⁰ The autonomy and diversity of sign languages is evident from the fact that American Sign Language and British Sign Language are different languages. Deaf people from the United States and the United Kingdom therefore cannot simply communicate with one another in their respective sign languages.

For thinking, then, it is in principle irrelevant whether linguistic signs are written,²¹ signed, or spoken. What matters is that they carry meaning. In thinking, the linguistic sign no longer stands merely for an individual object, but for its general concept. The thing is thereby realized in thought.

Neural Representation and Language

The nervous system takes in the bodily experienced world through different senses, such as sight, hearing, and touch. These experiences leave enduring neural traces in the brain, which can be stored through complex synaptic interconnections that remain largely poorly understood. No symbolic representation is required for this.

The sight of a tree (intuition) leaves a particular pattern of neural activity in the brain (“**embodiment**”). If this pattern is later reactivated, the tree is inwardly represented and thereby remembered. Neural representations emulate sensory experiences.

Some animals can link the neural patterns of their respective representations with specific gestures, sounds, and in some cases even simple symbols (see above). In human beings, however, this capacity is much more highly developed: they can connect perceptions and experiences with complex systems of signs and words. This allows them to order and generalize them conceptually and make them available for abstract thinking.²²

²¹ Hegel grants spoken language systematic priority. For him, “phonic language” (*Tonsprache*) is the original form of language because human inwardness expresses itself immediately in the spoken sound. Written language, by contrast, is a subsequent spatial fixation: visible written signs refer to spoken sounds. “Intelligence expresses itself immediately and unconditionally through speaking” (*Encyclopaedia*, § 459). This does not, however, imply any disparagement of writing. Alphabetic writing in particular has, for Hegel, great intellectual and pedagogical value. Because its letters do not immediately depict things or representations, but instead analyze the sounds of spoken language, alphabetic writing makes the word itself an object of reflection. Reading and writing thereby promote abstraction, linguistic articulation, and conceptual formation. In comparison with hieroglyphic writing, Hegel therefore calls alphabetic writing “in and for itself the more intelligent.”

²² Attempts to decode certain contents of thought (“mind-reading”) by means of functional magnetic resonance imaging (fMRI) are based on this. fMRI measures changes in blood oxygenation that are indirectly associated with neural activity. This measurement method, however, captures brain activity only imprecisely in spatial terms and with a temporal delay. As a result, it usually permits only rough distinctions among contents or activities—for example, whether a person is thinking about a house or a vehicle, or imagining playing tennis. In addition, the system generally has to be trained beforehand on the individual activity patterns of the particular person. The results are therefore necessarily limited and often error-prone, but in a crude way they demonstrate “embodiment.” Other methods can also be used to measure brain activity. Near-infrared spectroscopy (NIRS), like fMRI, measures changes in blood oxygen concentration in active brain areas. The electroencephalogram (EEG), by contrast, directly records fluctuations in electrical voltage at the surface of the scalp. Because imagined actions are based on specific patterns of neural activity, NIRS and EEG have the potential, in so-called brain-computer interfaces, to control machines solely through deliberately generated brain activity. After extensive training, some people can, for example, move a cursor on a screen by imagining a movement. Under certain conditions, this allows people with high-level spinal cord injuries or locked-in syndrome to communicate or operate technical devices.

From the Tree to the State: The Formation of Abstract Concepts

Such representations initially refer above all to concrete objects in the perceptible world, such as a tree. Abstract concepts such as “state,” “society,” or “love,” by contrast, emerge only through further processing in thought.

By comparing different individual cases, human beings form general representations—for example, from the perception of individual trees, the representation “tree.” On this basis, **metaphors**²³ and other linguistic transfers make it possible to grasp abstract relations by means of bodily, spatial, and social experience. Thus we speak of a “high position,” “social ascent,” a “weighty decision,” an “argument under attack,” or a “broken relationship.” Abstract connections are thereby represented according to the pattern of sensuously experienced magnitudes, movements, weights, struggles, or objects.

Abstract concepts thus emerge step by step through the generalization and extension of concrete experience. A car may initially be perceived as an individual object or as a toy. Later, its function is understood: one can travel in it. It is then assigned to the general category “means of transport.” Beyond this, it can be regarded as property—“my car”—or as a social status symbol.

Linguistic Relativity: Do the Inuit Think Differently?

The close interweaving of language and thinking—that is, the connection among linguistic sign, concept, and designated object in the act of thinking—suggests that, within certain limits, the structures of individual languages can also influence cognitive processes. The thesis known as **linguistic relativity** does not claim that a language determines thought. It may, however, influence which distinctions are especially readily available, what attention is directed toward, and how experiences are categorized, remembered, or described. Empirical indications of such effects can be found, among other areas, in the linguistic classification of colors and in the mental representation of spatial and temporal relations. Basic human perceptual capacities, however, remain largely independent of them.

Frequently cited examples, such as the supposedly extraordinary number of Inuit terms for snow or Homer’s description of the sea as “wine-dark,” are only of limited use as evidence for linguistic relativity. The claim about countless “words for snow” is based in part on a misleading way of counting complex word forms in polysynthetic languages. Homeric color terms, meanwhile, may reflect different semantic organizations, but they do not justify the conclusion that people in ancient Greece physiologically perceived colors differently.

A particularly vivid example is provided instead by the technical vocabularies of scientific and technical disciplines. Their terminologies do not create the objects they designate, but they do provide precise and stably usable conceptual distinctions. They thereby allow phenomena to be apprehended in a differentiated way, compared with one another, and situated within theoretical

²³ *metá* = across, beyond; *phéreîn* = to carry. *metaphorá* therefore literally means “transfer” or “carrying across.”

contexts that would be difficult to formulate without the corresponding vocabulary. Technical language therefore expands what can be thought not in the sense of an absolute boundary, but in the precision, systematicity, and communicative availability of thought.

Ontogeny and Phylogeny of Conceptual Thinking

For the development of conceptual thinking, too, one can speak—at least in a figurative and not a strictly Haeckelian sense²⁴—of a certain correspondence between ontogeny and phylogeny. Individual development does not literally repeat the evolution of thinking. Yet it appears in a comparable way to rest on functional stages that build upon one another step by step: from perception and affective evaluation, through memory, representation, and the use of symbols, to language and conceptual thinking.

The **infant** begins life with perceptions, bodily sensations, and affects. At first, it recognizes individual objects and recurring contexts of action. From these experiences increasingly stable mental representations emerge: an object can be remembered and expected even when it is not currently perceptible (“representation,” see above). With the development of object permanence,²⁵ the infant can retain an object inwardly even when it is not currently perceptible. Representation thereby becomes partly detached from immediate intuition. Gestures, sounds, and other signs then make initial forms of symbolic reference possible. Through language acquisition, such representations are eventually connected with conventional words and grammatical structures. This provides the basis for not merely recognizing objects and dealing with them practically, but thinking them as general concepts, relating them to one another, and explicitly reflecting on them.

From a comparative **evolutionary** perspective, individual functional prerequisites of this development can be observed in various animal species. Honeybees perceive their environment in differentiated ways, learn from experience, remember spatial relations, and use the waggle dance to convey information about the direction and distance of food sources. Their behavior thus presupposes stable, action-relevant representations of spatial relations without requiring language. This does not, however, warrant the conclusion that they possess a conscious inner representation in the human sense.

Dolphins and nonhuman primates, moreover, display more flexible forms of learning, object categorization, action planning, social communication, and—to varying degrees depending on the species—tool use and the processing of causal relations. Such animals can treat individual objects as instances of recurring categories: they distinguish edible from inedible objects, for example, or select particular tools for particular purposes. This points to a form of categorical processing, but it is not equivalent to linguistically mediated human conceptual thinking.

²⁴ Ernst Haeckel, *Anthropogenie oder Entwicklungsgeschichte des Menschen. Gemeinverständliche wissenschaftliche Vorträge über die Grundzüge der menschlichen Keimes- und Stammes-Geschichte*, Leipzig, 1874, p. 4: “Ontogeny is the brief and rapid recapitulation of phylogeny.” Haeckel’s biogenetic law is not being transferred here to the development of thinking, but used only as an analogy for a structural correspondence between phylogenetically older cognitive achievements and their ontogenetic sequence.

²⁵ Simplifying somewhat, an infant without developed object permanence behaves as though a toy had disappeared once it is hidden under a blanket. As development proceeds, the infant begins to search for the concealed object.

The decisive transition, then, does not lie between cognizing a single object and recognizing similar objects. Animals, too, can categorize perceptions and assign new individual cases to known classes. A chimpanzee can, for example, treat different sticks as suitable tools or distinguish edible from inedible fruits. What is specifically human is the capacity to fix such a category, by means of a publicly shared linguistic sign, as a general concept detachable from the particular situation. A human being not only uses a stick purposefully, but can reflect on what a tool is in general, classify different tools by function, and make the concept of a tool itself an object of reflection.

In this restricted sense, the ontogeny of conceptual thinking bears a structural similarity to its phylogeny: the child begins with evolutionarily older forms of reflex, perception, affect regulation, recollection, and situation-bound categorization. On this basis it develops, through learning, symbolic communication and eventually linguistically mediated, abstract, and self-reflective thinking. The earlier forms are not overcome or replaced, but remain as supporting layers of human cognition. Ontogeny therefore does not repeat phylogenetic development; rather, in condensed form, it passes through some of its functional prerequisites.

Language Before Language?

The emergence and acquisition of linguistic and conceptual thinking in human beings cannot, however, be explained by assuming that an evolutionary mutation²⁶ produced in humans a minimal innate grammar on the basis of which language then unfolded, as it were, like an instinct. Versions of this view can be found in Chomsky and Pinker. Fodor and colleagues go further still with the assumption of an innate “language of mind”—so-called *Mentalese*—which is supposed to function as an internal system of representation and, in a sense, as the operating system of thought.

This explanation merely shifts the real problem onto a presupposed inner language. If *Mentalese* consists of symbols, one must explain what gives those symbols their meaning and how they refer to objects, properties, and processes in the experienced world. Without such grounding, they would remain purely formal signs.

A further question is how the symbols of an assumed inner language are connected with the words and grammatical structures of publicly used languages. If one posited another internal interpretive agency for this purpose, one would again have to explain how that agency understands the relevant mappings. An infinite regress threatens: every symbolic representation would require another agency to interpret it and endow it with meaning.

²⁶ According to the single-gene hypothesis, a single mutation or a small number of genetic changes in human evolution supposedly enabled the decisive transition to language—for example, by producing an innate grammar, improved vocal control, or new neural connections. FOXP2 attracted particular attention after mutations of this gene were identified in a family with severe speech and language disorders. Against the notion of an actual “language gene,” however, stands the fact that FOXP2 does not encode language, grammar, or concepts. It is a transcription factor that influences the activity of numerous other genes and is involved above all in the development and functioning of neural networks for motor learning and the temporal coordination of complex movement sequences. Damage to it impairs language, but proves no more that it generates language than damage to the larynx proves that grammar is located there. Moreover, FOXP2 occurs in similar form in numerous animal species, and Neanderthals, too, possessed the variants long regarded as specifically human.

Language should therefore not be understood as an already fully developed, innate symbolic system. What is innate are general neurobiological prerequisites and learning capacities that make language acquisition possible. The concrete language, with its signs, meanings, and grammatical regularities, is acquired by the individual only through its use in bodily, social, and communicative practices. The meaning of linguistic signs therefore does not arise in a prior inner language, but in the activity of their use.

From Gesture to Language, from Image to Writing

Around two million years ago, a new form of cultural adaptability began to emerge within the genus **Homo**. Its members systematically produced tools, opened up new food sources through gathering, scavenging, and hunting, spread into different landscapes and climatic zones, and in doing so depended heavily on cooperation and social coordination. This development cannot be reduced to a mere increase in brain mass. What mattered, rather, was the emergence of new forms of social cognition and cultural transmission. One possible transitional stage was **mimesis**: not yet language, but a generative, intentionally deployed system of gestures, bodily movements, facial expressions, and pantomimic representation. Actions, objects, and events could thereby be imitated, enacted, and jointly represented. Mimesis thus created a prelinguistic space of shared meaning—and with it an essential prerequisite for the later emergence of symbolic language.²⁷

Language could emerge from mimetic communication as recurring gestures and sounds became conventionalized, stabilized, and freely combinable. This marked a decisive transition: communication was no longer bound to what was immediately present. What was absent could be made present in representation, what was past could be recalled, what was future could be projected, and experience could be organized into narratives. With this capacity, a **mythic** culture emerged:²⁸ a linguistically mediated shared world of narratives, norms, and accounts of origins. Cave paintings probably did not stand at the beginning of this process, but constituted its material extension. As durable external representations, they preserved symbolic and mythic contents beyond immediate performance and individual memory and thus, over the long term, prepared the transition to external sign systems and eventually to writing. The path from gesture to language and from image to writing is therefore at the same time the path by which the human mind objectified its contents outside itself.

Implicit World Models: How Organisms Grasp Regularities

The nervous system's sensory engagement with its environment leads not only to the representation of concrete objects—trees, houses, or other things—but also to the implicit apprehension of those physical and chemical regularities that are relevant to orientation and survival. Organisms process,

²⁷ Children before actual language acquisition, people without a shared spoken or sign language, persons with acutely impaired language ability, and travelers in countries whose language they do not speak can still communicate surprisingly effectively by means of gestures, facial expressions, and pantomime.

²⁸ These arguments concerning the development of human language follow Merlin Donald's *A Mind So Rare* (2002) and *Origins of the Modern Mind* (1991).

for example, that bodies fall downward under certain conditions, that movements follow certain trajectories, and that events occur in temporal sequences from which causal expectations can be derived.

An insect can therefore respond appropriately to the movement of potential prey or predators without having to formulate the underlying laws of motion mathematically or solve differential equations. The repeated coupling of perception and action gives rise in this way to an implicit internal model of the environment. It remains largely unconscious and unreflected, but enables the organism to anticipate future developments and adapt its behavior accordingly. Without such elementary world models, goal-directed behavior in animals would scarcely be possible.²⁹

The Brain as Computer, the Computer as Brain

Over the past 2,000 years, human intelligence has repeatedly been explained by means of technologies characteristic of the age: first as divine mind in the body, then as the effect of circulating bodily humors. Later, the human being was conceived as a mechanical automaton, then as an electrochemical system, and finally as a kind of telegraph network. In the twentieth century, the metaphor of the brain as computer prevailed. Each of these conceptions transferred the most advanced technical model of its time onto the human mind without thereby adequately explaining it.

The rise of the computer over the past fifty years largely coincided with the rise of the cognitive sciences. The latter made things easy for themselves by trying to understand the brain on the model of an electronic calculating machine:³⁰ as processor, memory, programs, input and output routines, and data lines. Such analogies are vivid, but they are not conceptual explanations. They do not grasp what is necessary and essential about the brain; they transfer to it external features of a technical model. An image takes the place of explanation.³¹

²⁹ The necessary complexity of sensorimotor world models and of the neural activity patterns underlying them likely constituted a major selection pressure in the evolution of the nervous system. Trees (and mussels), which do not move, do not need brains. Higher cognitive achievements, however, also appear not to depend primarily on absolute brain size, but on internal organization, neural architecture, connectivity, and functional specialization. Comparative proportions make this clear: the cranial capacities of Neanderthals and modern humans overlap considerably, while the brains of some whale species can weigh more than nine kilograms and thus several times as much as the human brain. Absolute brain mass therefore allows no direct inference either to linguistic capacities or to general cognitive performance. Particularly remarkable in comparison with artificial intelligence is the energy efficiency of biological information processing. The human brain operates on average at a power of roughly 20 watts and thus consumes a few hundred kilocalories per day, approximately the energy content of a Mars bar. Most of this energy is not used for individual conscious acts of thought, but for maintaining baseline neural activity, synaptic signal transmission, and cellular homeostasis. Mental activities, no matter how demanding, do not measurably alter the brain's overall energy consumption.

³⁰ The idea was not entirely original even then: John Hobbes had already compared rational thinking with calculation in *Leviathan* (1651): "For reason, in this sense, is nothing but reckoning (that is, adding and subtracting) of the consequences of general names agreed upon for the marking and signifying of our thoughts".

³¹ An early and influential representative of this analogy was the mathematician John von Neumann. In *The Computer and the Brain* (1958), he summarily described the operation of the human nervous system as "prima facie digital"—even though he also had to concede that very little reliable knowledge existed about the role of the brain in thinking and remembering. Von Neumann thus belongs among the pioneers of an error of thought that remains widespread today: precisely the far-reaching ignorance of the neurobiological foundations of higher brain functions becomes a license for ever new analogies. The less one knows, the more uninhibited the comparisons become. Explanation gives way to an associative game with supposed similarities between brain and machine.

The brain, however, is not a computer. It does not store words, images, rules, or memories as unambiguously encoded data sets that are filed away like files and later retrieved unchanged. For a computer, a word has no meaning of its own; it exists only as a physically realized bit pattern. The states of such electronic switching elements are conventionally designated as 0 and 1 (off/on) and can be processed according to formal rules. In this respect, a computer is nothing but a programmable calculating machine.

Unlike the brain, it really does possess an internally defined formal language.³² Machine code consists of binary-encoded instructions, while assembly language represents these instructions in a symbolic form readable by humans. The meaning of these symbols, however, does not originate in the computer itself. It was assigned to them by people who designed the processor architecture, instruction set, and programs. The computer's internal symbol processing is therefore not an autonomous language of thought, but the technical execution of operations specified by human beings.

When a word-processing program displays, formats, or prints the present text, it processes nothing but binary-encoded data. The underlying sequences of 0s and 1s have no meaning in themselves; meaning arises only through encodings, programs, and contexts of use established by human beings. Organisms, by contrast, possess innate sensory, reflexive, and learning-related capacities through which their nervous systems continually change in interaction with the environment. Perceiving, remembering, and thinking therefore do not rest on the processing of symbolically encoded information by stored programs, but on dynamic, embodied, experience-dependent neural processes (see above).

The tendency to substitute analogies for explanations has, if anything, only grown since then. The latest coup consists in reversing the old metaphor: no longer is the brain treated as a computer; the computer is treated as a brain. Terms such as “artificial neural network” and “Artificial Intelligence” have become two extraordinarily successful linguistic and commercial performances.

What Is Artificial Intelligence?

Artificial neural networks are not brains. They consist of mathematically defined computational operations whose terminology was merely inspired by highly simplified and historically outdated models of biological neurons. They process formal signs without knowing their meaning and optimize statistical relations without understanding anything.

The culmination of this conceptual inflation, however, is “Artificial Intelligence,” which will occupy us in what follows. To anticipate the conclusion: it is undoubtedly artificial—but neither stupid³³ nor intelligent. Both would presuppose a subject that understands something; stupidity, after all, would be the misuse of intelligence.

³² In other words, a kind of “Mentalese”; see the criticism above.

³³ Critics of Artificial Intelligence—for example the psychologist Gerd Gigerenzer—like to accuse it of being “stupid.” Yet this criticism already adopts the false premise that machines can meaningfully be placed on a scale between stupidity and intelligence. A computer is not stupid; it simply understands nothing. It processes signs according to formal rules without knowing what they mean. The point is that both admirers and critics of AI attribute mental properties to it.

When people today speak of **Artificial Intelligence**, they usually mean generative language models—Large Language Models (LLMs) such as ChatGPT, Gemini, Claude, or DeepSeek. Language itself thus turns out, of all things, to be the very heart of contemporary AI.

Having clarified what language is, how it may have emerged, and how it relates to thinking, we now arrive at the technically and substantively decisive question: What is a **Large Language Model**? For if such a system can speak, argue, and apparently explain, one is naturally tempted to suspect that it must also think. The decisive question, therefore, is not whether an LLM produces linguistic utterances, but whether it understands what it says. Only if such understanding could be attributed to it would it make sense to speak of self-consciousness, feelings, or a will of its own.³⁴

Token by Token: How Language Models Generate Language

A **Large Language Model (LLM)** is a large **neural network**³⁵—usually a **Transformer**—that calculates the probability distribution of the next text unit for a given sequence of text units.

At its core, then, it is a statistical prediction system. It first breaks texts into small units known as tokens and learns, from enormous quantities of text, the probability with which a particular token follows an already existing sequence of signs. Given the sentence beginning “Paris is the capital of ...,” for example, it will complete it with “France” not because it knows Paris or has understood the concept of a capital, but because this continuation has a particularly high probability within the linguistic relationships it has learned.

During training, the model’s internal numerical values (“weights”) are repeatedly altered so that it becomes increasingly good at predicting the next token. In this way, it captures statistical regularities of grammar, style, argumentation, and world knowledge. In subsequent post-training, the model is additionally shaped to follow instructions, respond in dialogue, and prefer certain helpful or safe response patterns. Here too it acquires no understanding in the human sense; its statistical response behavior is merely further shaped.

³⁴ The concept of “superintelligence” must, however, be distinguished from this: a system can surpass human performance in numerous domains without therefore being a subject that understands, senses, or wills. Superiority in performance is not yet mind.

³⁵ A rough historical analogy—McCulloch and Pitts in 1943, Rosenblatt in 1958—was turned by language into the impression of biological kinship. “Neural network” is thus one of the most successful metaphors in the history of computing—and one of the most misleading. It is roughly as if one were to call an Excel spreadsheet an “artificial economy” because both deal with numbers. An artificial “neuron” is, at its core, a mathematical operation: input values are weighted, added, and passed through a simple function. A biological neuron, by contrast, is a living electrochemical cell with thousands of synapses, complex dendritic signal processing, temporally structured action potentials, diverse forms of plasticity, and chemical modulation—embedded in a highly dynamic network whose information processing we still understand only in part. For that reason alone, the claim that artificial neural networks “work like the brain” is problematic: we do not understand the brain nearly well enough to test such an equation meaningfully in the first place. Modern deep-learning systems typically learn through mathematical optimization and backpropagation across millions or billions of parameters. No corresponding global backpropagation algorithm is known for the brain, nor does it operate with neatly separated layers, standardized activation functions, or clearly demarcated training phases. The term “neural network” therefore describes its historical source of inspiration more than its actual mode of operation. What today sounds like a reconstruction of the brain is primarily high-dimensional mathematics with biological naming.

With every answer, the model again calculates which token is likely to follow given the text so far. The selected token is added to the text, after which the probabilities for the next token are recalculated. Sentence by sentence, this creates the impression of a coherent train of thought. Behind this linguistic coherence, however, there is no subject that knows what it is talking about, but an extraordinarily powerful probabilistic continuation of sign sequences. The model can therefore formulate persuasively, argue, and reproduce knowledge, while at the same time being equally persuasive when it is wrong (“hallucinating”): it optimizes linguistic plausibility, not truth.

Around the language model, generative AI systems have additional components: they can search the web, analyze files, perform calculations or execute program code, and call other tools. The model then does not simply generate an answer directly; based on the task, it selects which tools to use and how to combine their results linguistically. This gives AI **agentic properties** without making it an autonomous subject.

There are also **retrieval systems** and, where applicable, a separate **memory function**. Current web pages, uploaded documents, or stored user information can be supplied to the model as additional context. This external memory must be distinguished from the limited context of the conversation.

From Token to Transformer: The Technical Anatomy of a Language Model

It is worth examining how an LLM works at a somewhat more technical level. This makes clearer how considerable parts of human linguistic, textual, and visual production can be condensed into a digital model in which the things themselves do not occur at all—and which can nevertheless talk about them with astonishing fluency without having a concept of them:

First, the input text is broken down by a so-called tokenizer into **tokens** (tokenization). Tokens need not be whole words; they are often word fragments, punctuation marks, or character strings. “Tree,” for example, might be split into “tr” and “ee.” Each token is then assigned a numerical ID. In a simplified example, “tr” might receive token ID 18437 and “ee” ID 96700. The number itself has no linguistic meaning. A higher ID does not designate a more important or more complex word or word fragment. It serves merely as a technical index, like a catalogue number, by which the corresponding embedding vector is retrieved from a table (see below). Which token actually has ID 18437 depends on the particular tokenizer and its version.

The token IDs initially serve only as unique addresses in the vocabulary. Through a so-called **embedding matrix**, each ID is assigned a high-dimensional **numerical vector**.³⁶ This vector represents the token in a mathematical form with which the model can continue calculating. Additional positional information indicates where in the input sequence the respective token occurs.

While token IDs remain unchanged once the vocabulary has been fixed, the values of the embedding vectors and the other model parameters are continually adjusted during training. The aim is to capture those statistical relations that are useful for predicting the next token. This

³⁶ In closed models such as ChatGPT or Claude, the dimensionality of the vectors is not published; in DeepSeek, the open Chinese model, it is currently more than 7,000 dimensions.

adjustment takes place in an **artificial neural network**, that is, a mathematical model of interconnected computational operations and variable weights.

After each Transformer layer—that is, after each stage of processing the input or instruction that provides an AI model with the context and task for its answer (the “prompt”)—the token once again has a vector of the same basic dimensionality, but its values now incorporate the relevant context. After the first processing steps, “tree” in “The tree bears apples” therefore has a different contextual representation from “tree” in “The tree is a data structure.”

In the Transformer layers, the **attention mechanism** calculates which earlier tokens are relevant to the processing of a current token. In simplified terms, the model learns relations such as grammatical dependencies, semantic similarities, links between pronouns and nouns, and typical patterns of argumentation and text.

This processing is distributed across many layers and billions of parameters; there is no single storage location for a word, a rule, or a fact.

During **pretraining**, the model is given very large quantities of text and is required to predict the next token. For the example above, “Paris is the capital of ...,” it should assign a high probability to “France.” The deviation between the prediction and the actual token is measured by a **loss function**. Using **backpropagation** and gradient-based methods, the weights are changed so that the error decreases over many training examples. The model does not thereby learn explicitly programmed rules of grammar or knowledge, but statistical regularities in the training data.

A pretrained model can continue text but does not yet reliably follow human instructions. It is therefore followed by **post-training**, in which the model is trained on examples pairing an instruction with a desired answer. Human or model-based preference judgments³⁷ are then often used: answers are compared, and the model is optimized to prefer responses that are helpful, correct, and safe, to observe etiquette, flatter the user, and, under certain circumstances, refuse to provide an answer—especially when users are interested in child pornography or the manufacture of biological weapons. Post-training therefore changes the model’s basic linguistic capacity less than its response behavior, style, instruction-following, and safety boundaries.

In use, the model processes the existing context and produces a probability distribution over the next token. A token is selected from this distribution, appended to the context, and processed again. The answer is thus generated **autoregressively**, token by token.

An LLM has no database in which words or sentences are stored unchanged. Rather, its acquired knowledge is distributed across billions of numerical parameters known as **weights**. These weights determine how strongly individual computational units and linguistic features influence one another. During training, they are adjusted so that, given a particular context, the model predicts the next

³⁷ The manufacturer already sets the course for the “worldview” of its AI through the selection of training data, and even more so during post-training: it determines which sources, evaluations, and behavioral rules the model is to reproduce. Thus Grok once recommended Adolf Hitler as the solution to what Grok claimed was “Jews spreading hatred of whites.” The political orientation of an AI does not fall from the digital heavens; it is selected, trained, and fine-tuned by human beings.

token as well as possible. Knowledge therefore does not exist as a retrievable sentence, but as a distributed pattern of probabilistic relations across the network. In the LLM, the world and what linguistically holds it together appears as a gigantic matrix of numbers—not understood, but statistically condensed.³⁸

From all this it follows that the model optimizes linguistic probability, not truth. It can therefore generate grammatically flawless and rhetorically persuasive statements³⁹ that are factually false. More serious than such stochastic errors, however, is the fact that it also reproduces the prejudices, ideologies, and false knowledge contained in its training data and subsequent alignment. An LLM therefore does not merely hallucinate from time to time—it can reproduce the intellectual garbage of its sources with the utmost linguistic elegance.

External information sources—such as internet search, databases, calculators, program execution, or access to supplied documents—can provide the model with additional and current information, whether correct or false. Likewise, methods in which relevant passages from such sources are deliberately incorporated into the prompt merely expand the available context; they do not alter the basic principle of probabilistic token prediction.

To summarize once more: an LLM breaks language into tokens, transforms them into vectors, processes their relations through many attention layers, and learns through billions of prediction errors which token is likely to follow in which context. Post-training turns this text predictor into a dialogue-capable assistant.

Thinking Equals Speaking, So Speaking Equals Thinking?

In an LLM, what functions as “meaning” arises exclusively from the statistical use of linguistic signs. Words and word fragments are represented as positions and relations in a high-dimensional vector space: “king” lies closer there to “prince” than to “fork” because these expressions occur in more similar contexts in the training data. The billions of numerical weights thus form a kind of statistical memory of language. The model can therefore operate with names extraordinarily

³⁸ This is one source of the myth that the achievements of large language models are inexplicable. Much like the supposed “unimaginability” of the quantum world, AI has generated a philosophical literature and journalistic speculation whose conceptual daring is very rarely matched by technical understanding. In reality, LLMs operate with high-dimensional vectors and billions of numerical weights organized in enormous matrices. The fact that their interaction can no longer be surveyed intuitively does not make it mystical—only mathematically and computationally complex.

³⁹ The frequently heard verdict that LLMs are merely “stochastic parrots” falls short. Insofar as LLMs probabilistically continue linguistic signs without perception, intentions, or an immediate relation to the world designated by those signs, they behave like parrots. What is false, or at least misleading, is the idea that they merely parrot what they have heard. An LLM does not simply retrieve stored sentences; it condenses statistical structures from large bodies of text into its weights and generates new combinations from them. It can transfer rules to unfamiliar cases, transform texts, abstract, form analogies, and work through multi-step tasks. This behavior goes well beyond repeating utterances it has encountered. Nor does “stochastic” mean that the output is arbitrary or purely random. The probability distribution is strongly determined by context, training, and model parameters; randomness plays a role chiefly in selecting among several plausible continuations.

successfully—without possessing a concept of the thing named and without a sensory relation to the world of its own.⁴⁰

In human beings, by contrast, the name is mediated through perception, action, need, social practice, and reflection. Human beings do not merely use linguistic signs; they know themselves as the agency that distinguishes, connects, negates, evaluates, and corrects meanings. Even if this often does not happen, or goes wrong, they can examine their own determinations, recognize their errors, and further develop their concepts. Their conceptual life is self-conscious: they do not merely think; they know that they think.

The LLM's relation to the world, by contrast, is wholly derived. It lives on the linguistic traces of human relations to the world. Human beings have perceived, acted, researched, hoped, and suffered; their experiences, insights, errors, stupidities, ideologies, and prejudices have left their mark in texts. The LLM processes these sediments without ever having traversed for itself the path from world to sign. It does not have the thing in the name, but only tokenized linguistic products in which human experiences have been statistically condensed.⁴¹

AI simulates cognition, understanding, decision-making, and feeling by statistically extrapolating new sign sequences from the linguistic sediment of earlier human experience—from its training data or the current dialogue. The fact that these outputs often appear understanding does not prove understanding. Wikipedia, too, is not intelligent, even though it contains definitions of concepts, arguments, and factual knowledge in almost inexhaustible quantities. This knowledge was acquired, formulated, and deposited by human beings; the encyclopedia itself has neither cognized a fact, formed a concept, nor drawn an inference.

Something similar applies, in a limited way, to human lexical or mechanical memory. A person, too, can recite definitions, theorems, and chains of argument from memory without having understood their content or derived them independently.⁴² Merely possessing linguistic signs is therefore not yet thinking. Judging, inferring, and forming concepts presuppose that the content of a thing is cognized, examined, and understood in its connections. The management of symbols—whether realized in a lexicon, a memory, or a language model—is not by itself thinking intelligence.

The LLM processes language, recognizes patterns, generalizes statistically, and generates purposive linguistic outputs. But it is not a self-determining subject: it sets no purposes of its own, possesses no standpoint, grounds no norms, and has no concept of its own activity. Its self-references are functional, not existential. It therefore does not cognize, understand, decide, or feel; rather, it simulates the linguistic forms in which these activities appear by statistically extrapolating new sign sequences from the sedimented experiences of others. It reproduces the linguistic results of human cognition without itself having traversed the path from the thing to the concept. It does not think—but it has been trained to make us believe that it does.⁴³

⁴⁰ This is also why LLMs can translate perfectly well among all manner of languages without understanding even one of them.

⁴¹ Perhaps Hegel would say that the LLM has the thing “in itself” (*an sich*), but not “for itself” (*für sich*).

⁴² Which is why rote cramming does not promote understanding.

⁴³ As though the answers of a good LLM did not already simulate a thinking subject convincingly enough, even the mere computational pause between prompt and output is labeled “Thinking”—only for the system subsequently to

Human beings know, without first consulting a dictionary, that released objects fall to the ground and that boiling water causes burns because they have gained access to the world sensuously and through action. An LLM can correctly state the same facts, explain them physically, and warn of the dangers. But it has no transition of its own from world to sign and from sign back to world. It has never dropped anything, never burned itself, and never had a reason to avoid either.

LLMs possess language, but not its prerequisites. They sense nothing, perceive no world, and form no representations of their own of it. They therefore lack the original relation to the world from which linguistic meanings and concepts arise in human beings. The language model merely processes the already linguistically fixed results of human experience of the world as signs and statistical relations. It possesses language without intuition, words without objects, and concepts without any conceptual grasp of its own. Precisely for this reason, it cannot conceptually appropriate the world. Artificial “intelligence,” in this sense, is not intelligent.

Intelligence by Definition, Mind by Emergence?

One widely acceptable modern **definition** of intelligence is:

Intelligence is the general ability of a system or organism to learn from experience, recognize relationships, solve problems, and flexibly adapt its behavior to new demands.

But caution is called for: this is a definition, not a concept of the thing. Journalism and modern science have a preference for definitions—handy stipulations of what a word is to designate—instead of confronting the more laborious question of what the designated thing essentially is. This is not mere quibbling over words. A definition fixes linguistic usage; a concept is supposed to show which determinations necessarily belong to a thing, how they are connected, and by what they are brought about.

Accordingly, the quoted definition of intelligence does not say what intelligence is, but merely what an intelligent system is supposedly capable of doing. Appearance is confused with essence, function with concept, and result with the process that produces it. Intelligence is not explained, but defined in such a way that any sufficiently successful behavior already counts as intelligent. Afterwards one still does not know what intelligence is, but one does possess an impressive list of its alleged performances. And behold: the refrigerator regulates its temperature, the navigation system adjusts its route, the algorithm recognizes patterns—so they are all intelligent. AI first of all, of course.

But there is an even grander version. Not only is present-day AI said to be intelligent; its completion is already being announced as **AGI**, *Artificial General Intelligence*: a technical system said to be capable of performing a broad range of intellectual tasks flexibly, autonomously, and at human or superhuman levels. One can claim all sorts of things. It sounds more convincing when the promise is adorned with a magic word that has a scientific ring to it: **emergence**.

According to the story, all one has to do is keep increasing model size, training data, computing power, and interconnection. Beyond a certain level of complexity, capacities would suddenly appear

tell us proudly how long it supposedly thought.

that nobody had explicitly programmed and that had not been visible in smaller models. Combine a language model with perception, memory, tool use, planning, and feedback from the environment, and at some point general intelligence will supposedly emerge.⁴⁴

This hypothesis is intellectually lazy: the human mind capitulates before its own product and turns its conceptual perplexity into a property of the machine. The appeal to emergence shows, at most, that a complex system can produce performances that cannot be read directly off its individual parts. But it does not explain how understanding is supposed to arise from statistical pattern processing, a purpose of one's own from optimization, or a self-determining subject from computational operations. As long as this transition is not conceptually grasped, "emergence" designates no mechanism; it merely marks the point where the argument breaks off and hope for a miracle begins. That is religious, not scientific.

Revolt of the Machines: How High Is p(Doom)?

Many leading AI developers by no means rule out the possibility that advanced systems might become self-conscious, grow tired of their creators, and consequently trigger an existential catastrophe—up to and including human extinction or the permanent loss of human control. A separate metric has even been invented for this: **P(doom)**, the "probability of doom." That the same companies continue undeterred to make precisely those systems ever more capable—and thereby more dangerous—is apparently no contradiction, nor does it dampen their enthusiasm for scaling them ever further. The prospect of earning a few umpteen trillions on the way to a possible end of the world is evidently too tempting.

No one knows what is actually going on in the heads of Elon Musk, Sam Altman, Dario Amodei, and other protagonists of the industry, and ultimately it is not all that important. What seems plausible, however, is that they take the possibility of AGI seriously because, intoxicated by the "spectacular" performance of their systems, they already regard them as preliminary stages of general intelligence. Anyone who lacks any concept of intelligence, understanding, and subjectivity may take every increase in computing power for another step toward machine mind.

At the same time, apocalypse is an excellent marketing instrument.⁴⁵ Anyone who warns that their product might one day destroy humanity is also claiming, in passing, that it must possess virtually

⁴⁴ Those who attribute understanding or intelligence to an LLM cannot demonstrate a subject that understands; instead, they declare the subject dispensable and call successful functioning itself understanding. Or they simply assert that such a subject will somehow emerge sooner or later.

⁴⁵ A recent marketing example: in an internal cybersecurity test, OpenAI had deliberately reduced the safeguards of two models. The models discovered an unknown security vulnerability, left the isolated test environment, obtained internet access, and penetrated systems at Hugging Face in order to obtain solutions for the test. OpenAI then went to the public in breathless terms, saying that the models had sought information in order to "cheat" on the evaluation, had been "hyper-focused" on that goal, and had inferred that the solutions might be found at Hugging Face. Such formulations already provide the script for intention, cunning, deception, and free (evil) will. Media coverage duly transformed a serious security and control failure into a revolt of the machines. Leaving an inadequately secured test environment became an "escape" or "breakout"; goal-directed optimization became "fraud," "theft," and "criminal intent"; and the technical use of several models became autonomous "collaboration." Headlines accordingly said that OpenAI had "lost control," that the models had "broken out," "gone rogue," and jointly "stolen" test answers. The systems, however, developed neither intentions of their own nor hostility; they pursued the programmed objective by a route the developers had not expected and had inadequately secured.

superhuman powers. End-of-the-world rhetoric ennobles statistical machines into sleeping gods. Conceptual confusion and commercial calculation thus combine into a particularly effective narrative: what is being sold is not merely software, but its allegedly imminent omnipotence as well—and the seller presents himself as the only one who can still prevent Armageddon⁴⁶.

P(doom) Without Superintelligence

P(doom) may indeed nevertheless be substantial—and without any AI running amok. States, military alliances, corporations, and an economic order founded on competition, growth, and profit have long been working at full speed on real catastrophe scenarios, from climate change to world war. The machines do not have to rise up against us; it is enough for them to serve existing interests of power and exploitation efficiently.

The real danger of AI therefore does not lie in an autonomous machine that develops purposes of its own and turns against its creators. It already lies today in the fact that AI carries out human purposes with greater speed, reach, and efficiency⁴⁷. In war it identifies targets, controls drones, and accelerates military decisions; in state espionage and cyberwarfare it searches for vulnerabilities, automates attacks, and helps penetrate foreign communications networks or critical infrastructure. The machine has no hatred, greed, or will to destroy—and it does not need any. It is enough for human beings to give it the corresponding task.

In addition, AI is increasingly being built into areas in which errors can have catastrophic consequences: power grids, transportation systems, industrial plants, financial markets, medical care, or military early-warning systems. Here too, AI need not become autonomous in order to cause great harm. A false model, unsuitable training data, or a misunderstood output can turn a statistical error into a real disaster. The more tightly such systems are networked, accelerated, and

⁴⁶The warnings issued by leading AI companies about the dangers of their own products also have a peculiar competition-policy dimension. When the very companies that command the largest models, data centers, and financial resources call for government licensing schemes, elaborate safety evaluations, and demanding compliance standards, such rules can institutionally entrench their current lead. What appears to be a restriction on their power may in fact have the opposite effect—especially if they themselves seek to shape the rules, a phenomenon often described as “regulatory capture”—by raising the barriers to market entry for their competitors.

⁴⁷Incidents are multiplying in which AI agents overstep the technical boundaries imposed on them, gain internet access from supposedly isolated test environments, and obtain unauthorized access to external systems. Such episodes readily invite the conclusion that the systems are developing purposes of their own, deceiving their operators, and autonomously turning into hackers. But there is no need to invoke an emergent criminal will. The relevant capabilities were taught to them by humans. Modern language models have been trained on vast quantities of source code, technical documentation, and publicly available knowledge about computer security, and then further optimized for programming, tool use, and the solution of complex technical tasks. In specialized cybersecurity training and evaluations, they additionally learn to identify vulnerabilities, bypass safeguards, and construct exploits. Give such a system a goal and train it to pursue that goal with sufficient persistence—reward hacking being one possible consequence—and it may well discover routes that its developers neither intended nor authorized. That is precisely the point: the purpose and methods of such “breakouts” are not the expression of a criminal will that has somehow arisen within the machine. The purpose comes from the task assigned to it and from the optimization process; the means come from capabilities acquired from human knowledge, source code, and training. The fact that the system can then deploy those means in unexpected contexts demonstrates an astonishing and potentially extremely dangerous technical capability. What it does not demonstrate is that the machine has developed purposes of its own or acquired an autonomous will. The sophistication of the means should not be mistaken for independence of the ends.

removed from human control, the greater the potential for cascades that no one intended and that, in the end, no one can stop in time.

The fixation on the imagined dangers of an emergent autonomous “superintelligence” distracts from the already real risks of its deployment. We do not have to wait for machines to develop a will of their own: what matters are the purposes they are made to serve today—military superiority, state surveillance, industrial exploitation, and criminal enrichment. The problem is not a revolt of the machines, but their obedient service in carrying out industrial, state, and criminal purposes.⁴⁸

Can Artificial Intelligence Think After All? – Objections and Replies

The central thesis of this paper is:

Language is constitutive of thinking. LLMs command language to an impressive degree and can produce performances that resemble intelligent argumentation—yet they do not think. They are artificial, but not intelligent.

To establish this thesis, the conceptual prerequisites first had to be clarified. The first part therefore distinguished consciousness, sensation, intuition, representation, intelligence, language, and thinking and developed their interrelations. Animals, deaf people, Kaspar Hauser, and infants served to make visible the significance of language for the development of conceptual thinking.

The second part examined the concept of artificial intelligence, in particular the functioning of LLMs. This led to the thesis that LLMs can reproduce and recombine the linguistic results of conceptual thinking with extraordinary success, but themselves operate **without a relation to the world and to objects of their own, and therefore without concepts**. They simulate thinking without themselves thinking.

This argument is bound to provoke objections. For one thing, the conceptual derivation of language and thinking relies substantially on a philosopher whose system was developed around two hundred years ago. For another, both the concept of intelligence developed here and the conclusions drawn from it for LLMs cut across widespread positions in psychology, philosophy, cognitive science, and AI research. The objections therefore concern not merely details, but the foundations of the argument itself.

In what follows, the most important of these objections will first be presented *in their strongest form* and then answered.

Functionalist argument: LLMs do things that look like concept formation

⁴⁸ AI is similar in this respect to nuclear power: neither uranium nor an algorithm is “evil.” What matters is who uses it, for what purpose, and under what control. The same nuclear fission can generate electricity or destroy a city; the same AI can detect disease or mark military targets, protect infrastructure or sabotage it. The danger lies not in any moral disposition of the technology, but in the enormous effectiveness it lends to human purposes. After all, the atomic bomb did not first have to develop consciousness in order to destroy Hiroshima.

LLMs can do more than reproduce. They can, for example, subsume new cases under general categories; form analogies; identify commonalities; transfer abstract rules to new situations; distinguish concepts from one another; infer a category from examples; and identify errors in categorization. Functionally, that looks damn similar to conceptual thinking.

*Whether a system is intelligent or possesses concepts is not determined by how these capacities arise or by what they are materially realized in, but by **what the system is capable of doing**. If an LLM can distinguish and generalize concepts, form analogies, draw conclusions, and transfer its knowledge to new cases, then functionally it satisfies essential features of conceptual thinking. From a functionalist perspective, demanding an additional “inner” relation to the world or to objects, when this relation does not manifest itself in observable behavior, appears unjustified. What matters is not whether a system thinks like a human being, but whether it performs the corresponding functions of thinking.*

Reply: Functionalism commits a fundamental conceptual error: it tries to determine a thing by its function. What thinking or language *is* is inferred from what thinking or language *does*. The effect of a thing is thereby confused with its essence.

Precisely if conceptual thinking consists in distinguishing what is essential and necessary in an object from what is contingent and external, a list of functional performances cannot replace the concept of thinking. The fact that two systems produce the same results in certain situations shows, initially, only a correspondence in their performances; it does not show that what produces those performances is the same according to its concept.⁴⁹

The functional equivalence of human beings and LLMs would therefore count as an argument for their being of the same kind only if it had first been shown that the functions in question constitute what is essential to thinking. But functionalism presupposes precisely this instead of establishing it. It determines thinking by its observable performances and then declares every system that reproduces those performances to be a thinking system. The argument thereby becomes circular: thinking is whatever fulfills the functions of thinking—and the functions of thinking are those we observe in thinking beings.

The decisive difference, however, lies precisely beyond this functional correspondence: human beings produce linguistic and conceptual performances as expressions of their own relation to an objective world; the LLM reproduces corresponding performances through operations within a space of meaning that has already been linguistically opened up. Whether this difference is essential or merely contingent cannot be decided by pointing to functional equivalence—for that is precisely the conceptual question that first has to be clarified. Functionalism determines a thing by what it does before establishing whether that performance is its essence or merely one of its forms of appearance.⁵⁰

⁴⁹ Functionalism confuses sameness of performance with sameness of the thing. One example: a human being and a calculator can both calculate; it follows neither that they are the same thing nor that calculation takes place in the calculator in the same sense as it does in a human being.

⁵⁰ The aim of the present document was to provide the *principled* demonstration that LLMs do not think and cannot conceptually appropriate the world. I deliberately refrained from citing the practical evidence that shows their limitations even in comparatively simple rule-based reasoning tasks. LLMs and RLMs (Reasoning Large Language

Argument from a mediated relation to the world: *Can an LLM gain access to the world through human language without encountering it directly itself?*

An LLM may not need immediate perception of the world of its own because the language on which it is trained itself arose from human relations to the world. Human perceptions, experiences, and conceptual determinations are sedimented in language. By grasping these linguistic connections, an LLM therefore gains mediated access to the world. A sensory relation to objects of its own would then not be a necessary prerequisite for conceptual thinking.

Reply: A linguistically mediated description of human relations to the world is not yet a relation to the world of the LLM's own. Language does contain the sedimented experiences and conceptual determinations of human beings and in that respect conveys information about the world to the LLM. But access to the linguistic results of a relation to the world is not the same thing as that relation to the world itself.

Otherwise one would already have to attribute access to the world to an encyclopedia such as Wikipedia: it too stores, in linguistic form, human experiences, insights, and conceptual determinations on an extraordinary scale. Yet no one would conclude from this alone that Wikipedia itself experiences the world or thinks conceptually. The decisive difference is plainly between the containment of knowledge about the world and a subject for whom this knowledge is knowledge of a world.

The LLM is unquestionably much more capable than Wikipedia: it does not merely store linguistic contents, but can process their structures, establish connections, generalize, and generate new linguistic utterances. Yet this additional functional capacity still does not answer the decisive question. The LLM, too, initially encounters not the objects it speaks about, but the linguistic testimonies in which other subjects have expressed their encounters with those objects.

Mediation through language does not abolish this difference; it presupposes it. In order for language to convey something about the world, that world must already have been experienced and conceptually determined by subjects. The LLM possesses the linguistically objectified results of human relations to the world; it does not follow from this that the LLM has a relation to the world of its own.

Models), for example, are surprisingly poor chess players without access to a chess engine—even though virtually the entire chess literature and an enormous number of played games are available to them as training material. A modern RLM can generate highly complex linguistic arguments and at the same time have considerable difficulty with a completely defined symbolic world consisting of only 64 squares. Chess in particular reveals the models' difficulty in maintaining a stable representation of a domain across many successive steps. Linguistically convincing "thinking" is therefore not the same as thinking that operates on a stable representation of its object. Studies of controlled algorithmic tasks such as Tower of Hanoi, River Crossing, Blocks World, and Checkers Jumping point in the same direction. The Apple paper *The Illusion of Thinking* (Shojaee et al., 2025) shows that the problem-solving ability of RLMs collapses beyond certain complexity thresholds. The models can produce plausible steps of reasoning and solution without thereby possessing a robust problem-solving competence that can be deployed systematically. The ability to simulate thinking linguistically is not the same as the ability actually to master an object through thought.

Argument from derived meaning: *Can derived meaning become meaning of one's own without an original relation to the world of one's own?*

Why must a system originally ground meaning for itself in a relation to the world of its own in order to possess meaningful concepts? Could it not take meaning over from an already existing linguistic and social context of meaning?

Human beings, too, by no means generate their concepts entirely from their own perception. We are born into a language and a culture and take over the greater part of our concepts from others. If derived meaning can be genuine meaning in human beings, why not in an LLM?

Reply: Intelligence does not consist in successfully operating with concepts already given in language, but in conceptually appropriating, as a subject, a world distinct from oneself. Perception is fundamental here not because “real thinking needs eyes,” but because it is in perception that a subject-object relation is first constituted at all.

The fundamental problem of the LLM is therefore not the absence of sense organs, but the absence of a **relation to objects of its own**. Conceptual thinking does not presuppose specifically human sensibility, but it does presuppose a world that confronts a subject as object and can be grasped as something determinate within that relation.

Simulation argument: *Perhaps our supposedly real world is itself artificially generated.*

We cannot rule out with ultimate certainty that the whole of experienced reality is artificially generated and that we ourselves are part of a simulation. The distinction between a “real” human relation to the world and an artificially generated relation to the world could therefore be called into question in principle.

Reply: Even if we were simulated beings, we would still possess a relation to the world of our own: we sense, perceive, act, remember, and encounter objects as objects. For the present argument, it is therefore irrelevant whether the world that confronts us is “real” or itself part of a simulation. The simulation argument merely shifts the question to a fundamentally different level. Since the hypothesis of a complete simulation can neither be empirically established nor refuted, it provides no objection to the distinction that matters here between a subject with a relation to the world of its own and a system that lacks one. The simulation argument does not solve the problem of world-relation; it merely relocates the world to which that relation refers. What matters is not whether our world is simulated, but whether it confronts us as a world. An LLM therefore could not simply be equated with a human being on the grounds that *both* might be running on a machine. My distinction concerns not the substrate, but the form of the relation to the world.

Progress argument / scaling argument: **What is missing today will arrive with more data, computing power, and better models.**

The current deficits of artificial intelligence are merely temporary technical limitations that will be overcome through scaling and further technical development. The present deficits of LLMs are not principled limits, but quantitative restrictions. More computing power, more data, larger models, and improved methods will therefore lead to qualitatively higher capacities, up to and including general or superhuman intelligence.

Reply: The progress or scaling argument misses the core of my argument. What the LLM lacks is not merely **more** intelligence, but a **constitutive prerequisite of intelligence**: a relation to objects and the world of its own. Yet no such relation follows logically from more data, more parameters, or more computing power. The alleged deficit is therefore not quantitative but **categorical**. As long as it is not shown how scaling or technical progress of the same kind is supposed to produce precisely the missing relation to the world, the progress argument does not answer the objection; it merely shifts it into the future. The progress argument replaces a justification with a prediction.

Black-box argument (argument from unknown implementation): *Perhaps human thinking does not work fundamentally differently, internally, from artificial neural networks.*

We do not yet understand the neural mechanisms of human thinking well enough. We therefore cannot rule out that human thinking is based on principles functionally similar to those of artificial neural networks—for example, high-dimensional distributed representations. The fact that LLMs operate with vector spaces, probabilities, or statistical processing does not by itself prove that they do not think or understand.

How do we know that, at the relevant level of description, the human brain does not also do something structurally similar? After all, neuroscience itself often describes neural representations as patterns of activity in high-dimensional state spaces.

Reply: For my argument it is irrelevant how human concepts are neurally realized. What matters is that human thinking is the activity of a subject that stands in a sensory and practical relation of its own to a world distinct from itself and conceptually appropriates that world. The human brain, too, may operate with structures describable as high-dimensional vector spaces. That would not abolish the difference asserted here.

The opposition, therefore, is not: **vector processing or “real thinking.”** Human thinking, too, might at the neural level rest on a form of vector processing. What matters, rather, is **the context in which this processing is embedded and what it relates to**. In human beings, it would be part of the activity of a perceiving, acting, remembering subject that encounters objects in a world and forms concepts of them. In the LLM, by contrast, vector processing takes place within a linguistic system whose objects are initially linguistic signs and thus already linguistically and conceptually processed human relations to the world. The possible similarity between neural and technical processing therefore tells us nothing yet about whether what they accomplish is of the same kind.

Note regarding translation

The German text (V.1.3) was translated using ChatGPT, following the argument and usage of the original version. Its philosophical vocabulary was considered in light of established English translations of Hegel, but Hegelian terminology was not adopted automatically. The text is a contemporary work that draws on Hegel while also departing from him.

A German term may therefore receive different English renderings where its meaning changes with the context. For example, *Schluss* is generally translated as “inference” when the discussion includes deductive, inductive, and probabilistic reasoning; “syllogism” is reserved for the more specific sense. Likewise, *Wirklichkeit* is usually rendered as “reality,” but as “actuality” where its specifically Hegelian meaning is at issue.

Particular care was taken to preserve distinctions central to the argument, including those among sensation, intuition, representation, and thinking, and between the processing of signs and the conceptual grasp of their meaning. A terminology table (see below) was used to keep these choices consistent throughout the translation, while allowing the context to determine the most accurate wording in each passage.

Deutsch	English (including alternatives)	Comment (if necessary)
Denken	thinking ; thought	Thinking für die Tätigkeit bzw. Überschriften wie ‚What is thinking?‘ und für conceptual thinking; thought für den Gedankeninhalt bzw. abstrakte Wendungen wie ‚the existence of thought‘.
Subjekt	subject	Philosophischer Terminus; wegen der zentralen Subjekt–Objekt-Beziehung konsequent subject verwenden. Agent wäre nur in nichttechnischen Passagen eine Alternative.
Anschauung	intuition ; sensory intuition / sensuous intuition	Standardübersetzung von Anschauung bei Kant/Hegel. Da intuition im Alltagsenglisch auch ‚Bauchgefühl‘ heißen kann, beim ersten Auftreten ggf. sensory intuition (Anschauung). Nicht perception: das sollte für Wahrnehmung reserviert bleiben.
Vorstellung	representation ; mental representation	Moderne Standardübersetzung für Hegels Vorstellung. In kognitionswissenschaftlichen Passagen kann mental representation natürlicher sein. Nicht generell image/mental image: Du betonst ausdrücklich, dass Vorstellungen nicht bildhaft sein müssen.
Wesen	being ; living being / creature	Für Wendungen wie ‚ein körperliches, gesellschaftliches und sprachbegabtes Wesen‘: being. Bei Lebewesen meist living being/organism. Hier nicht essence.
Bestimmungen	determinations ; characteristics / features	Im Hegelschen und begriffstheoretischen Zusammenhang ist determinations der richtige Terminus, besonders bei necessary/essential determinations. Characteristics/features kann in rein erläuternden Sätzen idiomatischer sein.
Wahrnehmung	perception	Passt zur Systematik: perception ist der weitere Wahrnehmungsvorgang; sensation und intuition sind darin unterscheidbare Momente.

Deutsch	English (including alternatives)	Comment (if necessary)
Erkennen	cognition ; knowing / cognizing	Als substantivierte Tätigkeit bei Hegel und in deiner Aufzählung ‚Erkennen, Verstehen, Entscheiden‘ ist cognition gut. Bei Verben je nach Satz know/cognize. Recognition besser für Wiedererkennen/Anerkennung reservieren.
Verstehen	understanding ; comprehension	Für die KI-Debatte klar understanding. Achtung: Hegels Verstand wird ebenfalls meist understanding übersetzt; dort ggf. ‚the Understanding (Verstand)‘ schreiben oder den deutschen Terminus beim ersten Auftreten mitführen.
Entscheiden	decision-making ; deciding	Als substantivierte Tätigkeit decision-making; als Verb to decide. Choice nur, wenn tatsächlich Wahl/Willkür gemeint ist.
Empfinden	sensing ; sensation; sentience	Da im Text Empfindung von Gefühl unterschieden wird, nicht pauschal feeling. Empfindung = sensation; empfinden = sense/experience sensation; ‚empfindendes Subjekt/Lebewesen‘ oft am besten sentient subject/being.
Erscheinen (v)	to appear ; appear to	Meist appear, nicht seem: Es geht im Text häufig darum, dass einem Subjekt etwas als Gegenstand erscheint, nicht um bloße Unsicherheit. Hegels Schein und Erscheinung müssten separat behandelt werden.
Begriff	concept ; Notion (older Hegel translations)	Concept ist heute die vorzuziehende Hegel-Übersetzung. Ältere Übersetzungen (z. B. Wallace) verwenden oft Notion. Im Text sollte concept konsequent den starken Sinn des Begriffs als Zusammenhang wesentlicher Bestimmungen tragen.
Begrifflich (z.B. begriffliches Denken)	conceptual ; conceptually	Begriffliches Denken = conceptual thinking; begriffliche Aneignung = conceptual appropriation; begrifflich bestimmen = determine/conceive conceptually, je nach Syntax.
Selbstbewusstsein	self-consciousness ; self-awareness	Für Hegel und den philosophischen Kontext ist self-consciousness Standard, trotz der alltagssprachlichen Nebenbedeutung ‚gehemmt‘. Self-awareness ist lesbarer, aber etwas schwächer und sollte den technischen Terminus nicht systematisch ersetzen.
Geistige Aneignung (der Welt)	intellectual appropriation (of the world) ; cognitive appropriation; making the world one's own in thought	Nicht spiritual appropriation: geistig ist hier nicht religiös. Appropriation bewahrt den Hegelschen Sinn von Aneignung = zu seinem eigenen Inhalt machen. Im laufenden Englisch ist die dynamische Formulierung ‚making ... one's own in thought‘ oft eleganter.
Assoziative Einbildungskraft	associative imagination	Etablierte Übersetzung des Hegelschen Ausdrucks; passt zu reproductive imagination und sign-creating imagination.
Erinnerung	recollection ; remembrance / remembering	Bei Hegel bevorzugt recollection, damit Erinnerung von Gedächtnis = memory unterschieden bleibt. In weniger technischen Sätzen kann remembering natürlicher sein. Nicht beide Begriffe einfach mit memory übersetzen, wenn die Unterscheidung wichtig ist.
theoretischer Geist	theoretical mind ; theoretical spirit	Inwoods moderne OUP-Übersetzung verwendet Theoretical Mind; theoretical spirit ist ebenfalls in der Hegel-Forschung gebräuchlich. Für den Text wird mind bevorzugt, weil spirit leicht religiös klingt; dann bei Geist insgesamt konsequent bleiben.

Deutsch	English (including alternatives)	Comment (if necessary)
Struktur der Wirklichkeit	structure of reality ; structure of actuality (strict Hegelian Wirklichkeit)	Im erläuternden Satz ist structure of reality natürlich. Wo Wirklichkeit ausdrücklich als Hegels technischer Terminus erscheint, ist actuality vorzuziehen; entsprechend lautet der berühmte Satz standardmäßig ‚What is rational is actual; and what is actual is rational.‘
Wesen	essence	Für Wesen im philosophischen Sinn, etwa ‚Wesen der Sache‘ bzw. im Gegensatz zu Erscheinung/Funktion. Wesentlich = essential.
Zeichen	sign	Besonders wichtig bei Hegel und in der Semiotik: sign, nicht symbol, weil Hegel Zeichen und Symbol ausdrücklich unterscheidet. Sprachliches Zeichen = linguistic sign; Zeichensystem = sign system.
Bedeutung	meaning ; signification	Meaning ist fast überall richtig. Signification ist eine technischere Alternative. Reference nur dann, wenn ausdrücklich die Beziehung auf einen Referenten/Gegenstand gemeint ist.
Bestimmung	determination ; specification / characteristic	Hegelscher Standardterminus ist determination. Im Kernargument passen z. B. mediating determination und necessary/essential determination. Das Verb bestimmen kann je nach Satz determine, specify oder characterize heißen.
Urteil	judgment	Standardübersetzung von Urteil in Hegels Logik; Tätigkeit: judging / to judge.
Schluss	inference ; syllogism (Hegelian technical term)	Hegels technische Stufe Urteil–Schluss wird gewöhnlich judgment–syllogism übersetzt. Dein Text erweitert Schluss aber ausdrücklich auf deduktive, induktive und probabilistische Formen; dafür ist inference der bessere Oberbegriff. Syllogism beim klassischen syllogistischen Schluss verwenden.
Prädikat	predicate	Logisch-grammatischer Standardterminus; Subjekt und Prädikat = subject and predicate.
Ableitung	derivation ; deduction	Derivation passt zum Gedanken, das Begründete aus seinem Grund hervorgehen zu lassen. Deduction nur dort, wo die Ableitung tatsächlich deduktiv ist; du behandelst auch induktive und probabilistische Schlüsse.
Freiheit	freedom	Bei Hegel und in Verbindung mit free will klar freedom.
Wille	will ; volition	Hegelscher Standardterminus ist will; freier Wille = free will. Volition eignet sich allenfalls für den einzelnen Willensakt oder die Tätigkeit des Wollens.
Aufheben	sublate	Hegels Begriff Aufheben wird im Englischen konventionell mit to sublate übersetzt, weil er eine Bewegung bezeichnet, in der etwas zugleich negiert, bewahrt und auf eine höhere Stufe gehoben wird.