PHILOSOPHY & DIGITALITY
AN OPEN ACCESS, PEER-REVIEWED JOURNAL

On the Very Idea of a (Synthetic) Conceptual Scheme

Shane Denson

Abstract: This article critically engages with M. Beatrice Fazi's provocative argument that contemporary forms of generative AI, particularly large language models (LLMs), produce genuine language through acts of philosophical synthesis akin to human cognitive processes. While Fazi proposes a Kantian-inspired model wherein LLMs synthesize self-enclosed, coherent "worlds," I challenge this idea, questioning the coherence and implications of positing AI as creating separate worlds detached from shared human experience. Drawing on Donald Davidson's critique of conceptual schemes, along with insights from the phenomenological tradition, I argue instead for a distributed and open model of synthesis. Without equating computational processes with human agency, I emphasize the inherently mediated and social nature of both human and computational language generation, suggesting that generative AI mediates worldly significance not by synthesizing separate worlds but by intervening in this, the only world. Ultimately, I advocate a nuanced recognition of generative AI's transformative role within shared forms of life.

Keywords: generative AI; Large Language Models; synthetic media; philosophical synthesis; phenomenology; conceptual schemes

1 In an article titled "The Computational Search for Unity: Synthesis in Generative AI," M. Beatrice Fazi argues for an important and provocative reorientation in our thinking about contemporary generative AI and large language models (LLMs) in particular (Fazi 2024). Whereas the texts (and images) produced with such models—and sometimes the models themselves—have been referred to as "synthetic media" in order to emphasize their artificiality, or the fact that they are "contrived artifacts or manufactured manipulations" (Fazi 2024, 32), Fazi urges us to consider them instead in the light of philosophical concepts of synthesis. On this suggestion, LLMs are synthetic media not in the sense of producing "fake" or "simulated" (as opposed to "natural" or "real") language; rather, they are synthetic in the sense that in order to create their outputs, they must first complete operations of "amalgamation, […] composing and putting together" (Fazi 2024, 33).

Contact: Shane Denson <shane.denson@stanford.edu> ORCiD: 0009-0005-4638-7700
https://doi.org/10.18716/pd.v2i1.11664

2 So far, this reorientation may seem rather innocuous, as it simply "stresses the integrative and combinatory aspects of anything synthetic" (Fazi 2024, 32). However, the turn away from the idea of synthetic-in-the-sense-of-simulated significantly signals an acceptance that what synthetic media produce is in fact *real*. The latter idea has occasionally, and sometimes vehemently, been contested: LLMs are suspect, according to some critics, because they don't produce real language, or genuine linguistic expressions, but only the semblance thereof.[1] Lacking the motivation of intention, the outputs of LLMs are not categorically different from the marks made by a wave crashing on the shore, apparently spelling out verses from Wordsworth; we project meaning onto them, but they don't actually mean anything on their own.[2] (On this view, we could say that although it is we who "prompt" LLMs to produce new texts, those outputs in turn serve as "prompts" for us to provide them with sense. All the meaning-making remains squarely on the side of the user/reader.) Against this background, Fazi's reorientation is first of all significant because it ascribes linguistic meaning to LLMs' outputs themselves and therefore contradicts the argument from intention.

3 Importantly, though, Fazi does not ascribe intention to the "speaker" of those outputs, so her ascription of meaning is rooted differently, and this is where the concept of synthesis comes in. According to Fazi, contemporary AI is capable of what she calls "*synoptic computing*," which she clarifies as follows: "by 'synopsis' I do not mean a precis or a digest but, etymologically, a 'seeing all together,' so a mode of computing that affords a view of the whole" (Fazi 2024, 34). Fazi explicates this act of synopsizing by way of philosophical theories of synthesis (exemplified in Aristotle, Kant, and Hegel), as well as via the concept of "compositionality" in neuroscience, cognitive science, and linguistics. In both cases, mental and linguistic acts of representation are seen to rest on processes of structuring and unifying a manifold or multiplicity into a new, synthetic whole. Sidestepping the question of intention, as well as that of referentiality, representation and meaning are here approached in terms of the ability to create a holistic or "coherentist" model of the world. (Note: Fazi does not use the term "coherentist." I use it here to suggest an analogy with coherence theories of truth and knowledge, and thereby to emphasize a contrast with correspondence theories. Whereas the latter seek to articulate conditions of agreement between a statement or proposition and a state of affairs in the world, coherence theories suggest that everything must be taken together, that each statement has its meaning only in relation to the others, and that it is this system itself which must face the test of reality.)

4 Ultimately, then, Fazi's reorientation is radical indeed: If LLMs are synthetic media not in the sense of producing fake language but in the sense of "composing and putting together" real language, i.e. producing genuinely meaningful representations, then this is because they are able to execute synthesizing operations much like those that characterize the way human minds work.

1. See, for example, (Bender and Koller 2020) and (Bender et al. 2021).

2. The "wave poem" image derives from (Knapp and Michaels 1982); it has recently been taken as a touchstone for a dossier on meaning and LLMs in (Kirschenbaum 2003).

(Here we might pause to anticipate an objection: obviously, LLMs often fail spectacularly to generate genuinely meaningful statements. But so do humans. In any case, occasional or even frequent failures do not militate against the claim that these models sometimes succeed in producing meaningful language and that these successes, which would appear to be increasing in frequency along with the sophistication of new LLMs, are owing to acts of synthesis in the relevant sense.) Like the mind, which Fazi characterizes (extrapolating from Kant) as a "structure that structures—or a unity that unifies" (Fazi 2024, 37), LLMs accordingly synthesize new wholes out of discrete parts, and in this way they transcend the merely mechanical operations of statistical reconstruction (or the stochastically modified regurgitation of training data), in effect achieving their own representations of the world.

5 Clearly, this is a profound argument, and one with far-reaching consequences. Though Fazi does not impute any form of subjectivity to AI, the homology of synthesis between mind and machine might be put forward as a necessary condition for artificial general intelligence (AGI) and thus taken as a sign of the latter's plausibility. I will bracket discussion of this possibility in the following. Moreover, I will argue that we can provisionally bracket the notion of representation as it functions in Fazi's argument as well. This is possible because I understand Fazi to be arguing that synthesis is a necessary and not sufficient condition for representation, however we might understand that term. For example, she writes: "Synthesis is the active principle driving computation's search for unity, which, in turn, is fundamental to the large language model's management of its distributed representations" (Fazi 2024, 53). Here it is clear that (the management of) representation depends on synthesis, that the latter is a necessary condition for the former—and hence we should be able to consider synthesis on its own terms. As we shall see, however, it is not quite so simple a matter to separate the two as they operate in Fazi's argument, and I will therefore have to argue for their (provisional) severability. In any case, I take Fazi's article to be part of a larger project, and I expect that she will spell out further the meaning of representation in LLMs, perhaps in ways that will complicate my argument here. For now, though, I would like to focus solely on synthesis and its connection to world-modeling.

6 For Fazi, the synopsizing activity of an LLM, while homologous to the synthesizing activity of the human mind, is different in at least one respect: the LLM does not sense the external world, and therefore it does not refer to the external world. Rather than *the* world, the LLM's synthesized model is an insular model of *a* world (Fazi 2024, 47). In the following, I would like to question the coherence, so to speak, of this coherentist vision. Following Donald Davidson's classic "On the Very Idea of a Conceptual Scheme" (Davidson 1984), we might question whether it makes sense to posit this difference between "the world" (our world) and "a world" (the self-enclosed world of the machine). If we reject the meaningfulness of that distinction, we might go on to question the model of synthesis that has given rise to it. In other words, we could simply reject Fazi's contention that LLMs are "synthetic media" in the quasi-Kantian sense of a "structure that structures." If we did so, it would seem that (in the absence of another rebuttal to the argument from intention) we would then have to revert to the more common sense of the term: LLMs are "synthetic media" because they produce only the semblance of meaning. But if we in-

stead agree, as I do, that Fazi's reorientation offers something significant to our thinking about AI, we might instead follow a different tack: perhaps what we then discover is that an LLM's synthesizing activity is not limited to an insular model of *a* world but ranges in fact over *the* world, or some portion(s) of it—a world the computer shares in common with us. The model's reference to the world may be imperfect, and it may or may not be proper to say that it *represents* that world (again, I will argue that we can and should bracket this question for now), but its words (taken as real, i.e. meaningful linguistic expressions) would nevertheless pertain to the same world that we intend when we speak about common objects around us or things that happened on the Internet. Later, I will consider what conditions would have to obtain for this to be the case and suggest how they might plausibly be fulfilled. First, I turn to the details of Fazi's argument.

Synthesis and Representation

7 How does Fazi arrive at the conclusion that "synoptic computing" involves LLMs in an act of world-modeling? Though she refers to Aristotelian and Hegelian understandings of synthesis as well, Fazi's philosophical touchstone here is Kant. She points to Kant's understanding of "synthesis [as] a mental unification of multiplicity" (Fazi 2024, 36), as exemplified in the three syntheses of intuition, imagination, and understanding. Because, as we have seen, Fazi admits that LLMs do not have sensory access to the external world, the synthesis of the understanding, which deals in concepts, would seem to be most pertinent in this context (since intuition pertains to sensory information, which the imagination, equipped with generalizing schemata, intercepts and passes along for conceptual labeling). In today's multimodal models, which are capable of processing and generating descriptions of images, video, and audio, for example, it falls to an LLM to provide the concepts, and the model thus functions similarly to the understanding. It is open to question whether the three syntheses (or their objects) are in fact severable, but in any case, Fazi is less interested in a particular synthesis, such as that of the understanding, and more interested in the general form of the synthesizing activity.

8 Noting that, for Kant, "[w]hat is apprehended in intuition, reproduced in imagination, and recognized in concepts are representations (i.e., intuitions, images, and concepts are representational)" (Fazi 2024, 37), Fazi goes on to quote Kant: "Wherever our representations may arise, whether through the influence of external things or as the effect of inner causes, […] they must all be ordered, connected, and brought into relations." (Kant 1998, 228) This is synthesis "in the most general sense"—as, according to Kant, "the action of putting different representations together with each other and comprehending their manifoldness in one cognition." (Kant 1998, 210) For Fazi, "organizing information," or "structuring it," is key (Fazi 2024, 37). For human minds, "this structuring is the unification, in thought and by thought, of a multiplicity that is given to mind" (Fazi 2024, 37). For LLMs, on the other hand, we have to abstract this structura-

tion away from the reference to mind in order to avoid begging the question about AI's subjectivity or cognitive "intelligence." As I stated above, I think we should also bracket the question of representation, and for similar reasons. Note that for Kant, synthesis both *acts on* representations (it is "the action of putting different representations together"), but it also appears to *constitute* them, or at least serves as a *horizon for their appearance* ("wherever our representations may arise, […] they must all be ordered" etc.—i.e. they do not appear outside of the structuring operation of synthesis). This suggests, if not mere circularity, then something like a transductive relation between synthesis and representation—i.e. a relation that is fundamental to the existence of the relata, which are thus constitutively and originarily related (Simondon 2020). But perhaps this only holds, if indeed it does, for the human mind, or for a definition of information that is inherently representational. In any case, synthesis as "organizing information," or "structuring it," would seem to require neither cognition nor representation in any sense that would have to be construed as subjective, intentional, or even referential.

9 Fazi holds on to the language of representation throughout her article, but in a sense that would seem to fulfill these requirements. She explicitly distances her notion of representation from the communicative intent that critics like Emily M. Bender, Timnit Gebru, and coauthors marshal as a criterion of meaningful expression (failure of which allegedly renders LLMs nothing more than "stochastic parrots" (cf. Bender et al. 2021) (Fazi 2024, 43-44), and she even attests that "[l]arge language models lack what linguistics calls a *referent*. Reference here is the relationship between a linguistic expression and what, in the world, the expression is supposed to represent" (Fazi 2024, 46). Accordingly, representation is here divorced from the relationship that is supposed to underwrite representation itself! And this separation would indeed seem necessary in the context of today's generative AI, which operates not with symbols and rules but subsymbolically, hence by decomposing the linguistic units that are central to human expression and reference into mathematical and computable, but hardly cognizable, vectors. Hence, according to Fazi, "while these systems might not refer to an outside reality, they still need to build their representational reality, which is a sum of word embeddings and their internalized relations but also a network of evaluations and elaborations of vector representations via the processing of signals across weighted nodes and layers in the net. This is not the direct encoding of representations of the external world, but the encoding of representations nevertheless" (Fazi 2024, 46-47). We begin to see here how the Kantian model, which explicates the synthetic unity of apperception entirely in terms of a system-internal processing of phenomena, apart from the radically exterior noumenal realm, guides Fazi's thinking about representation in the LLM: "The philosophical concept of synthesis holds substantial speculative relevance to addressing the prospect that this representational reality, internal to the language model, could be a stable (if imperfect) whole, a togetherness of distributed representations—not *the* world as such but *a* world" (Fazi 2024, 47).

10 Fazi identifies an "exciting prospect" here, basically a form of nonhuman representation: "It is evidently true that, because of the absence of such an external anchoring, a large language model is unlike a human being in its use of language. A large language model is a mechanism designed

to produce text statistically, estimating probabilities. Once that is acknowledged, however, and thus, once philosophers, linguists, and computer scientists accept that they can turn their attention away from attempting to appraise machines against human standards, the same philosophers, linguists, and computer scientists are offered the exciting prospect of investigating a mode of representation that does not rely on human interpretation and whose synthetic generativity is thus self-sufficient" (Fazi 2024, 48-49). I agree with Fazi that this is an exciting prospect, but I worry that the Kantian model is taking too strong a lead here, and that a very human-centric notion of representation remains at the heart of this picture. This is because Fazi seeks a formal homologue to Kant's apperceiving mind as the basis of synthesis: "minds […] have themselves to exhibit a distinctive unity that would allow for such synthesizing activity to occur in the first place (this is the Kantian *transcendental unity of apperception*)" (Fazi 2024, 37). Fazi suggests that LLMs similarly have to exhibit such structure, that they must be "a structure that structures—or a unity that unifies" (Fazi 2024, 37), and thus the LLM's world model answers to the "computational search for unity" (Fazi 2024, 42) that would enable synthesis.

11 There is an elegance to this argument, which is effectively a transcendental argument in the Kantian vein—an argument for the conditions of possibility of synthesis. But by taking the transcendental unity of apperception as the model for the computational unity that enables machine synthesis, representation goes from a secondary phenomenon (enabled by synthesis as a necessary condition) to a foundational and originary one (transductively and inseparably tied to synthesis as its correlate). The unity of apperception, as a synthesizing synthesis, turns every synthesis into a representational affair, both subject and object of the synthesizing operation. And representation, in turn, establishes itself as a reflexive correlate of the "I think." But, given all the differences between human and computational uses of language, is this really the proper model? Would it not be better, in the case of LLMs, to dispense with the search for unity, which would allow us to bracket representation in this sense and focus instead on synthesis as a pre-personal, pre-subjective, or at least pre-reflective condition of having or being in the world?

12 The early Sartre, in *The Transcendence of the Ego*, drew on Husserl's phenomenology (turning it at times against Husserl's own explicit thinking) to argue that the unity of the "I think" was dispensable, a self-objectifying reification that distracts from the openness and indeterminacy of conscious experience (Sartre 2004). And while it is beyond the scope of this essay to pursue this line of thinking in detail, Sartre (and phenomenology more generally) opens up a space for thinking synthesis without closure, unity, or reflective representation. Synthesis, on this account, is a matter less of *re*-presentation than of the pre-personal presencing [*Gegenwärtigung*] that Husserl describes: the gradual coalescence of minute "adumbrations" [*Abschattungen*] revealed to ambulatory experience as I walk around a tree, or the primary retentions that come together, prior to subjective reflection, in my temporal experience of auditory phenomena (Husserl 2012; 1964). I suggest that this provides a better model for the synthetic operations of LLMs, whose subsymbolic workings might be thought of as impersonal *Gegenwärtigungen* or presencings of an open model. Perhaps this is not far from what Fazi has in mind with the "togetherness of distributed representations"—they are distributed in the sense of diffusely co-

present, not focused subjectively as we might think of symbols and concepts when they serve as the objects of expressive consciousness.

13 But without further careful specification, it seems better for now to bracket the idea of representation and to focus instead on the relation of synthesis to worldhood and its modeling. Is synthesis alone enough for the togetherness of distributed presencings, and is this synthetic togetherness sufficient for a world? The ego-less flux of impersonal experience described above is certainly a necessary condition, in the phenomenological tradition, for the emergence of and participation in a world, but it is unclear that it is sufficient. In *Being and Time*, Heidegger describes a forgetting of self in one's practical involvement in the world, but that world itself is marked out, as becomes clear when one is forced to reflect, by a referential order (including signs and other representational devices) within which our underlying projects are inscribed and by which they are oriented (Heidegger 1962). By foregrounding an existential openness of experience and contrasting it with the unity that, in Fazi's argument, is provided by representation, I might be seen to beg the question: is such openness really at odds with the LLM's constitution of a bounded world, the "synthetic generativity" of which, as Fazi says, is "self-sufficient" (Fazi 2024, 49)? Putting representation aside, let us now consider the unity of such a world on its own merits.

Worldhood and Alterity

14 "Strikingly, a large language model is also its own model" (Fazi 2024, 47). This is how Fazi describes the self-sufficiency of the LLM as a synoptic togetherness of a world. She elaborates: "ChatGPT's use of language, for example, does not originate from being alive and accumulating experiences over space and time; its use of language is not geared toward social action or communication, either, for it does not handle referential features and is not meant to act upon them. Extraordinarily, this is a 'symbolic order' (to use the Lacanian expression, although, in this connectionist setting, we should perhaps talk of a *subsymbolic* order) that lacks a referent culture and the beliefs, values, desires, goals, norms, intentions, and conventions that society and culture express through that language-mediated order" (Fazi 2024, 47). How are we to assess this claim?

15 Let us begin by observing that what Fazi calls a "world within" (Fazi 2024, 47) seems to be a species of what (Davidson 1984) calls a "conceptual scheme"—something like the *Weltanschauungen*, epistemes, or incommensurable paradigms that are advanced in discussions of conceptual relativism. Such schemes purport to describe a deep alterity, not unlike Fazi's LLM as a world within as opposed to our outer world; however, if the LLM's world is synthesized *without* a "referent culture," the conceptual schemes at stake in most discussions are offered as incommensurable *because of* their attachment to, or better: entrenchment in, radically different referent cultures. Whether the Sapir-Whorf hypothesis and claims that the Hopi language structures time differently than European languages, or Kuhn's suggestion that scientists before

and after a scientific revolution "work in different worlds" (Kuhn 1962, 134), as quoted by Davidson (Davidson 1984, 187), cultural-linguistic differences are posited as ontological differences. In his "On the Very Idea of a Conceptual Scheme," originally published in 1974, Davidson famously argues that we cannot make sense of such radical difference because 1) we lack a neutral standpoint from which to judge such difference, and 2) we would not be able to detect a lack of intertranslatability between schemes so long as we assume that we are in fact dealing with language users (or beings capable of genuine expression). I think he is right, and I think this applies also to the notion of the LLM's synthesis of a separate "world within." But what I take from this argument is neither that LLMs are not really generating language (I agree with Fazi that they are, and that synthesis is the process by which they do so), nor that LLMs lack alterity (clearly, they do not have embodied senses like we do). Rather, the encounter with Davidson's argument will help us to understand how this alterity actually fits within a single shared world, but not without transformative effects for that world and our relation to it.

16 A key move in Davidson's argument is to stipulate that "[w]e may identify conceptual schemes with languages, then, or better, allowing for the possibility that more than one language may express the same scheme, sets of intertranslatable languages" (Davidson 1984, 185). The test, then, is to determine whether, and under what circumstances, we could identify a failure of translation—not just a punctual failure but a global one, or one catastrophic enough so that the languages in question could not, on the whole, be said to be intertranslatable. According to Davidson, to imagine such a scenario often involves us in "a dualism of total scheme (or language) and uninterpreted content" (Davidson 1984, 187)—hence a distinction between a kind of phenomenal-conceptual overlay and a noumenal reality that it does not touch. But if that is so, how could we know that two conceptual schemes were incompatible with one another? Significantly, one of the things that LLMs are particularly good at is translation. Fazi emphasizes this fact in arguing for the models' self-sufficient enclosure: "While a large language model is short of the kind of grounding that derives from direct interaction with an exteriority, it nonetheless maintains plenty of linkages with internal statistical patterns inherent in the use of language. For instance, a large language model can connect the English term 'sea' with the Italian word 'mare,' although obviously, it will not relate those words to childhood memories of summers by the Mediterranean Sea. The language a large language model produces is thus still distinctively relational while implying a cut from anything that is other to itself" (Fazi 2024, 49). This sounds exactly like the dualism of total scheme and an inaccessible outside.

17 But how are we to characterize the relation between the LLM's faithful translation between English and Italian and that which is left out (in this case, childhood memories)? This is the site of alleged incommensurability, where translation fails: not between languages, but between the linguistic-rendered-mathematical and the empirical-qua-sensorially-embodied. In a sense, we might say that a cut is made precisely between language's analytic dimensions (those that are computable) and its more robustly empirical content (which resists computation). On the other hand, though, if an LLM were trained on sufficient data, including poetry and literary texts from the Mediterranean, for example, or perhaps even fine-tuned and personalized via texts written

by people who grew up along the coast, would the model really be any worse than the vast majority of English- and/or Italian-speakers at "associating" the words "sea" or "mare" with childhood memories? This is not to say that the computable/mathematical vs. incomputable/sensorial split is irrelevant here. Indeed, it might be of central relevance.

18 While not, of course, talking about LLMs, Davidson explores how distinctions not unlike those being considered here are involved in attempts to make sense of, or to establish, conceptual relativism. On the one hand, he entertains the possibility that one might attack the analytic-synthetic distinction itself, which would be "to give up the idea that we can clearly distinguish between theory and language" (Davidson 1984, 187)—and hence to advance the idea that since all language is theoretical or theory-laden (i.e. non-neutral with respect to conceptual meaning), then linguistic changes are pivots of conceptual-theoretical shifts that are global in nature, that pertain to differences at the scale of worldviews. However, it becomes difficult to know whether speakers on either side of such a split, e.g. before and after a paradigm change, might not simply be using new words with old meanings (or vice versa). In a preeminently pragmatic (and anti-skeptical) gesture, Davidson suggests: "Instead of living in different worlds," these speakers may, "like those who need Webster's dictionary, be only words apart" (Davidson 1984, 189). Grabbing the bull by the other horn, Davidson continues: "The analytic-synthetic distinction is however explained in terms of something that may serve to buttress conceptual relativism, namely the idea of empirical content" (Davidson 1984, 189). That is, if we drop analyticity, or the idea that some truths are justified independent of any relation to the external world, and if we instead hold that "*all* sentences have empirical content" (Davidson 1984, 189), then "we get the dualism of conceptual scheme and empirical content. The new dualism is the foundation of an empiricism shorn of the untenable dogmas of the analytic-synthetic distinction and reductionism—shorn, that is, of the unworkable idea that we can uniquely allocate empirical content sentence by sentence" (Davidson 1984, 189).

19 Significantly, LLMs dispense, as a fundamental condition of their operation, with the notion of a "sentence by sentence" allocation of empirical content—or a sentence by sentence allocation of *any* content for that matter. For one thing, this is because the sentence is not the operative unit. LLMs don't approach meaning as we do; they don't revise their estimation of a sentence's significance based on an unexpected context. (For example, I might take "the cat is on the mat" as a linguistic placeholder in a philosophical discussion about language, but my philosopher friend points toward the door where I see an actual cat on a mat, or maybe just a picture of a cat printed on the mat, or perhaps a cat on top of our friend Matt's head—in each case, the intrusion of empirical content radically revises the meaning). LLMs don't have access to such data, however. This is why they are said to be lacking in reference to the external world (and hence either incapable of genuine meaning or, following Fazi, capable in a modified way—with reference to "a world" rather than "the world"). But, and this is the crucial point, large language models' operation also depends crucially on a lack of distinction made between the empirical and the non-empirical: because they operate subsymbolically, an allegedly analytic truth is no more or less probable for the LLM than any synthetic/empirical statement. The latter indistinction sug-

gests a kind of empiricism without the need for an external world, and perhaps this is a good way of reading Fazi's claim about "a world." The synthetic totality of an LLM's "world," which encompasses both logical relations and (what, from our point of view, can only be described as) empirical claims about the world, is indiscriminately treated as empirical data. Could we say, then, that these algorithmic systems, which treat all meaning as a statistical probability, and never as a logical necessity, actually embody a material proof of the argument against the two aforementioned dogmas, the analytic-synthetic distinction and reductionism? In the LLM, a paradoxical pan-empiricism without external reference is conjoined with a holism that militates against any atomistic or otherwise punctual reductionism—the model's synthesizing scheme organizes and operates on an indiscriminate mass of linguistic/computational data, which serves as the "stuff" of its artificial empiricism.

20 But, according to Davidson, "this second dualism of scheme and content, of organizing system and something waiting to be organized, cannot be made intelligible and defensible. It is itself a dogma of empiricism, the third dogma. The third, and perhaps the last, for if we give it up it is not clear that there is anything distinctive left to call empiricism" (Davidson 1984, 189). And perhaps the "proof" that LLMs deliver to us is not a material refutation of the first two dogmas but simply the following: these models' "artificial empiricism" is not just paradoxical in the ways I have outlined above, but in fact contradictory. That is, the very idea of a self-sufficient world and the conceptual relativism it entails collapses as the idea of empiricism folds in on itself. Davidson outlines what he takes to be the three essential ingredients of this would be empirical relativism: 1) "language as the organizing force," 2) the content or "what is organized," and 3) "finally, the failure of intertranslatability" (Davidson 1984, 190). "The idea then is that something is a language, and associated with a conceptual scheme, whether we can translate it or not, if it stands in a certain relation (predicting, organizing, facing, or fitting) experience (nature, reality, sensory promptings). The problem is to say what the relation is, and to be clearer about the entities related" (Davidson 1984, 191). What Fazi calls the "self-sufficiency" of the LLM's "synthetic generativity" (Fazi 2024, 49), and which I have just elaborated as the lack of distinction between an internal framework and external data, points to difficulties in saying what the relation is supposed to be between the first two elements. But it is in terms of the third element, the failure of intertranslatability (which is to say: the difficulty of demonstrating such a failure) that the LLM both perpetuates familiar problems of conceptual relativism and introduces new ones of its own.

21 In terms of familiar problems: As Davidson argues, we have to assume a common ontology between languages/conceptual schemes in order to identify local breakdowns (Davidson 1984, 192). And this is certainly true for our everyday dealings with LLMs: Mistakes, up to and including "hallucinations," can be identified as such because we take the bulk of the models' output to be composed of genuine language, which is to say language that individuates largely the same objects that are posited in my own language (and linguistically inflected experience of the world). Of course, the appropriateness of this assumption—that LLMs synthesize real and commensurable language—is precisely what is at issue here, so it won't do simply to appeal to the assumptions made in the course of what might be more or less unreflective interactions with

these systems. Besides, the assumption of commensurability flies in the face of what we know about the mismatch between our own symbolic use of language and the computational models' manipulation of subsymbolic vectors. In order not to beg the question, we thus need to turn to the matter of how we might identify a global incommensurability.

22 The (familiar) problem here is that recognizing such a total breakdown would require "a criterion of languagehood that did not depend on, or entail, translatability into a familiar idiom" (Davidson 1984, 192). In the absence of such a criterion, how could we know, or why would we assume, that we were dealing with language in the first place? An incommensurability might plausibly be assumed between the symbolic use of natural language (in *the* world) and the subsymbolic processing of linguistic data (in the LLM's "world within"), but it is unclear that this is an incommensurability of the relevant sort, for the simple reason that it is unclear that there could ever, in principle, be a translation between these. And if there could be no translation, then there is little reason to assume that the LLM has its own conceptual scheme (or model of a world). Of course, there can be a transformation, or perhaps transposition, from the subsymbolic to the symbolic level (and vice versa)—and such movements are precisely those operations that enable the training of models, to begin with, as well as our interactions with them later. However, these are not acts of translation, but rather: synthesis—and they are not interchangeable. Subsymbolic synthesis requires that natural language, whether as training data or as a prompt, be appropriated and assimilated into the computational system before new outputs can be computed. My own synthesis of linguistic meaning when I read the machine's output is completely divorced from the underlying system; I apprehend it, as I said above, as language that is (in the main) commensurable with my own. There is simply no translation between the two levels, but neither can I assert a failure of translation—because I have no reason to impute to the computer the use of language (except in the strictly extra-linguistic sense of "using" language as computational data, which, from the point of view of a language user, is precisely a *misuse* of it). I no more expect a computer—a hardware device designed for the extremely fast processing of electrical voltage differentials and, on that basis, algorithms—to be proficient in the use of language than I do a human brain—whose electrochemical activity is certainly essential to, while wholly distinct from, my ability to synthesize linguistic meaning.

23 We are getting at a fundamental issue here. Mutual understanding and translatability requires (the supposition of) commensurability in terms of what Wittgenstein called *Lebensformen*, or "forms of life" (Wittgenstein 2009, 11e), (Wittgenstein 2009, 94e), (Wittgenstein 2009, 238e). This is apparently not given with LLMs, which lack embodied sensation and direct reference to the outside world. Recall Wittgenstein's aphorism: "If a lion could talk, we wouldn't be able to understand it" (Wittgenstein 2009, 235e). The mystery here is that we apparently *are* able to understand the LLM's linguistic output, despite the system's radical alterity. This is not an easy problem to solve, and the difficulty is what gives rise to claims that the LLM is not actually generating language, or that it is but is encapsulated in its own distinct world. As I have stated, I don't think that either one of these solutions is correct, and I think that each of them is the result of emphasizing an encounter between a human subject and a radically alterior computa-

tional system, an encounter between a speaker of words and calculator of vectors. This is an encounter in which communication could never happen—but also an encounter that, practically speaking, never actually takes place. That is, our encounters with LLMs are always mediated by an interface, typically one that produces what looks like (and I think is, for the most part) genuine language. Without discounting the underlying alterity, we have to interrogate the encounter at this level, where we experience (or at least assume) the common ground of linguistic meaning—a shaky ground, to be sure, and one that we may be forced to question at every turn as we encounter hallucinations and other odd, uncanny, or unsettling artifacts. But the latter, as local misfires, can only appear as such against the background of (assumed) commonality. Since we have seen, following Davidson, that there is no way to make sense of a global incommensurability, it is not question-begging to credit these assumptions with being, by and large, correct. That is, we have no alternative, if we wish to single out local breakdowns, but to assume a common ontology, as expressed in a common or mutually intertranslatable language. This assumption also seems much simpler than the idea that the LLM's textual outputs simply fail at genuine linguisticality. This is not to suggest that we should place too much faith or trust in the truth-value of individual sentences generated, but simply that we believe our own eyes: *this is language we are dealing with.* To accept the latter is, in a sense, to accept the LLM into the world of significance—to accept an alien being into *our* world. And this brings us back to synthesis.

Back to the World

24 The only way to solve the mystery outlined above is, I suggest, to come to terms with the way the alterity of large language models fits into our own forms of life, and to consider what implications this might have for our own experiential synthesis of the world. It will help to return to Davidson's article and to his considerations of partial failures of translatability, as this will assist in understanding issues of truth, reference, and—perhaps, yes—representation as they pertain to our dealings with LLMs.

25 Having dismissed the idea of global incommensurability, Davidson's turn to the "more modest approach" of partial failure brings with it "the possibility of making changes and contrasts between schemes intelligible by reference to the common part" (Davidson 1984, 195). Accordingly, "[w]hat we need is a theory of translation or interpretation that makes no assumptions about shared meanings, concepts, or beliefs" (Davidson 1984, 195). It becomes clear that there are a number of interdependencies involved in our interpretive practices: speech is only interpretable if we know a lot about the speaker's beliefs, but we can't clearly distinguish said beliefs unless we can understand their speech. Thus, "we must have a theory that simultaneously accounts for attitudes and interprets speech, and which assumes neither" (Davidson 1984, 195). To get a foothold, then, we make certain minimal assumptions about our would-be interlocutor. We assume "the attitude of accepting as true, directed to sentences, as the crucial notion" in radical in-

terpretation, which implies nothing about what the speaker means or believes by holding those sentences true (Davidson 1984, 195-196). From there, "[t]he process is that of constructing a viable theory of belief and meaning from sentences held true" (Davidson 1984, 196). This involves forming and revising hypotheses about what words mean, and what the speaker believes, with respect to context, surroundings, etc., including anything we might know about the speaker's "referent culture," as Fazi put it. My friend is a philosopher, so when she utters "the cat is on the mat," I assume she is trying to make a point about language, but a gesture towards the door, where I discover the cat on the mat or the cat on Matt, for example, causes me to revise this assumption. As Davidson points out, "[s]uch examples emphasize the interpretation of anomalous details against a background of common beliefs and a going method of translation" (Davidson 1984, 196). More generally, "[w]hat matters is this: if all we know is what sentences a speaker holds true, and we cannot assume that his language is our own, then we cannot take even a first step towards interpretation without knowing or assuming a great deal about the speaker's beliefs. Since knowledge of beliefs comes only with the ability to interpret words, the only possibility at the start is to assume general agreement on beliefs" (Davidson 1984, 196).

26 Clearly, LLMs call all of this into question. Apart from instances of the so-called "ELIZA effect," we dispense in our interactions with LLMs with the idea that the model believes anything at all or even holds anything to be true (in any ordinary sense). Nevertheless, we accept that the sentences produced are meaningful and thus capable of being true (or false) even in the absence of belief. That is, belief on the part of the model is bracketed, and yet we still manage to interpret. We might say this is because of the output's grounding in the beliefs of the many speakers and writers whose words serve as the model's basic content or foundation (i.e. the training data, largely scraped from the Internet). There is, I think, some truth in this. But the transformation of the training data is radical: because of the intercession of subsymbolic processing, which as we have seen dissolves language into mathematically computable vectors, the original speakers' beliefs can hardly be said to be connected to the newly generated textual outputs. By emphasizing this point, I am not reverting to an insistence on the (non-)encounter between the speaker of words and the calculator of vectors, as I put it in the previous section. Rather, I am merely emphasizing how it comes about that we find ourselves in a position of interpreting language that lacks the motivation of belief. But even if we do manage to interpret these outputs, I take Davidson's point that interpretation is always connected with a general agreement—not about specific statements or beliefs per se, but about the mass or aggregate of them. And this agreement, both generally and in the case of LLMs, is secured by way of reference to a shared background: reference, in other words, to *the* world.

27 This is, of course, just the beginning of interpretation. Progress depends very much on evaluations of truth and reference with respect to context. "We get a first approximation to a finished theory by assigning to sentences of a speaker conditions of truth that actually obtain (in our own opinion) just when the speaker holds those sentences true. The guiding policy is to do this as far as possible, subject to considerations of simplicity, hunches about the effects of social conditioning, and of course our common-sense, or scientific, knowledge of explicable error" (Davidson

1984, 196). LLMs make us reflect, often and quite explicitly, about these conditions of interpretation, and they confront us repeatedly with questions about the fallibility of our hunches, guesses, and beliefs about explicable error—the latter being a major focus of our interactions with LLMs, which marks a significant shift from most of our everyday communicative interactions. That is, it is only right that we should exercise vigilance in our interactions with LLMs, approaching the truth-value of their outputs with a healthy skepticism, and this to a degree that would prove awkward and even rude in our interactions with a fellow human being. Nevertheless, the fact that interpretation still happens points to the necessary occurrence, in these interactions, of assigning such "conditions of truth" and seeking to verify them by way of a shared world to which the texts (attempt to) refer. As I said above, I think there is a kernel of truth in the notion that the interpretability of LLMs' outputs is grounded in the beliefs of the speakers and writers whose linguistic production was used as training data. Although those beliefs do not transfer to the outputs in any sort of punctual manner (i.e. individual beliefs about individual statements are severed in subsymbolic processing and do not find their way back together in the computed outputs), they do nevertheless inform the aggregate shape of those outputs. They do so only indirectly, but those beliefs—held on the part of embodied speakers whose *Lebensform*, we must assume, is not so different from our own (again: this assumption is necessary for the recognition of them as language users)—serve ultimately to connect LLMs' outputs broadly to this world. Truth and reference cannot be assumed on a sentence by sentence basis—but this, as we have seen, is hardly different from language generally.

28 What *is* different is the underlying synthesis, which occurs by means of "synoptic computing"—an aggregating "seeing all together," as Fazi puts it—rather than by looping through the contextually situated and materially embodied vision (and sensation more broadly) of a speaking subject. Nevertheless, as I have attempted to show, we have no alternative but to see LLMs' linguistic outputs as rooted in and broadly expressive of a form of life commensurate with our own. As synthesizers of language, large language models thus refer, by and large, to the same world we live in. And they do so without the need for a Kantian "I think" attached to their outputs (as we have seen, our interpretation proceeds without the ascription of belief). As I argued earlier, Kant's transcendental unity of apperception, which served as a model for Fazi's "computational search for unity" and hence as a basis for understanding computational synthesis, brought with it problems related to the conception and functioning of representation in LLMs. In place of that model, I recommended a pre-personal and pre-reflective model of presencing [*Gegenwärtigung*] (as suggested by Husserl and picked up by Sartre) as more suitable for a subjectless form of synthesis—a model that is itself rooted in human experience and thus ensures at least a formal commensurability between human and nonhuman synthesis. If, at that time, I suggested giving up the search for unity in favor of the openness of ego-less presencing, then the discovery of LLMs' grounded aggregate reference to this world reestablishes the possibility of unity—not on the model of a Kantian subject or a closed world, but that of the common world itself. Significantly, however, this distributed unity is itself open and disunified in an important sense: it is not focused subjectively and does not limn the borders of a conceptual scheme;

rather, it is the socially shared and technically distributed background for synthesis itself. This is to say that human syntheses have never been so neatly encapsulated as Kant's model would suggest, and neither should we expect computational ones to be. If we are "structures that structure," as Fazi puts it, then we are so on the basis of a shared, nonlocalized network of distributed meanings and agencies. Nietzsche's thought, typed on a primitive typewriter—that "our writing utensils collaborate in the formation of our thoughts" (*Unser Schreibzeug arbeitet mit an unseren Gedanken*)—never rang truer than it does in an age of LLMs (Nietzsche 2002, 18). For us and for any other synthesizers of meaning, the world itself provides the requisite unity—itself disunified and distributed, and hence in need of synthesis, across speakers, communities, and even technologies. And this because, despite the dispersal or de-concentration of significance, there is only one world.

29 My inclusion of LLMs and other technical agencies in this shared world may be foreign to Davidson, but it accords largely with what he has elsewhere called his "rejection of subjectivist theories of epistemology and meaning, and [his] conviction that thought itself is essentially social" (Davidson 1993, 608). It thus contributes to the dismantling of the third dogma of empiricism, which is precisely what is needed to reestablish contact with the world: "In giving up the dualism of scheme and world, we do not give up the world, but re-establish unmediated touch with the familiar objects whose antics make our sentences and opinions true or false" (Davidson 1984, 198). In accordance with the above, I will simply add that "unmediated" cannot mean "without media," and the new medium of generative AI, with its own characteristic antics, demonstrates this amply: not just because it presents a new medium through which we can come to know things about the world, but, more fundamentally, because it shows that knowledge—and even the faculties though which we come to know—can be transformatively exteriorized. I learn about the world through AI because the model synthesizes (more or less reliably) information about the same world—information that conforms, fallibly but in the aggregate, to the modes of sensing and knowing that characterize my own form of life. This is because the LLM depends on humans as its own exteriorized senses, as communicated (which is to say: exteriorized) through language (and, with multimodal models, through sound and image as well). We as humans have always been in the game of exteriorization as a basic condition of our contact with and involvement in the world (cf. Stiegler 1998). AI continues this game rather than introducing it. In its synthesis of language—genuine language with indirectly grounded reference to the world—it demonstrates mediation as a condition of having a world in the first place.

30 As we have seen, lest we fall into solipsism or conclude that communication is impossible, we have to assume that all human language users exist in the same mediated world. And unless we give up on the idea that the texts produced by LLMs can be intelligible, we have to conclude that these models too exist in and synthesize this same world. So now that we have reestablished contact with the world, we can at long last reconsider representation, which we have bracketed up until to now. The foregoing reflections on the inherently mediated world point back to Fazi's notion of the "togetherness of distributed representations" (Fazi 2024, 47), which she associated with an LLM's "world within," but which we can now reinterpret as a synthetic unity of repre-

sentations mediately distributed across sites "within the world"—including, among others, the linguistic expressions (and their speakers) that served to train the model, the linguistic outputs of the model (and of those who prompted their creation), and the model itself (which mediates, by way of synthesizing, in order to participate in this worldly network). Whereas Fazi identified the "speculative relevance" of synthesis with regards to "the prospect that this representational reality, internal to the large language model, could be a stable (if imperfect) whole" (Fazi 2024, 47), my suggestion is instead to associate synthesis with distribution and exteriority—it is the means by which disparate sources of significance and referentiality are focused and fused into coherent worldly agencies, and vice versa: whereby those agencies exteriorize themselves, producing the networks and relations that constitute the world. Think, for example, of the way that, for Heidegger, the self can dissolve into an absorptive practice or, given the right prompting, consolidate itself into a reflective agency, suddenly aware of the referential order in and through which its worldly projects were inscribed and oriented (Heidegger 1962). As a condition of worldhood, representation is and has always been distributed; it is bound up with mediation and exteriorization as a condition of our ecstatic mode of existence. In a Stieglerian vein, exteriorization is re-presentation itself, which is to say: a memorialization or record of presencing past and, through it, the future possibility of a presencing deferred (cf. Stiegler 2011). By means of their synthetic displacements, transformations, and syntheses of expressive language, LLMs are obviously involved in these circuits of representational worldmaking.

31 One last word, though, lest there be any misunderstanding. It is not my intention to put large language models on a par with the synthetic agencies of humans. I would suggest that the indirectness of LLMs' grounding in the world—by means of their collating, compressing, and rendering computable the otherwise (mediately) unmediated referentiality of embodied expression—significantly differentiates, even as it partially commensurabilizes, human and computational forms of synthesis. I agree with Fazi that it will not do to write these models off as "stochastic parrots"—a phrase designed to deny their synthetic power altogether in a way that, as Fazi rightly remarks, "might be a little bit unfair toward birds" (Fazi 2024, 43) as well as toward machines. But just as I wish to de-emphasize the role of individual human subjects and challenge the claims they might make with regard to mastery and control over significance and representation, so too would I warn against inflating the agential scope of generative AI in shaping the world. Humans and machines are not on equal footing, and this is not a flat ontology. Though problematic in many respects, I offer in closing Heidegger's rather infamous theses on involvement in the world: "[1.] the stone (material object) is *worldless*; [2.] the animal is *poor in world*; [3.] man is *world-forming*" (Heidegger 1995, 177). The speciesist poverty that Heidegger ascribes to animals is probably also at least a little bit unfair toward birds, but it might be just right for machines.[3]

3. Without explicitly arguing that LLMs are "poor in world," David Bates has also recently brought Heidegger's dictum into proximity with AI (Bates 2024).

Conclusion

32 Finally, then, we can affirm the strong, revisionary sense that Fazi proposes for the term "synthetic media." As a matter of practice, it hardly makes sense to question whether LLMs' outputs constitute "real" language. As a result, we have to take seriously the claim that LLMs are involved in processes of synthesis not unlike the organizational or compositional operations that characterize human cognition and understanding. And while this raises a number of difficult questions, such as how we account for LLMs' material alterity without exaggerating their symbolic incommensurability, it also helps us see how generative AI operates as a transformative force within the distributed networks that constitute and mediate worldhood in an existential sense. The reorientation that Fazi proposes, away from synthesis as contrivance or simulation and towards synthesis as composition, therefore constitutes an important and foundational step towards a reckoning with the aesthetics and politics of artifice in its deepest sense—the making of the world.

References

Bates, David W. 2024. *An Artificial History of Natural Intelligence: Thinking with Machines from Descartes to the Digital Age*. Chicago: University of Chicago Press.

Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜." In *FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–23.

Bender, Emily M., and Alexander Koller. 2020. "Climbing Towards NLU: On Meaning, Form, and Understanding in the Age of Data." In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL 2020)*, edited by Dan Jurafsky, Joyce Chai, Natalie Schluter, and Joel Tetreault, 5185–98. Stroudsburg, PA: Association for Computational Linguistics.

Davidson, Donald. 1984. "On the Very Idea of a Conceptual Scheme." In *Inquiries into Truth and Interpretation*, 183–98. Oxford: Clarendon Press.

———. 1993. "The Third Man." *Critical Inquiry* 19 (4): 607–16.

Fazi, M. Beatrice. 2024. "The Computational Search for Unity: Synthesis in Generative AI." *Journal of Continental Philosophy* 5 (1): 31–56.

Heidegger, Martin. 1962. *Being and Time*. New York: Harper and Row.

———. 1995. *The Fundamental Concepts of Metaphysics: World, Finitude, Solitude*. Bloomington: Indiana University Press.

Husserl, Edmund. 1964. *The Phenomenology of Internal Time-Consciousness*. Bloomington: Indiana University Press.

———. 2012. *Ideas: General Introduction to Pure Phenomenology*. New York: Routledge.

Kant, Immanuel. 1998. *Critique of Pure Reason*. Cambridge: Cambridge University Press.

Kirschenbaum, Matthew. 2003. "Again Theory: A Forum on Language, Meaning, and Intent in the Time of Stochastic Parrots." *In the Moment* (blog). June 26, 2003. https://critinq.wordpress.com/2023/06/26/again-theory-a-forum-on-language-meaning-and-intent-in-the-time-of-stochastic-parrots/.

Knapp, Steven, and Walter Benn Michaels. 1982. "Against Theory." *Critical Inquiry* 8:723–42.

Kuhn, Thomas S. 1962. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Nietzsche, Friedrich. 2002. *Schreibmaschinentexte*. Weimar: Bauhaus-Universitätsverlag.

Sartre, Jean-Paul. 2004. *The Transcendence of the Ego: A Sketch for a Phenomenological Description*. London: Routledge.

Simondon, Gilbert. 2020. *Individuation in Light of Notions of Form and Information*. Translated by Taylor Adkins. Minneapolis: University of Minnesota Press.

Stiegler, Bernard. 1998. *Technics and Time: The Fault of Epimetheus*. Translated by Richard Beardsworth. Vol. 1. Stanford: Stanford University Press.

———. 2011. *Technics and Time: Cinematic Time and the Question of Malaise*. Translated by Stephen Barker. Vol. 3. Stanford: Stanford University Press.

Wittgenstein, Ludwig. 2009. *Philosophical Investigations*. 4th ed. Oxford: Wiley-Blackwell.